

# Речевая аналитика на базе VOSK

---

# Обо мне:

---



- Богдан Барабанщиков
- Образование – высшее, техническое
- Работаю в Совкомбанк Технологии, до этого Восточный экспресс банк
- >9 лет в телефонии
- Веду команду разработки

# Voice perception – план

---

- Проблемы в КЦ
- Основные функции
- Обзор интерфейса
- Стек и схема развертывания
- Подробнее про распознавание и работу с текстом

# Проблемы оценки качества в больших

## КЦ

Необходимость прослушивать

### 01

Отдел анализа качества прослушивает максимум 5% звонков.

### Человеческий фактор

### 02

Не исключена субъективная оценка при прослушивании звонков операторов.

### Задержка отчетности

### 03

При ручном отслушивании затягивается получение оценки качества базы звонков.

### Медленная реакция на изменения

### 04

Скрипты разговора с клиентами обновляются реже, как и метрики оценки разговора оператора.

# Voice perception – основные функции

---

- Хранение записей звонков
- Перевод звонков в текст
- Диаризация моно записей
- Пофразное распознавание эмоций
- Построение отчетов
- Суммаризация на базе технологий GPT
- Автотематизация на основе семантического сравнения
- WEB интерфейс

Voice perception – WEB интерфейс

Прослушивание  
ЗВОНКОВ

☰

🔍 Поиск звонков

📈 Речевая аналитика

🔍 Текстовый поиск

📎 Отправка файла

▶

1001

84212409900

0:02: здравствуй

call\_uuid7f1c7fb5-0142-4a7a-90fc-3663782312b7

call\_start\_ts2023-08-26T12:25:04+00:00

call\_end\_ts2023-08-26T12:27:58+00:00

caller1001

callee84212409900

duration174

directionoutbound

UPDATE

endDate27.08.2023 19:38

StartDate27.08.2023 19:38

caller%

callee%

Звонки: 10

10

<

>

Страница: 1

| call_start_ts             | call_end_ts               | caller     | callee      | duration | direction |
|---------------------------|---------------------------|------------|-------------|----------|-----------|
| 2023-08-26T12:25:04+00:00 | 2023-08-26T12:27:58+00:00 | 1001       | 84212409900 | 174      | outbound  |
| 2023-08-26T12:25:04+00:00 | 2023-08-26T12:27:58+00:00 | 1001       | 84212409900 | 174      | outbound  |
| 2023-08-26T09:44:49+00:00 | 2023-08-26T09:44:49+00:00 | 1002       | 1001        | 5        | outbound  |
| 2023-08-26T09:43:13+00:00 | 2023-08-26T09:43:13+00:00 | 1002       | 1001        | 12       | outbound  |
| 2023-08-26T09:41:50+00:00 | 2023-08-26T09:41:50+00:00 | 1001       | 1002        | 15       | outbound  |
| 2023-08-26T09:39:57+00:00 | 2023-08-26T09:39:57+00:00 | 1002       | 1001        | 45       | outbound  |
| 2023-08-26T09:26:40+00:00 | 2023-08-26T09:26:40+00:00 | 1001       | 1002        | 12       | outbound  |
| 2023-08-26T08:06:29+00:00 | 2023-08-26T08:06:29+00:00 | 1001       | 1002        | 24       | outbound  |
| 2022-12-05T22:26:39+00:00 | 2022-12-05T22:26:39+00:00 | 9999999999 | 9879879665  | 10       | outbound  |
| 2022-12-05T20:09:05+00:00 | 2022-12-05T20:09:05+00:00 | 9999999999 | 9876543214  | 10       | outbound  |

# Voice perception – WEB интерфейс

Суммаризация /  
Теги



Мета информация  
**call\_uuid** 549c2f89-ca9c-4ea8-ab0e-5254ab3d04ec  
**call\_start\_ts** 2023-11-19T05:36:37.111000+00:00  
**caller** 9144092014  
**callee** 8001000800  
**duration** 58  
**direction** outbound

Тэги  
приветствие согласие **приветствие** дата отказ дата **продажа курса**

Transcription Вопрос по диалогу

8001000800 9144092014

с отправил сообщение в Ростелеком, что его роутер не может подключиться к сети. Он указал номер своего лицевого счета и фамилию. Ростелеком проверил наличие аварий и попросил отправить контактный номер телефона. После этого роутер подключился, но некоторое время был в автоматической настройке. Ростелеком свяжется с ним.

