

при отсутствии дорогостоящего оборудования; повысить наглядность и понимание сложных квантовых процессов; развить у студентов навыки работы с современной измерительной техникой; индивидуализировать обучение за счёт многократного повторения виртуального эксперимента.

1 Годлевская, А. Н. Физика атома и атомных явлений / А. Н. Годлевская, В. Г. Шолох. – Гомель: ГГУ им. Ф. Скорины, 2006. – 213 с.

2 Кислов, А. Н. Атомная и ядерная физика: учеб. пособие / А. Н. Кислов. – Екатеринбург: УрФУ, 2017. – 271 с.

3 Морозов, А. Н. Организация физического практикума в техническом университете / А. Н. Морозов, О. С. Еркович. – М.: МГТУ им. Баумана, 2014. – 14 с.

4 Положение о порядке подготовки, выполнения, оформления и защиты лабораторных работ. – URL: <https://www.bsuir.by> (дата обращения: 25.02.2026. – Текст : электронный.

ТЕХНОЛОГИИ ИНТЕГРАЦИИ LLM В ВЕБ-ОРИЕНТИРОВАННЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Шевчук М.С.,

магистрант 2 курса ВГУ имени П.М. Машерова, г. Витебск, Республика Беларусь

Научный руководитель – Алейникова Т.Г., канд. физ.-мат. наук, доцент

Актуальность темы обусловлена широким внедрением больших языковых моделей в корпоративные и веб-приложения и необходимостью обобщения подходов к их интеграции в информационные системы. При этом важно рассматривать не только алгоритмы моделей, но и архитектурные и интеграционные решения, обеспечивающие сценарии типа RAG, потоковой генерации и многоканального доступа в рамках веб-ориентированных систем.

Целью работы является систематизация технологий и архитектурных решений интеграции LLM в веб-ориентированные системы, а также разработка обобщенной модели их взаимодействия с традиционными компонентами бэкенда. В соответствии с этой целью формулируются задачи: уточнение понятия веб-ориентированной информационной системы в контексте интеграции LLM; классификация типовых сценариев применения LLM в веб-системах; анализ архитектурных ролей компонентов при интеграции LLM; анализ выбора и применения технологий интеграции, формулирование обобщенных рекомендаций и архитектурных паттернов по интеграции LLM в веб-системы.

Материал и методы. Объектом исследования является веб-ориентированная информационная система для работы с документами, реализующая сценарии генерации ответов с опорой на внешние источники данных (RAG). Система выбрана как репрезентативный пример, поскольку включает типовые компоненты современных LLM-интеграций: сервис оркестрации, векторное хранилище, модуль индексации неструктурированных данных и API-шлюз.

Материалами исследования послужили архитектурная документация системы, исходный код модулей интеграции, а также техническая документация подключаемых больших языковых моделей (генеративных, эмбединговых и мультимодальных).

Методика работы включала три этапа. На первом этапе проведен структурный анализ компонентов системы: выделены контуры обработки запросов и потоки данных, связывающие пользовательский интерфейс, механизмы подготовки контекста и внешние LLM. На втором этапе выполнена классификация сценариев использования (асинхронный диалог, потоковая генерация в редакторе, семантический поиск) и сопоставление их с конкретными архитектурными решениями. На третьем этапе, на основе анализа взаимодействий между сервисами, были выделены устойчивые архитектурные связки (паттерны). Для верификации выделенных паттернов проводилось их сравнение с рекомендациями эталонных архитектур и открытыми реализациями аналогичных систем.

Результаты и их обсуждение. В ходе анализа архитектуры системы выделены три сценария использования LLM: диалог с контекстом из активированного набора данных, генерация текста в блоке промпта внутри документа с привязкой к наборам и потоковая

запись ответа в редактор, а также предобработка данных перед индексацией с помощью промпта рафинирования. Оркестрация запросов к модели и подбор контекста сосредоточены в отдельном сервисе, подготовка данных и построение индексов — в сервисе индексации; единая точка входа для веб-клиента и сервис документов обращаются к оркестратору, не вызывая LLM напрямую. Один контур генерации обслуживает веб-интерфейс и иные каналы доступа при общих наборах данных; контекст формируется из предобработанных наборов с использованием эмбедингов и векторного поиска; ответ доставляется потоком, в том числе через очередь сообщений в документ.

