

гибридный подход с интеграцией реальных данных, обеспечивающий необходимый уровень достоверности и интерактивности.

Литература

1. Процедурная генерация / Wikipedia. – URL: https://ru.wikipedia.org/wiki/Процедурная_генерация (дата обращения: 09.03.2026).
2. Малышев, И. А. Методы динамической процедурной генерации ландшафта / И. А. Малышев, К. А. Кивва // Вектор научной мысли. – 2025. – № 5 (22).
3. Диаграмма Вороного и ее применения. – URL: <https://habr.com/ru/articles/309252/> (дата обращения: 30.03.2026).
4. Симуляция эрозии в 3D-рельефе. – URL: <https://habr.com/ru/articles/533990/> (дата обращения: 30.03.2026).
5. Андриевский, Т. Э. Некоторые вопросы генерации реалистичных ландшафтов / Т. Э. Андриевский, Б. Г. Мирвелов // Научный Лидер. – 2022. – № 20 (65).
6. Цветков, В. Я. Технология генерации виртуальных ландшафтов с применением метода вокселизации и волюметрических данных / В. Я. Цветков, В. А. Мордвинов, В. Т. Матчин // Образовательные ресурсы и технологии. – 2023. – № 1 (42). – С. 91–99.

ПРОГРАММНОЕ СРЕДСТВО АВТОМАТИЗИРОВАННОГО СОЗДАНИЯ УЧЕБНЫХ МАТЕРИАЛОВ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Д. А. Шаблинский

Учреждение образования «Гомельский государственный технический
университет имени П. О. Сухого», Республика Беларусь

Научный руководитель О. Д. Асенчик

Представлено программное средство для автоматизированного формирования учебных материалов на основе больших языковых моделей с интеграцией извлеченных знаний по технологии Retrieval-Augmented Generation. Описаны архитектура оркестрации в виде графа состояний, многоуровневая верификация ответов и результаты апробации. В ходе экспериментов по дисциплине «Нереляционные базы данных» подтверждено преимущество атомарной декомпозиции задач над монолитной генерацией.

Ключевые слова: большие языковые модели, Retrieval-Augmented Generation, RAG, автоматизация образовательного контента, многоэтапная верификация, рекурсивная генерация, семантические метрики.

Современные образовательные процессы требуют регулярного обновления и расширения учебных материалов. Развитие больших языковых моделей (LLM) и подхода генерации, дополненной извлечением (Retrieval-Augmented Generation, RAG), открывает возможности для автоматизации генерации образовательного контента [1].

Однако прямое применение LLM может приводить к галлюцинациям, несогласованности разделов и отсутствию встроенных механизмов верификации. В работе предложено программное средство, минимизирующее эти риски за счет многоуровневой оркестрации, интеграции внешних знаний через RAG и рекурсивного структурирования материалов.

Научная новизна заключается в методике адаптивной рекурсивной генерации, направленной на сохранение семантической последовательности и снижение дублирования определений в смежных главах. Разработанная архитектура на базе ориентированных графов состояний позволяет уменьшить проявления галлюцинаций за счет контура обратной связи и автоматизированной кросс-верификации знаний.

Система построена по модульному принципу и включает веб-интерфейс, серверную часть на асинхронном фреймворке, подсистему RAG, оркестратор процессов и движок рекурсивной генерации структурированных пособий.

Конвейер генерации ответа реализован как ориентированный граф состояний (с использованием библиотеки LangGraph), состоящий из пяти узлов: подготовка запроса, генерация, оценка качества, доработка и финализация. Переход к этапу доработки осуществляется условно (при оценке ниже 7,0 балла), а цикл ограничен тремя итерациями для предотвращения заикливания.

Графический интерфейс реализован как веб-приложение с тремя функциональными модулями: «Генерация», «База знаний RAG» и «Управление моделями». Вкладка генерации позволяет выбирать модель, настраивать тип верификации, вводить вопросы или загружать файлы, а также отслеживать прогресс выполнения задач в реальном времени с возможностью экспорта результатов в форматы Word/TXT. Модуль RAG обеспечивает загрузку и индексацию источников (PDF, DOCX, TXT, MD), просмотр статистики и очистку базы. Вкладка управления моделями позволяет динамически настраивать провайдеров и параметры LLM через интерактивные формы.

