

## **АВТОМАТИЗАЦИЯ ДОСТУПА К УЧЕБНЫМ МАТЕРИАЛАМ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОЙ МОДЕЛИ В ОБРАЗОВАТЕЛЬНОМ ЧАТ-БОТЕ: АРХИТЕКТУРА И РЕАЛИЗАЦИЯ**

**Сарычева А.А.<sup>1</sup>**

**Научный руководитель – кандидат философских наук, доцент Блейхер О.В.<sup>1</sup>**

<sup>1</sup>Университет ИТМО

alex.sarychewa2014@yandex.ru

### **Введение**

Современная цифровизация высшего образования ставит задачу преодоления ограничений традиционных платформ управления обучением (LMS). Существующие LMS, такие как Moodle, предлагают линейную навигацию и требуют значительных усилий для ручной адаптации контента [1], что критично для дисциплин со сложной иерархической структурой, например, высшей математики, тем самым снижая эффективность адаптации учебного процесса [2, 3]. Одновременно с этим попытки внедрения больших языковых моделей (LLM) в образование выявили их ключевой недостаток - склонность к генерации методологически некорректных ответов ("галлюцинаций"), что недопустимо в академической среде. Анализ материалов по данной проблеме демонстрирует эффективность гибридного подхода, заключающегося в интеграции LLM с внешними структурированными источниками информации в рамках парадигмы Retrieval-Augmented Generation (RAG). Такой подход позволяет повысить точность и достоверность ответов [4, 5]. Однако большинство решений ограничивается векторным поиском по тексту, не учитывая логические взаимосвязи между понятиями [6]. Таким образом, существует противоречие между потребностью в гибком естественно-языковом интерфейсе и необходимостью обеспечения достоверности учебного контента и логической связности материала.

### **Основная часть**

В качестве решения предлагается архитектура образовательного чат-бота, основанная на объединении графа знаний [7] и LLM в рамках подхода GraphRAG. Ядром системы выступает графовая база данных Neo4j, которая хранит семантическую сеть из дидактических единиц: определений, теорем, примеров, замечаний и логических связей между ними (например, "следует из", "определяется в", "используется в"). Применение графовых баз данных в образовании для структурирования учебных материалов обосновано в ряде исследований [8, 9], а необходимость персонализации структуры учебного контента, выявленная при анализе ограничений LMS [1], подтверждает актуальность такого подхода. Данная онтология позволит моделировать структуру рассматриваемого модуля "Комплексные числа" с высокой степенью связности.

Ключевой особенностью разработки является механизм трансляции запроса. Когда студент задает вопрос на естественном языке через Telegram-интерфейс, NLP-модуль на базе LangChain не выполняет поиск по тексту, а генерирует структурированный Cypher-запрос. LLM выступает в роли "ретранслятора" с человеческого языка на язык запросов в графовой базе. Полученный Cypher-запрос исполняется в Neo4j, возвращая точный подграф знаний, релевантный вопросу. Такой подход гарантирует, что ответ основан исключительно на верифицированных учебных материалах, загруженных преподавателем, что нивелирует проблему "галлюцинирования". Эффективность подобных гибридных систем с использованием Neo4j подтверждается современными работами в области GraphRAG [6].

Дополнительно архитектура включает модуль персонализации (overlay-графы), который накладывает на исходную структуру графа знаний информацию о прогрессе конкретного студента и его пожеланий об уровне изучаемого материала, позволяя адаптировать навигацию и рекомендации.

### **Выводы**

Предложенная архитектура была реализована в виде Telegram-бота для модуля "Комплексные числа", в котором студенты получают возможность кроме изучения материалов модуля задавать вопросы в свободной форме и получать структурированные ответы по тематике материала. Таким образом, система показала, что интеграция графа знаний и LLM позволят автоматизировать доступ к учебным материалам, сохраняя методологическую целостность материала. Кроме того разработанное решение масштабируемое, т.к. для смена дисциплины требует только загрузки нового графа знаний без перестройки логики работы NLP-модуля.

### **Литература**

1. Bartuskova A., Krejcar O. Personalization of learning content and its structure, based on limitations of LMSs //Recent Developments in Intelligent Information and Database Systems. – Cham : Springer International Publishing, 2016. – С. 455-465.
2. Абросимова Е. Е. Студенты и цифровизация высшего образования: социологический анализ //Ойкумена. Регионоведческие исследования. – 2023. – №. 4 (67). – С. 103-110.
3. Блейхер О.В., Снегурова В.И. Стратегии интеграции технологий искусственного интеллекта в образовательную практику с учётом профессиональных установок преподавателей // Письма в Эмиссия. Оффлайн: Электронный научный журнал - 2025. - № 8. - С. 3570
4. Lewis P. et al. Retrieval-augmented generation for knowledge-intensive nlp tasks //Advances in neural information processing systems. – 2020. – Т. 33. – С. 9459-9474.
5. Li Z. et al. Retrieval-augmented generation for educational application: A systematic survey //Computers and Education: Artificial Intelligence. – 2025. – Т. 8. – С. 100417.
6. Ghaffari E., Ibrahimi D. Evaluating Graph Retrieval-Augmented Generation with Neo4j and FAISS: A Performance Comparison in Hybrid Retrieval Systems. – 2025.
7. Hogan A. et al. Knowledge graphs //ACM Computing Surveys (Csur). – 2021. – Т. 54. – №. 4. – С. 1-37.
8. Yu X. et al. Design and implementation of curriculum system based on knowledge graph //2021 IEEE international conference on consumer electronics and computer engineering (ICCECE). – IEEE, 2021. – С. 767-770.
9. Guo Y., Song Y. Development of an Intelligent Educational Q&A System based on Knowledge Graph and Natural Language Processing //2025 2nd International Conference on Software, Systems and Information Technology (SSITCON). – IEEE, 2025. – С. 1-6.