UPDATE

StartDate

19.11.2022 08:42

endDate

19.11.2023 08:42

caller

%

callee

%

Звонки: 10

10 < > Страница: 1

| call_start_ts                    | caller     | callee      | duration | direction |
|----------------------------------|------------|-------------|----------|-----------|
| 2023-11-19T05:36:37.111000+00:00 | 9144092014 | 8001000800  | 58       | outbound  |
| 2023-08-26T12:25:04+00:00        | 1001       | 84212409900 | 174      | outbound  |
| 2023-08-26T12:25:04+00:00        | 1001       | 84212409900 | 174      | outbound  |
| 2023-08-26T09:44:49+00:00        | 1002       | 1001        | 0        | outbound  |
| 2023-08-26T09:43:13+00:00        | 1002       | 1001        | 0        | outbound  |
| 2023-08-26T09:41:50+00:00        | 1001       | 1002        | 0        | outbound  |
| 2023-08-26T09:39:57+00:00        | 1002       | 1001        | 0        | outbound  |
| 2023-08-26T09:26:40+00:00        | 1001       | 1002        | 0        | outbound  |
| 2023-08-26T08:06:29+00:00        | 1001       | 1002        | 0        | outbound  |
| 2022-11-21T23:00:46+00:00        | 9144092014 | 8126034161  | 10       | outbound  |

# Voice perception – WEB интерфейс

• Отчеты



Поиск звонков



Речевая аналитика

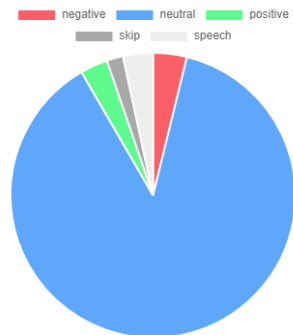


Текстовый поиск

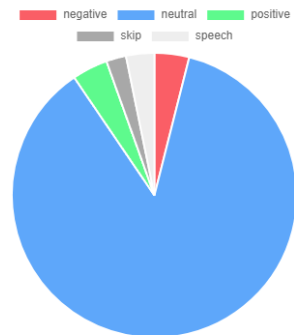


Отправка файла

emotion 2021



emotion 2022

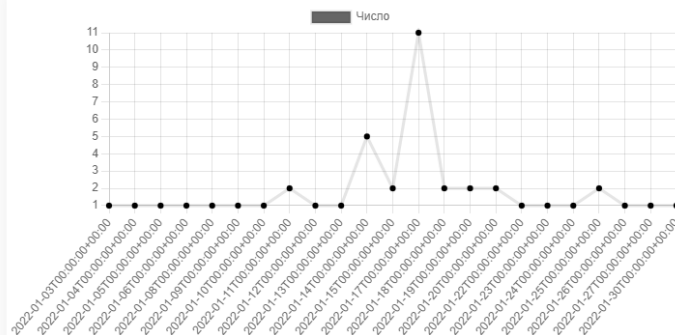


top words 2022 прилагательные

твой машин каждый готовый обязательный указанный открытый никакой **весь**  
маленький **один** разный **ЭТОТ** обычный бесплатный  
свободный **который** нормальный наш написанный сложный высокий  
сам далекий **какой-то** **такой**  
**хороший** ремесленный записанный быстрый общий  
**какой** плохой отдельный старый другой данный внутренний свой  
понятный **короткий** спокойный непонятный самый близкий  
третий **большой** сильный удобный детский тот ваш любой новый  
красный правильный прямой **второй** должный **добрый** первый  
мой **нужный** всякий возможный желтый нижний занятый большой