Подсистема RAG обеспечивает семантическое извлечение знаний из эталонных источников различных форматов. Процесс включает извлечение текста, предобработку, разбиение на фрагменты размером 800 символов с перекрытием 100 символов, векторизацию моделью sentence-transformers/all-MiniLM-L6-v2 (выбор обусловлен балансом качества и вычислительной эффективности для русскоязычных текстов) и хранение в векторной базе с метрикой косинусного сходства. Поиск контекста выполняется с пороговой фильтрацией релевантности.

Верификация реализована на двух уровнях:

- экспертная оценка языковой моделью (LLM-as-a-judge) по критериям точности, полноты, терминологии и наличия примеров;
- поисковая проверка: генерация уточняющих вопросов и кросс-проверка по внешним источникам с последующим синтезом результатов [2].

Итоговый индекс достоверности рассчитывается как взвешенная сумма оценок, где приоритет отдается результатам поисковой кросс-проверки по внешним источникам. Такой приоритет уменьшает зависимость итоговой оценки от возможных систематических ошибок, связанных с оцениванием ответа той же языковой моделью.

Для создания полноценных учебных пособий реализована рекурсивная генерация. Процесс включает:

- анализ перечня вопросов и формирование плана структуры пособия;
- рекурсивную генерацию глав с автоматическим разбиением на подзадачи при превышении лимита объема;
- управление терминологией для исключения дублирования объяснений;
- агрегацию результатов и формирование единого документа в формате Markdown.

Рекурсия ограничена двумя уровнями; по результатам апробации такой предел обеспечивает баланс качества текста и вычислительных затрат.

В ходе апробации системы была проведена генерация материалов по дисциплине «Нереляционные базы данных» на основе 22 тематически связанных вопросов. Для оценки эффективности предложенного метода выполнено сравнительное тестирование с контрольной группой, где генерация осуществлялась единым запросом (Single-shot) ко всему перечню вопросов.

Экспериментально подтверждено, что атомарная декомпозиция задач в рамках агентного графа обеспечивает более высокую семантическую точность по сравнению с генерацией единым запросом.

Для количественной оценки использованы семантические метрики: *BERTScore* (*F1*), вычисляемая по сходству контекстных представлений токенов, а также *Embedding Similarity* и *S-BERT Similarity*, определяемые как косинусное сходство векторных представлений текстов. Более высокие значения метрик соответствуют большей близости к эталонному источнику. На рис. 1 представлены результаты сравнительного анализа: метрики *BERTScore* (*F1*) и *Embedding Similarity* составили 0,99, тогда как при генерации единым запросом эти показатели составили 0,67 и 0,69 соответственно. Аналогичный разрыв зафиксирован по метрике *S-BERT Similarity* (рост с 0,72 до 0,99), что согласуется с гипотезой о снижении эффекта потери фокуса (*Lost in the Middle*) при перегрузке контекстного окна модели.

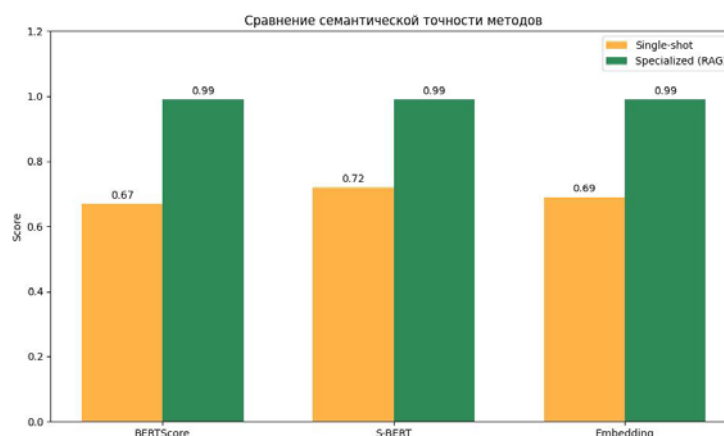


Рис. 1. Сравнительный анализ семантической точности генерации по метрикам

Средняя оценка качества ответов на отдельные вопросы, сгенерированных предложенной системой, составила 9,0 балла по десятибалльной шкале (LLM-as-a-judge по критериям точности, полноты и наличия примеров). Среднее время генерации одного атомарного ответа с полной верификацией составило от 60 до 90 секунд.

