



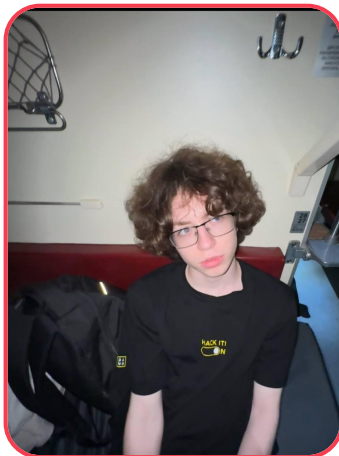
Ассистент отельера для создания описания отеля



ByteBusters



Наша команда



Сибик Георгий
Николаевич, 11 класс,
Тюмень, Team Lead
ML/python dev



Пахомов Роман
Иванович, 10 класс,
Санкт-Петербург,
CTO, ML researcher



Леонов Никита
Артемович, 11 класс,
Московская
область,
ML Engineer

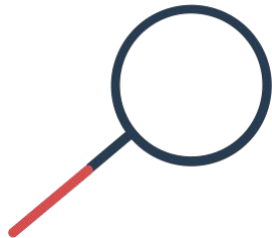


Буян Степан, 10
класс, Москва,
FullStack

Проблематика



Отельеры, публикуя свой отель в онлайн-системах бронирования, часто сталкиваются с трудностями при составлении качественного и привлекательного описания.



Трудности стоящие перед отельерами при составлении качественного описания:

- 1 Учёт всех ключевых параметров
- 2 Выделение уникальности отеля
- 3 Эмоциональная привлекательность
- 4 Грамотность

ByteBusters



Цель проекта

Разработать интеллектуального ассистента, который на основе анализа релевантной информации об отелях из внешних источников будет помогать отельеру сформировать привлекательное и полное описание своего отеля

Каким должно быть 'качественное описание отеля':

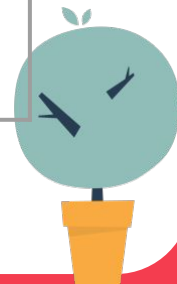
- Информативным
- Уникальным
- Ясным и лаконичным
- Привлекательным
- Правдивым

| Исследования показывают, что хорошо составленные описания повышают вероятность бронирования, так как помогают клиентам легче принять решение (данные Booking.com, TripAdvisor) |

