



Хакатон «Транснефть — Технологии» по искусственному интеллекту

Команда НефтеБоты

Кейс: ASR

Проект: ChronoVox AI

Федоров Максим

Федоров Дмитрий

Полубарьев Михаил

Самсоненко Станислав

ПОСТАНОВКА ЦЕЛИ И ЗАДАЧ

Цель проекта:

- Создать веб-приложение для автоматического распознавания речи и интеллектуальный тайм-трекер, которые преобразуют аудиофайлы в текст и фиксируют трудозатраты по голосовым командам.

Задачи проекта:

- Разработать веб-приложение для преобразования аудиофайлов в текст с поддержкой диаризации и транскрипции в реальном времени.
- Создать интеллектуальный тайм-трекер для фиксации трудозатрат сотрудников на основе голосовых команд, с выводом данных в структурированном формате.

ВЫБОР ТЕХНОЛОГИЙ РАЗРАБОТКИ



OpenAI Whisper



Perplexity AI



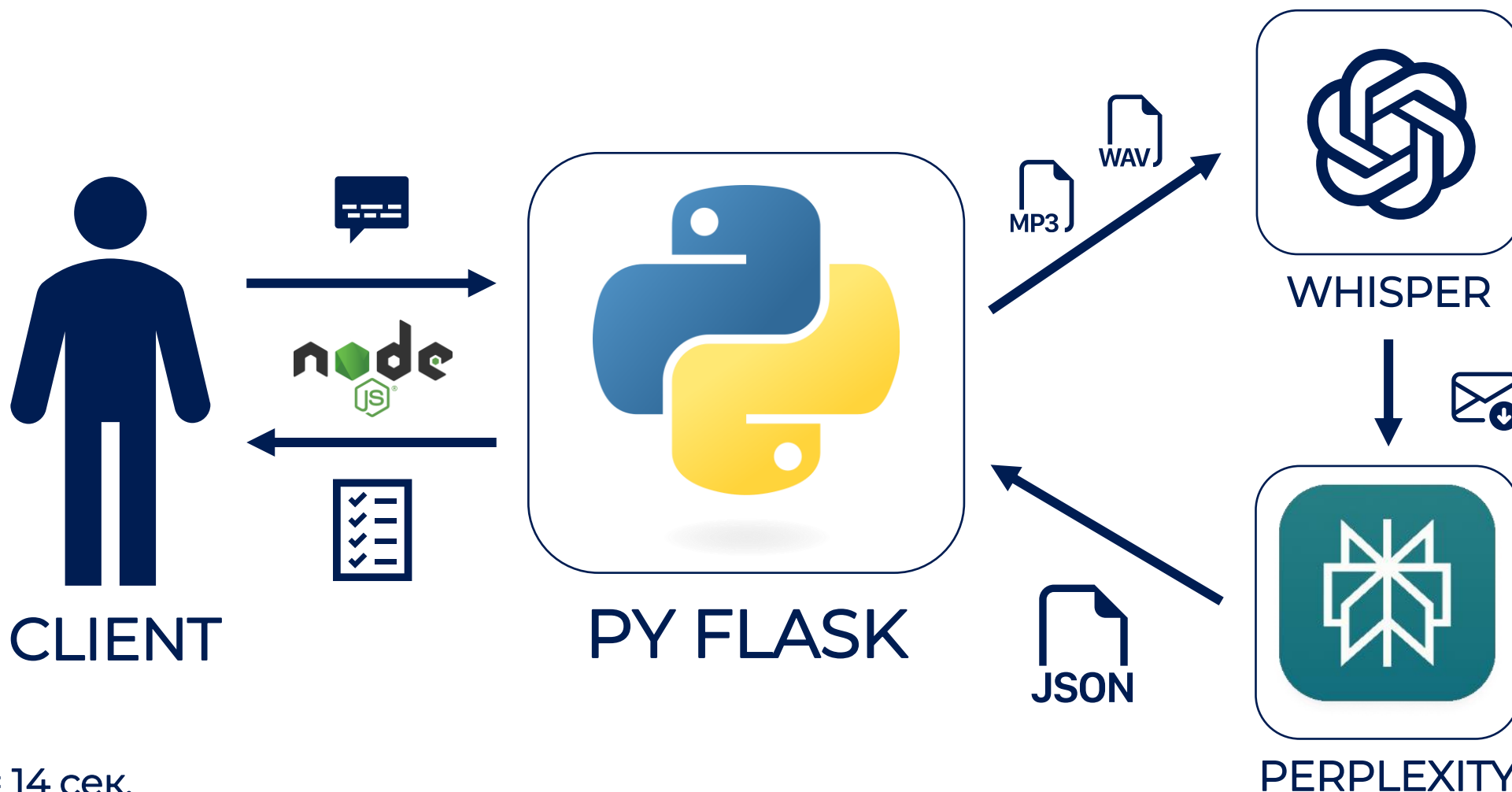
Python



Node.js

Whisper используется для автоматического распознавания речи и преобразования аудио в текст. Python обеспечивает реализацию серверной логики, включая обработку данных и интеграцию моделей. Node.js используется для создания веб-сервера и управления пользовательским интерфейсом. Perplexity применяется для анализа транскрипции и структурирования данных в удобном формате.

АРХИТЕКТУРА ПРИЛОЖЕНИЯ



🕒 = 14 сек.

СЕРВЕРНАЯ ЧАСТЬ ПРИЛОЖЕНИЯ

```
Используемая модель:  
Filename: ..\data\5.mp3  
Model_name: Whisper large  
Company: OpenAI  
Architecture: Transformer кодировщик-декодер  
Готово!  
Результат транскрипции сохранен в файл: ..\results\transcription.json
```

```
{  
  "text": " Я проводил обучение сотрудников 5 часа дня до половины третьего.",  
  "segments": [  
    {  
      "id": 0,  
      "seek": 0,  
      "start": 0.0,  
      "end": 3.66,  
