цифровой ↑ сезон ии

Q&А-бот с ИИ

Решение кейса для Корпоративного университета **"РЖД"**

Destination team



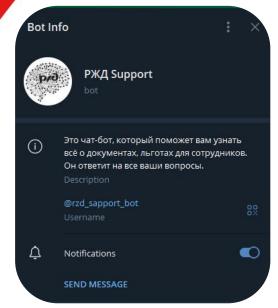




GitHub



Задачей было создание чат-бота с ИИ, формирующего ответы на вопросы пользователя на основании имеющихся в базе данных документов с минимальным привлечением к интерпретации документов, формулировке ответов и поддержанию актуальности базы данных человека.









- 1. Повысить информированность новых сотрудников о доступных льготах и возможностях.
- 2. Оптимизировать работу специалистов по управлению персоналом за счет автоматизации ответов на типовые вопросы.
- 3. Обеспечить актуальность и доступность информации о социальном пакете для сотрудников независимо от их местонахождения.









- 1. Информационная перегрузка новых сотрудников при ознакомлении с большим количеством нормативных документов.
- 2. Перегруженность и отсутствие времени специалистов по управлению персоналом для разъяснения всех возможностей и льгот сотрудникам.
- 3. Сложность постоянной актуализации нормативных документов, регулирующих социальный пакет.









Чат-бот с ИИ для ответа на вопросы, извлекающий информацию из документов, дающий точную информацию по интересующему сотрудника вопросу.
Функции:

- Просмотр частых вопросов пользователей
- Функция вопроса к ИИ
- Вопрос к оператору при возникновении необходимости
- Транскрибация голоса в текст

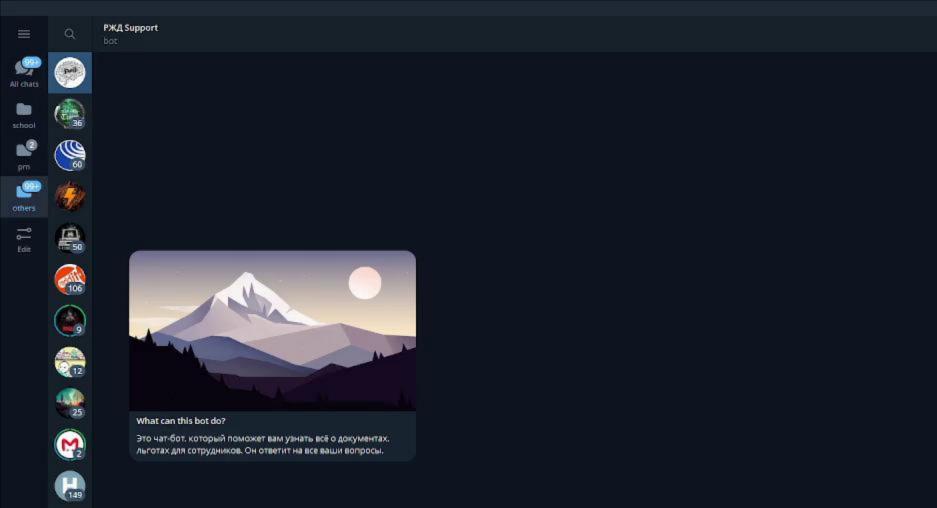




- ИИ: Ollama, qwen2 LLM (быстрее Llama на 30%), USER-bge-m3 LLM
- База данных: MongoDB, Beanie ODM, Qdrant
- Backend: Aiogram, Docker, openai-whisper STT, Ffmpeg, Python 3.12

В качестве платформы для размещения чат-бота используется мессенждер Telegram.







Уникальность предлагаемого решения в том, что оно превосходит аналоги на рынке благодаря:

- 1) использованию оптимизированной высокоточной модели Qwen2 с системным промптом;
- 2) использованию технологий, позволяющих точно и быстро отвечать на вопросы пользователей, с возможностью горизонтального масштабирования;
- 3) расширенным возможностям, включая speech-to-text и техническую поддержку. Комплексный подход и передовые технологии обеспечивают уникальное конкурентное преимущество на рынке.





В дальнейшей доработке проекта, мы бы хотели:

- Заменить Ollama на vLLM для повышения производительности
- Сделать кроссплатформенную систему для бота(например, Viber, WhatsApp, Vk Мессенджер(Vk Боты)
- Более развитая и продвинутая экосистема ручной поддержки клиентов.







Команда проекта









Велиев Тимур Спикер, автор презентации, Backend



Сусин Дмитрий ML-E, Backend



Емельянов Артём DevOps Backend



Колосков Вячеслав ML-E



