

Разработка умного гида для картинных галерей и музеев в формате чат-бота

Промежуточный отчет

Решаемая проблема

В современных музеях услуги гида и аудиогиды обладают значительными ограничениями с точки зрения персонализации и гибкости предоставления информации. Традиционные групповые экскурсии, проводимые гидами, не позволяют посетителям адаптировать маршрут под свои интересы. Аудиогиды обеспечивают индивидуальный формат, однако объем выдаваемой информации строго ограничен заранее записанными материалами и не предусматривает интерактивного взаимодействия и возможности получения справок на уникальные, частные вопросы. Современный посетитель заинтересован в свободном изучении экспозиций, выборе темпа осмотра и возможности получения расширенной информации об объектах в реальном времени.

Что реализовано на данный момент

Реализацию проекта можно разделить на три основные части:

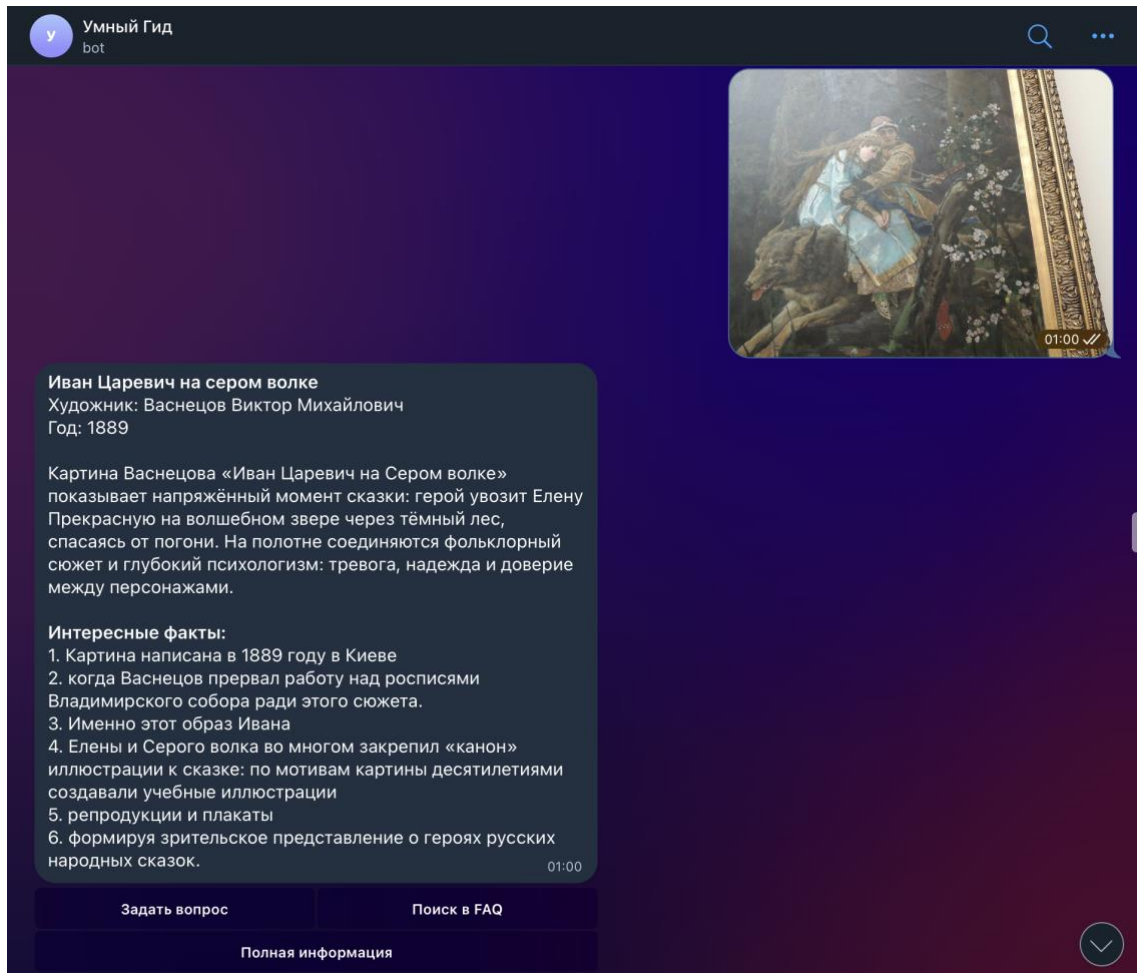
1. Реализация core-функционала
2. Реализация Telegram-бота
3. Наполнение базы знаний картин и вопросов

На данный момент были покрыты первые два пункта:

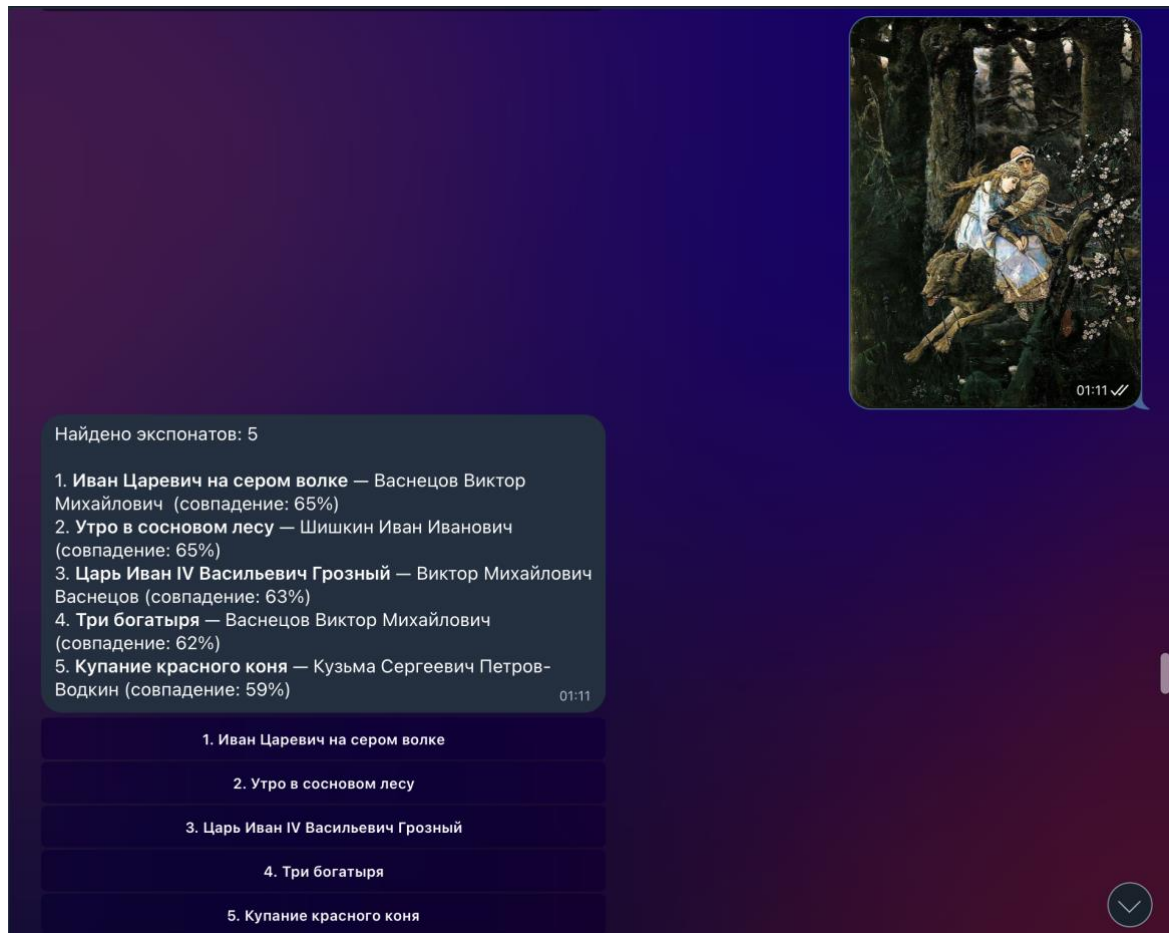
1. Модели
 - a. Реализован функционал Text Encoder
 - b. Реализован функционал Vision Encoder
 - c. VLM (ответ на вопрос по изображению)
2. Схемы данных
3. Реализация векторной базы
 - a. Добавление экспонатов (включая метадату)
 - b. Поиск экспонатов (с учетом метадаты)
 - c. Добавление вопросов в базу FAQ
 - d. Поиск вопросов в базе FAQ
 - e. Получение метадаты экспоната
 - f. Удаление экспоната
4. Основной «агент» – точка входа
 - a. Распознавание экспоната
 - b. Ответить на вопрос по изображению (экспонату)
 - c. Поиск по FAQ
 - d. Поиск экспоната по текстовому описанию
5. Бот
 - a. Реализованы утилиты (загрузка фото, форматирование)
 - b. Реализованы функциональные клавиши
 - c. Обработчики событий
6. Данные

- а. На данном этапе собрана тестовая коллекция экспонатов (10 сэмплов) для отладки функционала

Примеры работы



В тестовом формате реализован выбор из нескольких вариантов



Дальнейшие шаги

1. Необходимо провести дальнейшую отладку бота
 - a. Есть проблемы с инициализацией энкодер моделей (она сейчас происходит при первом вызове)
2. Произвести еще несколько улучшений
 - a. Заменить векторную базу Chroma на более гибкий вариант (например, Qdrant)
 - b. Добавить логику автоматической маршрутизации между базой FAQ и ответом с применением VLM
3. Собрать базы знаний
 - a. Собрать полную базу экспонатов и метаданных к ним
 - b. Собрать базу FAQ по экспонатам