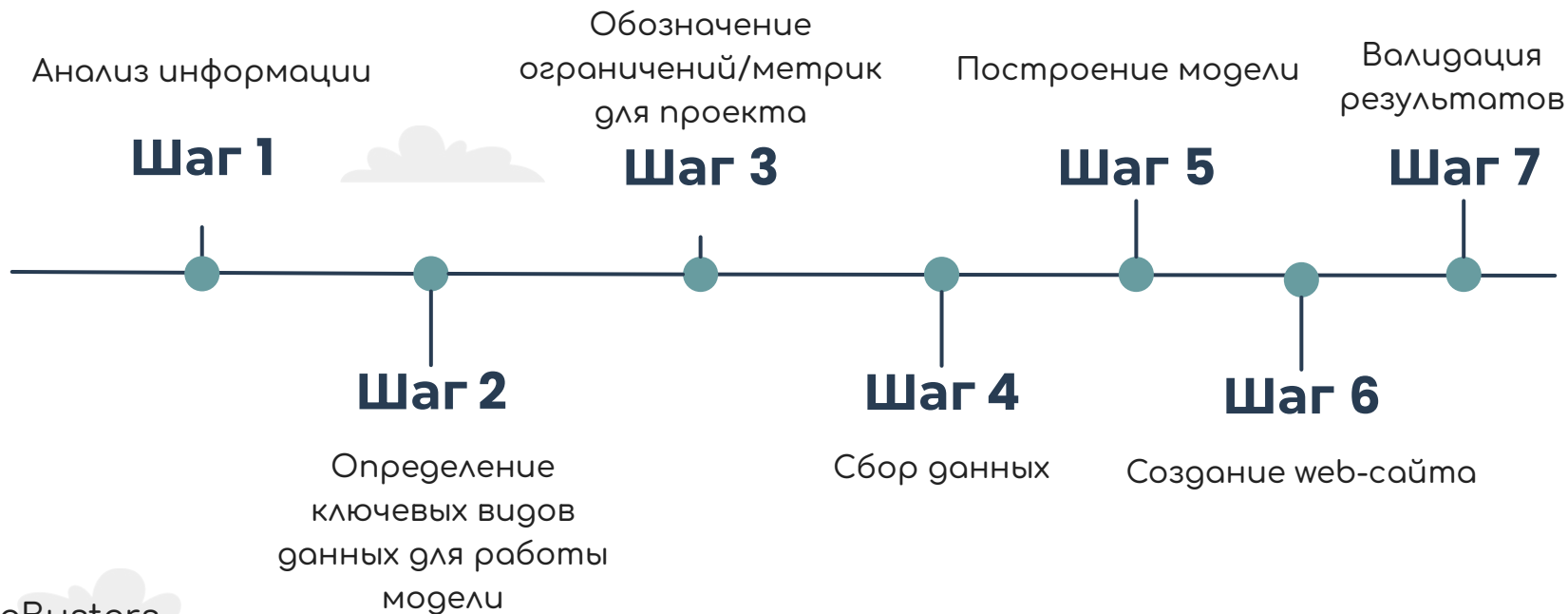


Анализ области

	Airbnb	Booking.com	Expedia Group	TripAdvisor
Методы	Индивидуальные описания от владельцев	Автоматизация + пользовательский контент	Интеграция с партнёрскими сервисами	Фокус на отзывы и рейтинги пользователей
Сильные стороны	Высокая персонализация и детализация	Экономия времени, актуальные отзывы	Широкий контент о местных достопримечательностях	Влияние на решение пользователей
Ограничения	Зависимость от писательских навыков владельцев	Машинные описания могут быть слишком общими	Сложная интеграция и возможные задержки в обновлении	Необходимость постоянного взаимодействия с пользователями