counts 2022-01



emotions 2022

positive **neutral** skip speech negative



# Voice perception – WEB интерфейс

Отчеты – фильтры

Поиск звонков

Речевая аналитика

Текстовый поиск

Отправка файла

emotion 2021

≡ ☒ x

UPDATE

Chart name

emotion 2021

startDate

01.01.2021

endDate

01.01.2022

caller

%

callee

%

spk

1

emotion 2022

≡ ☒ x

UPDATE

Chart name

emotion 2022

startDate

01.01.2022

endDate

01.01.2023

caller

%

callee

%

spk

1

counts 2022-01

≡ ☒ x

UPDATE

Chart name

counts 2022-01

startDate

01.01.2022

endDate

01.02.2022

caller

%

callee

%

spk

1

top words 2022 прилагательные

≡ ☒ x

UPDATE

Chart name

top words 2022 прилагательные

startDate

01.01.2022

endDate

01.01.2023

Part

ПРИЛАГАТЕЛЬНОЕ x | x v

caller

%

callee

%

spk

0

limit

70

# Voice perception – WEB интерфейс

• Настройка тегов



Поиск звонков



Речевая аналитика



Текстовый поиск



Отправка файла



Редактор Тэгов

Tag Name: приветствие Speaker: -1

добрый день x алло x здравствуйте x

Сохранить Удалить

Tag Name: дата Speaker: -1

месяц x дата x время x

Сохранить Удалить

Tag Name: продажа курса Speaker: -1

курс x обучение x

Сохранить Удалить

Tag Name: место Speaker: 1

город x регион x

Сохранить Удалить

Tag Name: согласие Speaker: 0

хорошо x да давайте x можно x да x ага x

Сохранить Удалить

Tag Name: отказ Speaker: 0

не надо x не хочу x нет x

Сохранить Удалить

Tag Name: страховка Speaker: -1

страхование x осаго x полис x полис страхования x

Сохранить Удалить

Tag Name: Speaker: -1

Select...

Создать

# Voice perception – WEB интерфейс

- Пример:  
Отображение  
эмоций

0:06: да скажите вот ситуация такая

0:10: хрущёвка

0:13: старые трубы там эти канализации и в туалете к соседям подтекает у нас воды нету нигде

0:23: так приходилось получается

0:26: сантехник от управляете смотрел говорит у вас там какой-то старый чугунный тройник треснул

0:47: не совсем соседи тоже старый нет там стоят там вроде бы не самостоятельно там не сам стояк менять надо то надо вроде как из стойка стойки тройник из него ещё тройник

0:04: да

0:34: заменить сможете это не такую монет многое вот эта вот канализационная труба чугунная диалого ну там жил соседи как соседи что-то вы не узнавали что вы не все

0:09: вот там я представляю компанию яндекс клауд удобно говорить вам

0:19: займёт пару минут

0:26: недавно вы посетили наше мероприятие учтите я хочу поблагодарить вас за участие

0:33: мероприятие было для вас полезным

0:41: возможно после мероприятий у вас остались вопросы в области речевых технологий которые хотели бы обсудить с нашим специалистом