      "text": " Я проводил обучение сотрудников 5 часа дня до половины третьего.",  
      "tokens": [...]  
    },  
    {  
      "temperature": 0.0,  
      "avg_logprob": -0.03095232021240961,  
      "compression_ratio": 1.2688172043010753,  
      "no_speech_prob": 0.05505261942744255  
    }  
  ],  
  "language": "ru",  
  "model_info": {  
    "filename": "..\\data\\5.mp3",  
    "model_name": "Whisper large",  
    "company": "OpenAI",  
    "architecture": "Transformer кодировщик-декодер"  
  },  
  "audio_file_name": "5.mp3"  
}
```

POST localhost:5000/transcribe-audio

Params Authorization Headers (8) Body Scripts Settings

Body Cookies Headers (6) Test Results

Pretty Raw Preview Visualize JSON

```
1 {  
2   "formalized_data": {  
3     "id": "ddf55de2",  
4     "tasks": [  
5       {  
6         "duration": "2 hours 30 minutes",  
7         "end_time": "15:30",  
8         "start_time": "13:00",  
9         "task": "Обучение сотрудников"  
10      }  
11    ],  
12    "timestamp": "2024-12-19 01:47:04"  
13  },  
14  "transcription": {  
15    "audio_file_name": "5.mp3",  
16    "language": "ru",  
17    "model_info": {  
18      "architecture": "Transformer кодировщик-декодер",  
19      "company": "OpenAI",  
20      "filename": "..\\data\\5.mp3",  
21      "model_name": "Whisper large"  
22    }  
23  }  
24 }
```

ВЕБ-ИНТЕРФЕЙС ПРИЛОЖЕНИЯ

ChronoVox AI



Преобразуйте голос в текст с помощью ИИ



Перетащите аудиофайл сюда или

Выберите файл

Обработать файл



 2.mp3 

Обработать файл

ChronoVox AI

Преобразуйте голос в текст с помощью ИИ


Обработка завершена успешно

 2.mp3 

Обработать файл

Результаты обработки

Информация о файле

 2.mp3

Транскрибированный текст

С 2 до 3.30 я занимался дизайном нового макета.