План разработки проекта



Бизнес метрики

01

Время работы

02

Требуемые
ресурсы

03

Качество
описания

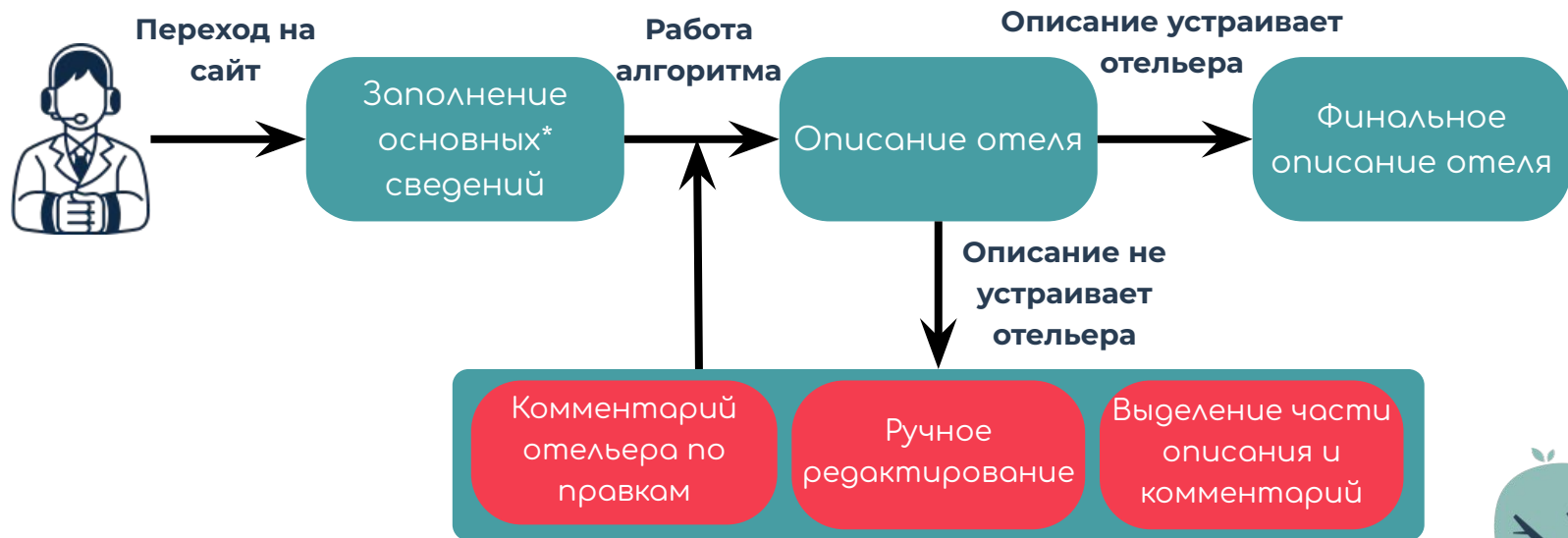
04

Стабильность
пайплайна

ByteBusters



Взаимодействие отельера с продуктом



Высокоуровневый план реализации решения



Веб

Реализация веб-интерфейса на Flask + Golang

Парсинг

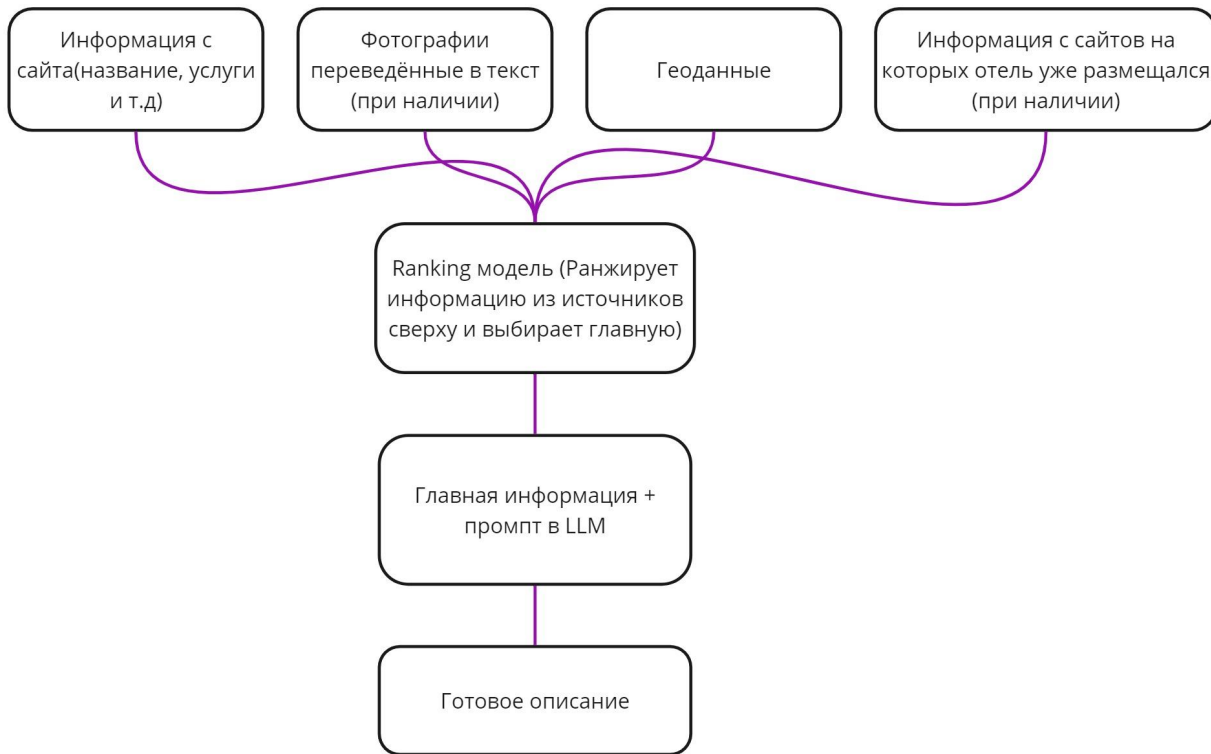
Парсинг сайтов с помощью библиотеки beautifulsoup и использования открытых api