Применяемые технологии интеграции сводятся к выделенному сервису оркестрации, RAG на основе индексированных наборов и векторного поиска, потоковой доставке ответа и унифицированному обращению к моделям по единому контракту API. Разделение индексации и генерации создаёт условия для асинхронной обработки данных и независимого масштабирования; выделенные решения применимы при проектировании иных веб-ориентированных систем с интеграцией LLM.

Заключение. В ходе выполнения работы была достигнута поставленная цель: на основе анализа референтной веб-системы проведена систематизация архитектурных решений и технологий интеграции больших языковых моделей (LLM). В отличие от существующих фрагментарных описаний, предложен целостный взгляд на архитектуру, объединяющую сценарии RAG, потоковую генерацию и асинхронную предобработку данных.

Анализ позволил выделить и описать три типовых сценария использования больших языковых моделей в веб-среде: диалог с опорой на предварительно собранные данные, встроенная генерация текста непосредственно в теле документа и подготовка исходных материалов для последующей индексации. Также была обоснована эффективность выделения специализированных составных частей.

Предложенная архитектура включает сервис оркестрации, механизм RAG на основе векторного поиска и унифицированный API-контракт. Обобщены ключевые технологические решения: потоковая доставка ответов, разделение контуров индексации и генерации, многоканальность при едином серверном контуре. Таким образом, предложенный набор архитектурных решений может служить типовой схемой при создании веб-ориентированных информационных систем, встраивающих большие языковые модели.

1. ТЕХНОЛОГИЯ RAG (RETRIEVAL-AUGMENTED GENERATION) КАК ИННОВАЦИОННЫЙ ПОДХОД В LLM / О. А. Науменко. – НИУ «БелГУ», 2025. – 10 с.
2. МЕТОДОЛОГИЯ СОЗДАНИЯ МНОГОСВЯЗНЫХ СТРУКТУР ДАННЫХ С ПРИМЕНЕНИЕМ LLM В РАБОЧИХ ПРОЕКТАХ / М. Ф. Зимнуров, И. А. Астраханцева. – ИГХТУ, 2025. – 8 с.
3. ДООБУЧЕНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ ДЛЯ РЕШЕНИЯ СПЕЦИАЛИЗИРОВАННЫХ ЗАДАЧ / Т. А. Молчанова. – УрФУ, 2024. – 79 с.
4. Причины галлюцинаций в современных языковых моделях и методы их устранения с помощью формализованных моделей знаний / Ю. Д. Лыкова. – БГТУ, 2025. – 5 с.
5. Об одном подходе к разработке информационно-справочных систем на основе больших языковых моделей / С. Е. Попов, В. П. Потапов, Р. Ю. Замараев. – ФИЦ ИВТ, Новосибирск, 2025. – 21 с.
6. Механизмы адаптации генеративных моделей искусственного интеллекта для персонализации взаимодействия с пользователями / А. В. Барейша. – Витебск, 2025. – URL:<https://rep.vsu.by/handle/123456789/48448> (дата обращения: 28.02.2026). – Текст: электронный.

ТРЕХМЕРНАЯ РЕКОНСТРУКЦИЯ ЦЕРЕБРАЛЬНЫХ НОВООБРАЗОВАНИЙ: ЦИФРОВАЯ ОБРАБОТКА МРТ- И КТ-ИЗОБРАЖЕНИЙ

Шуголь А.А.,

студент 4 курса ВГУ имени П.М. Машерова, г. Витебск, Республика Беларусь

Научный руководитель – Исаченко Ю.В., преподаватель

Каждый хирург, готовящийся к сложной операции на головном мозге, сталкивается с необходимостью мысленно воссоздать трехмерную архитектуру органов пациента, опираясь лишь на разрозненные плоские срезы КТ и МРТ. Однако человеческий мозг — это сложнейшая сеть, где опухоль может быть тесно переплетена с критически важными центрами и сосудами. В условиях ограниченного времени и плоской визуализации