Формализованные данные

Задача 1	
Описание	дизайн нового макета
Время начала	02:00
Время окончания	03:30
Длительность	1 hour 30 minutes

СОСТАВНЫЕ ВРЕМЕННЫЕ ИНТЕРВАЛЫ

POSTlocalhost:5000/transcribe-audio

ParamsAuthorizationHeaders (8)BodyScriptsSettings

BodyCookiesHeaders (6)Test Results

PrettyRawPreviewVisualizeJSON

```
1 {
2   "formalized_data": {
3     "id": "b52adb75",
4     "tasks": [
5       {
6         "duration": "2 hours 30 minutes",
7         "end_time": "11:30",
8         "start_time": "09:00",
9         "task": "Подготовка отчета"
10      },
11      {
12        "duration": "15 minutes",
13        "end_time": "11:15",
14        "start_time": "11:00",
15        "task": "Встреча с командой"
16      },
17      {
18        "duration": "1 hour",
19        "end_time": "15:00",
20        "start_time": "14:00",
21        "task": "Работа над презентацией для клиента"
22      }
23    ],
24    "timestamp": "2024-12-19 01:00:49"
25  },
26  "transcription": {
27    "audio_file_name": "6.mp3",
```

Результаты обработки

Информация о файле

6.mp3

Транскрибированный текст

С 9 утра до половины 11-го я занимался подготовкой отчета, затем с 11 утра до четверти первого провел встречу с командой, а с 2 часов дня до без 15-ти 4 работал над презентацией для клиента.

Формализованные данные

Задача 1	
Описание	Подготовка отчета
Время начала	09:00
Время окончания	11:30
Длительность	2 hours 30 minutes

Задача 2	
Описание	Встреча с командой
Время начала	11:00
Время окончания	11:15
Длительность	15 minutes

Задача 3	
Описание	Работа над презентацией для клиента
Время начала	14:00
Время окончания	15:00
Длительность	1 hour

ИТОГИ РАЗРАБОТКИ

- Создано веб-приложение для автоматического распознавания речи с использованием модели Whisper от OpenAI, реализованное на Flask и Node.js.
- Реализован интерфейс для загрузки аудиофайлов и отображения результатов транскрипции и анализа.
- Интегрирован анализатор для структурирования данных транскрипции через Perplexity API.
- Организовано хранение данных в формате JSON, включая результаты транскрипции и формализованные данные.
- Оформлено решение в виде GitHub-репозитория с подробной документацией и инструкциями по использованию.

<https://github.com/maxikfedorov/HACK-TRANSNEFT-ASR>

The screenshot shows a GitHub repository page for 'ChronoVoice AI'. At the top, there is a file list with columns for file names, commit messages, and commit times. Below this is the 'README' section, which includes a title 'Проект: ChronoVoice AI', an 'Описание' (Description) section, an 'Установка' (Installation) section, 'Предварительные требования' (Prerequisites), and a 'Languages' section with a progress bar and a table of language usage percentages.

File	Commit Message	Commit Time
web	update, preparation for final delivery of the pr...	2 minutes ago
.gitignore	init commit	5 days ago
README.md	update, preparation for final delivery of the pr...	2 minutes ago
requirements.txt	init commit	5 days ago
ТрансНефть.pdf	corrected a typo in a word	5 days ago
ТрансНефть.pptx	update, preparation for final delivery of the pr...	2 minutes ago

Проект: ChronoVoice AI

Описание

ChronoVoice AI — это веб-приложение для транскрибирования речи с использованием моделей OpenAI Whisper и Perplexity AI. Оно предоставляет удобный веб-интерфейс на Node.js и использует Flask для обработки данных на серверной стороне.

Установка

Предварительные требования

- Python 3.8+
- Node.js 14+
- pip и npm

Languages

Language	Percentage
Python	48.8%
JavaScript	23.7%
HTML	18.4%
CSS	9.1%

**СПАСИБО ЗА
ВНИМАНИЕ!**