Обработка текста

Использование saiga, llama, vikhr и их дообучение



Пайплайн работы модели



Парсинг



Введите сайт

Отельер вводит ссылку на сайт, где уже есть описание его отеля



Парсинг

С сайта парсится описание и услуги и обрабатываются LLM



Итог

Не нужно заполнять одни и те же данные, отельер экономит время



LLM



Введи текст

Отельер вводит текст,
например:
“Отель крутетецкий, недорогой,
даже зал есть новый и барчик
выпить алго”



LLM

Обрабатывает текст и
делает его
соответствующим
стандартам:
“Отель предлагает недорогое
размещение. В отеле есть новый зал
и бар, где можно приобрести
алкогольные напитки.”

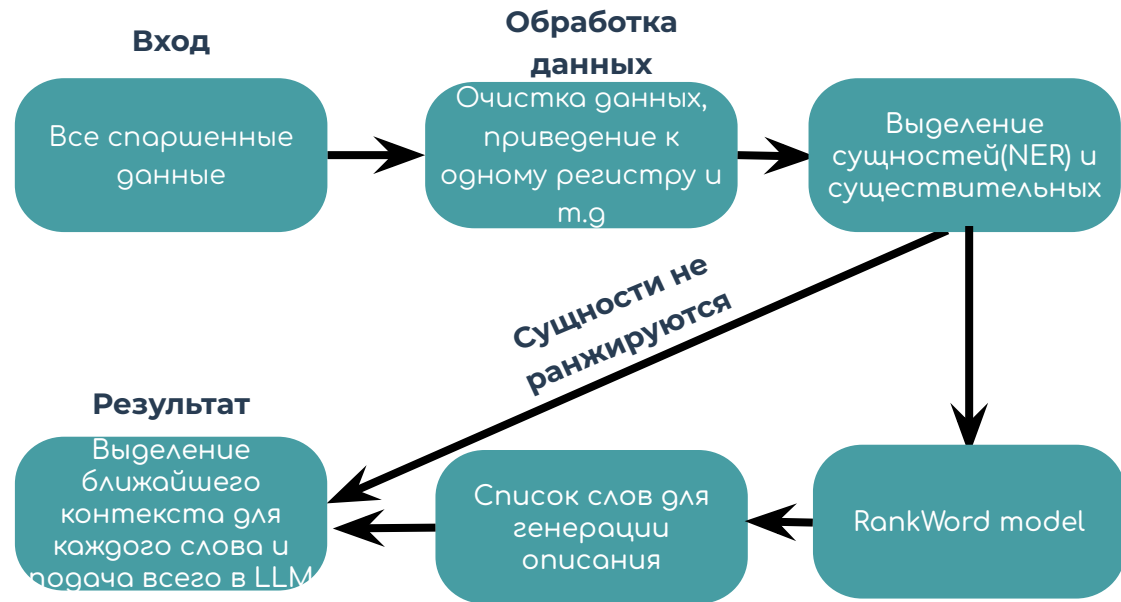


Итог

Не нужно ломать голову и
тратить уйму времени для
написания красивого
текста, достаточно
написать реальные факты
о отеле



Ranking model



Два подхода для создания RankWord model

Гипотеза один Гипотеза два

Датасет на основе отзывов

Из отзывов берутся пары слов и LLM размечает (А лучше для описания чем В)

Для каждого слова на основании сравнений присваивается ранг с помощью PageRank/Copeland rank

Строится модель для предсказания рейтинга слова на основе эмбедингов, затем подбирается трешхолд для определения значимых слов

Датасет на основе отзывов

Из отзывов берутся слова и LLM размечает (1- полезно, 0-нейтр., -1 - отриц. Влияние на привлеч. пользователей)

Строится модель для предсказания класса слова на основе эмбедингов, на инференсе берутся слова с классом один



ByteBusters



github