0:07: да

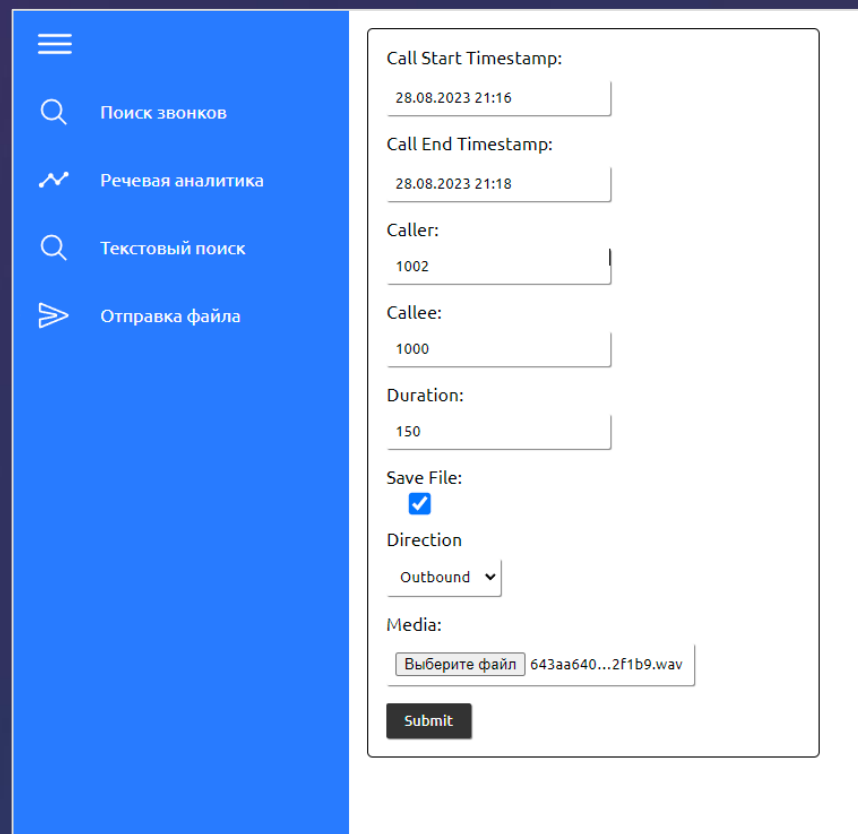
0:15: доктор

0:23: не давайте

0:37: было полезно

# Voice perception – WEB интерфейс

- Загрузка записей



The screenshot displays a web interface for voice perception. On the left is a blue sidebar with a menu containing four items: 'Поиск звонков' (Search calls), 'Речевая аналитика' (Speech analytics), 'Текстовый поиск' (Text search), and 'Отправка файла' (Send file). The main content area is white and contains a form for uploading recordings. The form includes input fields for 'Call Start Timestamp' (28.08.2023 21:16), 'Call End Timestamp' (28.08.2023 21:18), 'Caller' (1002), 'Callee' (1000), and 'Duration' (150). There is a 'Save File' checkbox which is checked, and a 'Direction' dropdown menu set to 'Outbound'. The 'Media' section has a file selection button labeled 'Выберите файл' and a text input field containing '643aa640...2f1b9.wav'. A 'Submit' button is located at the bottom of the form.

## Необходимый набор метаданных

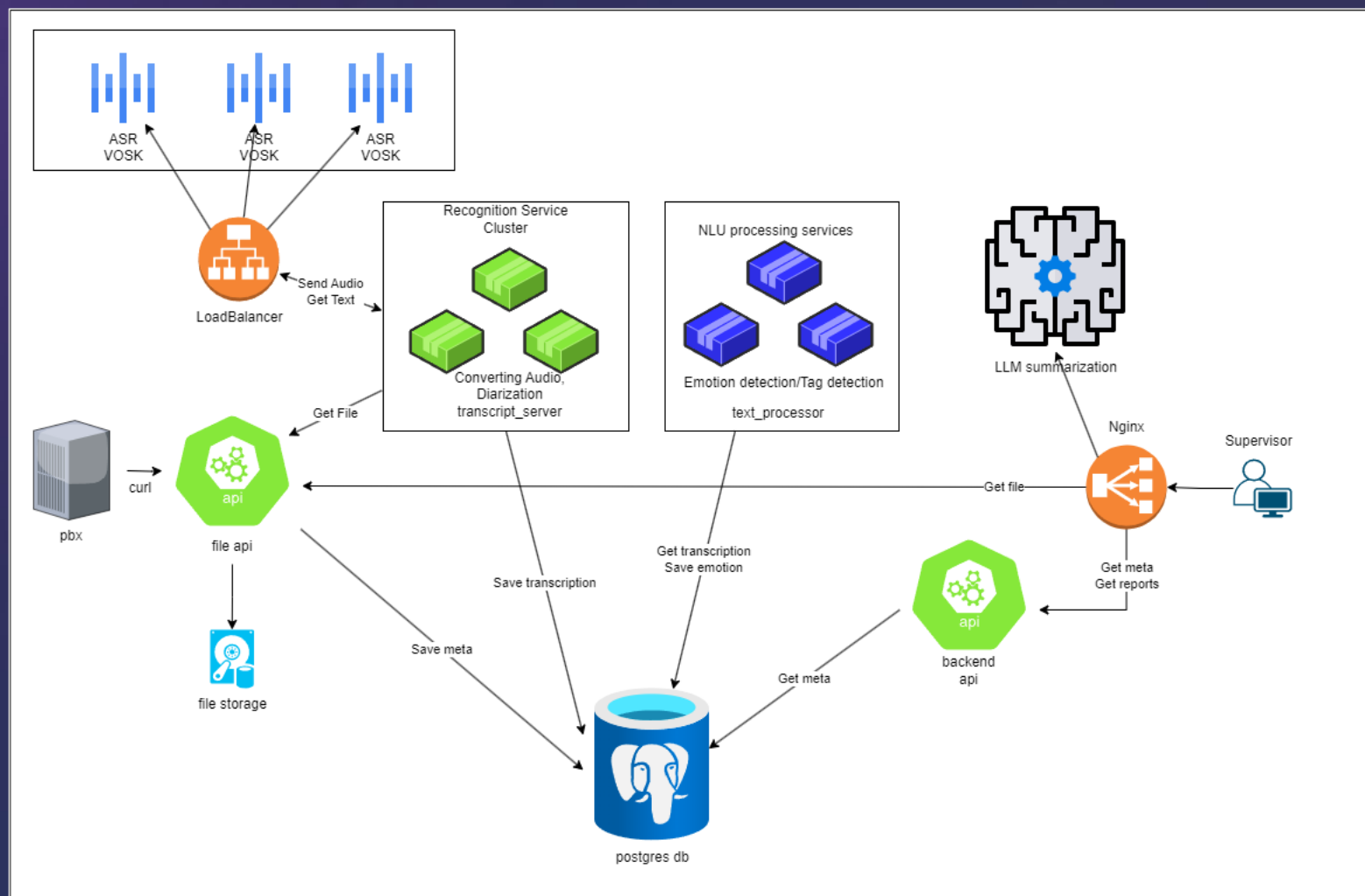
```
call_start_ts=2023-08-26 12:25:04  
call_end_ts=2023-08-26 12:27:58  
caller=1001  
calle=1002  
duration=174  
save_file=True  
media=@20230826122504.mp3  
direction=outbound
```

