

## **KEEL: МНОГОАГЕНТНАЯ ПЛАТФОРМА ДЛЯ СТРУКТУРИРОВАНИЯ ОБРАЗОВАТЕЛЬНОГО КОНТЕНТА С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ**

**Константинов А.А.** <sup>1</sup>

**Научный руководитель – канд. техн. наук Авдюшина А. Е.**<sup>1</sup>

<sup>1</sup>Университет ИТМО  
aakonstantinov@itmo.ru

### **Введение**

В современном образовательном процессе преподаватели сталкиваются с проблемой организации разрозненных учебных материалов различных форматов (PDF, DOCX, PPTX, TXT) в единую структурированную систему. Существующие AI-решения для образования в основном ориентированы на студентов или требуют глубокого анализа полного текста документов, что приводит к высоким вычислительным затратам и потенциальным ошибкам. В данной работе представлена платформа KEEL (Knowledge Engineering for Educators Layer) – многоагентная система, которая использует возможности больших языковых моделей (LLM) и технологии Retrieval-Augmented Generation (RAG) для автоматического структурирования образовательных материалов преподавателей.

### **Основная часть**

KEEL реализует инновационный подход к организации учебного контента, основанный на использовании структурных сигналов, уже присутствующих в документах преподавателей (заголовки, разделы, иерархии). Система состоит из четырех ключевых компонентов:

1. **Модуль обработки документов:** конвертирует разнородные форматы в единый Markdown-формат, сохраняя структурную информацию;
2. **Агент структурирования (Structure Agent):** использует LLM для классификации документов по образовательным функциям (лекции, лабораторные работы, организационные документы) и построения иерархической карты знаний на основе извлеченных заголовков;
3. **Система семантического индексирования:** создает векторную базу данных с использованием многоязычных моделей эмбедингов и библиотеки FAISS для эффективного поиска;
4. **RAG-система:** обеспечивает контекстно-зависимые ответы на естественно-языковые запросы, основанные на реальных учебных материалах.

Ключевое преимущество KEEL заключается в минималистичном подходе: вместо анализа полного текста система фокусируется на структурных элементах документов, что значительно снижает токентные затраты и повышает точность результатов. Экспериментальная оценка на реальных университетских курсах показала способность системы надежно извлекать педагогически значимую структуру даже из неоднородно отформатированных файлов.

## **Выводы**

Платформа KEEL представляет эффективное решение для автоматизации организации образовательного контента, предлагая преподавателям инструмент для трансформации разрозненных материалов в структурированные, интерактивные и доступные для поиска в пространстве. Система демонстрирует высокую производительность при работе с многоязычными материалами и различными форматами документов, обеспечивая прозрачность и модульность архитектуры.

## **Литература**

1. Wang S., Xu T. et al. Large Language Models for Education: A Survey and Outlook // arXiv preprint arXiv:2403.18105. – 2024.
2. Lewis P., Perez E. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // arXiv preprint arXiv:2005.11401. – 2020.
3. Chen P., Lu Y., Zheng V. W., Pian Y. KnowEdu: A system to construct knowledge graph for education // IEEE Access. – 2018. – Vol. 6. – P. 31553–31563.
4. Sun X., Li J., Wang Y. LLM4EduKG: Large language models for automatic construction of educational knowledge graphs // arXiv preprint arXiv:2403.08345. – 2024.