На втором этапе полученные материалы передавались в движок рекурсивной структуризации для формирования глав. Разделение процессов генерации и структурирования позволило устранить терминологические повторы и логические разрывы, которые в контрольной группе приводили к избыточности текста на 15–20 %. Благодаря автоматическому планированию содержания и исключению дублирующих объяснений на уровне глав итоговое пособие приобрело четкую иерархическую структуру с плавными межсекционными переходами, что недостижимо при попытке прямой генерации всего объема материала одним запросом.

Разработанное программное средство демонстрирует эффективность технологии RAG [2] и многоэтапной верификации для автоматизации создания учебных материалов. Рекурсивный движок генерации позволяет формировать структурированные пособия из накопленных знаний. Полученные результаты подтверждают практическую ценность подхода и открывают перспективы для дальнейшего развития системы в направлении поддержки полного цикла образовательного процесса.

Л и т е р а т у р а

1. Асенчик, О. Д. Использование больших языковых моделей для создания учебных материалов по дисциплине «Базы данных» / О. Д. Асенчик // Цифровая трансформация. – 2025. – Т. 31, № 4. – С. 5–14.

2. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / P. Lewis [et al.] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 9459–9474
3. Reimers, N. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks / N. Reimers, I. Gurevych // Proceedings of EMNLP. – 2019. – P. 3982–3992.

СЕРВИС ДЛЯ ЦЕНТРАЛИЗОВАННОГО ХРАНЕНИЯ ИЗОБРАЖЕНИЙ, ИНТЕГРИРОВАННЫХ В S3-ХРАНИЛИЩЕ, РЕАЛИЗОВАННЫЙ СРЕДСТВАМИ 1С:ПРЕДПРИЯТИЕ

В. А. Пинчукова

*Учреждение образования «Гомельский государственный технический
университет имени П. О. Сухого», Республика Беларусь*

Научный руководитель Е. В. Комракова

Рассмотрено проектирование централизованного сервиса для управления графическим контентом в рамках крупного предприятия. Исследование направлено на решение проблемы падения производительности систем на базе «1С:Предприятие», вызванной стремительным ростом объема бинарных данных. Предложена гибридная архитектура, в которой платформа выступает в роли шлюза для управления метаданными и API-менеджмента, а хранение файлов делегируется S3-хранилищу MinIO.

Ключевые слова: 1С:Предприятие, S3-хранилище, MinIO, масштабируемость, оптимизация баз данных.

В современных условиях цифровизации системы на базе «1С:Предприятие» вышли далеко за рамки классического бухгалтерского учета, превратившись в комплексные ERP-решения, оперирующие огромными массивами данных. Одной из наиболее острых проблем функционирования таких систем в РУП «Производственное объединение «Белоруснефть» является стремительный рост объема графической информации, такой как фотографии товаров, техническая документация и маркетинговые материалы. Хранение данных непосредственно внутри базы данных 1С неизбежно приводит к ее увеличению, что влечет за собой и увеличение времени работы, а также падению производительности.

Актуальность исследования обусловлена необходимостью перехода от монолитной архитектуры хранения изображений к сервисно-ориентированному подходу, где единый источник изображений должен быть доступен одновременно для множества потребителей: мобильных приложений клиентов, интернет-витрин, внутренних учетных систем и перспективных инновационных разработок. В частности, на предприятии ведется проектирование автоматизированной системы – роборуки, выполняющей функции продавца-консультанта, для которой мгновенный доступ к медиа-контенту является необходимым условием распознавания и позиционирования товаров.

На современном рынке существуют различные решения, предлагающие управление изображениями, но ни одно из них не сочетает в себе все преимущества, которые необходимы предприятию. Облачные хранилища, такие как Cloudinary и Amazon S3, предоставляют масштабируемость, но связаны с затратами на оплату облачных ресурсов, зависят от интернет-соединения и могут быть сложнее в управлении данными и настройке доступов. Системы управления контентом (CMS) с встроенными хранилищами, предлагают встроенные механизмы для управления изображениями, но они часто менее гибкие и масштабируемые, особенно для больших объемов данных и интеграции с ERP-системами.

В «1С:Предприятие» есть два классических подхода к работе с медиа-файлами, но и они имеют существенные ограничения при масштабировании. Первый метод