# Voice perception – Стек

---

- Front
  - React
- Back
  - Python/FastAPI
  - Postgres
- Окружение + ML/NN
  - Vosk, Llama.cpp
  - DeppPavlov rutokenizer rupostagger rulemma sklearn
  - Nginx
  - Docker

# Voice perception – схема развертывания



# Voice perception – функции сервера распознавания

- **Конвертация** – подготовка аудиозаписи под формат, подходящий для VOSK sample rate 8-16кГц(в зависимости от источника); 16бит
- **Разделение каналов** – Если каналов больше 1, запись разбивается на отдельные файлы для отдельного распознавания, если 1 канал, то после распознавания происходит диаризация
- **Взаимодействие с VOSK** – отправка конвертированной записи по websocket в сервер vosk с spk моделью и получение текста и векторов спикеров
- **Диаризация** – для моно записей происходит распределение фраз, полученных от vosk по каналам, максимальное число каналов 2. Распределение производится алгоритмом K-means плюс логика определения говорящего, если по фразе нет вектора спикера

# Voice perception – функции сервера процессинга текста

---

- **Лемматизация** – приведение слов к нормальной форме
- **Определение части речи** – распознавание по слову к какой части речи оно принадлежит
- **Определение эмоций** – на основе нейросети от deeppavlov rusentiment\_convers\_bert происходит определение по фразам 5 эмоций – negative, neutral, positive, skip, speech
- **Автотематизация** – проставление тематик разговора по каналам

# Voice perception – варианты развертывания

## 100-200 ч/сутки

1 сервер – все компоненты

### Ресурсы

8-16 core CPU  
16 GB RAM  
>250 GB HDD

### Компоненты

Postgresql  
File API server  
VOSK  
WEB api server  
NN server

## 1000 ч/сутки

6 серверов VOSK с CPU или 1 с GPU  
4 сервера процессинга текста  
1 сервер БД  
1 сервер для API/WEB/File storage

### Ресурсы Vosk/TP

8 core CPU/GPU gtx1080 - rtx3060  
16 GB RAM  
>250 GB HDD

### Ресурсы БД/API

4 core CPU  
8 GB RAM  
>1 TB HDD

## 1000+ ч/сутки

2+ серверов VOSK с GPU  
2+ сервера процессинга текста с GPU  
7 серверов - кластер БД  
2 сервера для API/WEB/File storage

### Ресурсы Vosk/TP

8 core CPU/GPU rtx3090+  
16 GB RAM  
>250 GB HDD

### Ресурсы БД + API

16 core CPU  
32 GB RAM  
>2 TB SSD

# Спасибо за внимание!

Буду рад ответить на все ваши вопросы  
сейчас или свяжитесь со мной в  
будущем:

— — —



**Богдан Барабанщиков**

`boltik1993@gmail.com`

`+7 914 409 20 14`

`tg://voiceperception`

