

Привет! Я — твой ассистент.  
Сейчас отвечу на твой вопрос



Где я могу найти реквизиты  
своего счета?



# AUDIOGRAM

Платформа синтеза и распознавания речи



**MTS AI  
ОДИН ИЗ  
ТЕХНОЛОГИЧЕСКИХ  
ЛИДЕРОВ**

в области искусственного  
интеллекта в России

# MTS AI



## Разрабатываем AI-решения и продукты

На основе технологий CV, NLP и Edge Computing



## Развиваем рынок в области DeepTech и AI

Инвестируем в перспективные проекты, помогаем стартапам превратить идеи в прибыльный бизнес



## R&D-центр

200+ экспертов и технологическая инфраструктура – используем самый мощный в телекоме суперкомпьютер

## Реализовали 20+ проектов

- Автоматизировали клиентский сервис МТС с помощью чат-бота
- Помогаем прогнозировать доходность от инвестиций в телеком-инфраструктуру с помощью ИИ
- Научили KION пропускать титры и определять место для рекламы

**АКТИВЫ MTS AI:**



**ИНФОРМАТИКО**



# Audiogram

Платформа синтеза и распознавания речи на базе нейронных сетей и методов машинного обучения

## ФУНКЦИОНАЛЬНОСТЬ

### ASR

Транскрибация аудиозаписи в реальном времени и получение результатов в текстовом формате

### TTS

Озвучивание текста синтезированным голосом

+ 20%

## ЭФФЕКТИВНОСТЬ

### Бизнес-эффекты

от использования сервисов синтеза и распознавания речи

### Дополнительные преимущества

Экономия на услугах дикторов

Уменьшение нагрузки на контакт-центры

**96%**

Повышение точности распознавания запросов пользователей

**20%**

Снижение затрат на оплату труда сотрудникам

**16–25%**

Увеличение продаж

**20%**

Увеличение повторных продаж

ЦЕЛЕВАЯ АУДИТОРИЯ

Кол-центры

Разработчики ботов и виртуальных ассистентов

Издательства, соцсети и мессенджеры

E-commerce

Медицина

# Audiogram

Платформа синтеза и распознавания речи на базе нейронных сетей и методов машинного обучения



Сканируйте QR-код и протестируйте Audiogram

## Художественный синтез речи



Расстановка ударений и интонаций на основе ML позволяет быстро и качественно озвучивать книги и ролики голосом, который близок к естественному

скоро

## Распознавание эмоций



Audiogram определяет тональность высказывания пользователя: позитивную негативную, нейтральную

скоро

## Голосовая биометрия



Выделение голоса конкретного спикера на фоне нескольких голосов в одном канале, определение принадлежности голоса спикеру

скоро

## Технология клонирования голоса



Создавайте голоса на основе нескольких минут речи диктора или генерируйте новые, не существующие в реальности голоса по заданным параметрам, например: мужской, низкий спокойный

## Работа в разных шумовых условиях



Audiogram успешно распознает голос, когда спикер разговаривает тихо или находится в месте, где есть посторонние звуки

## Мультиотраслевая модель



Модель распознавания речи не требует дополнительного обучения и может использоваться в любых сферах: в ритейле, телекоме, банках и других

# Audiogram

## ВОЗМОЖНОСТИ ДЛЯ КЛИЕНТОВ

### Легкая интеграция с системами заказчика

Audiogram поддерживает взаимодействие с внешними приложениями с помощью gRPC и REST API, также протоколов UniMRCP и SIP для интеграции с телефонией

### Дообучение доменных моделей

Возможность адаптировать модель распознавания речи под специфическую лексику бизнес-заказчика за 14 дней на основе 300 часов аудиозаписей

### Комплексное решение

Вы можете заказать комплексное решение, в состав которого входит Audiogram, NLP Platform для создания ботов, AI-аналитика коммуникаций, система видеонаблюдения и аналитики TenVision

## ПРОГРАММНЫЕ КОМПОНЕНТЫ

### Automatic Speech Recognition

Потоковое и файловое распознавание речи и транскрибация в текст

### Text-to-Speech

Для озвучивания текста предустановлены четыре голоса, есть автоматическая ML-разметка для литературной озвучки книг и статей и поддержка SSML для естественного звучания и управления интонацией, скоростью и ударениями

### Два типа ML-моделей

Доменная модель эффективно распознает речь по определенной тематике (телеком, банки, ритейл). Общая модель распознает лексику из любой сферы без дообучения и работает в широком диапазоне шумности

### Вспомогательные сервисы

Сбор статистики по использованию платформы и сервисы коннекторов для поддержания взаимодействия с внешними приложениями

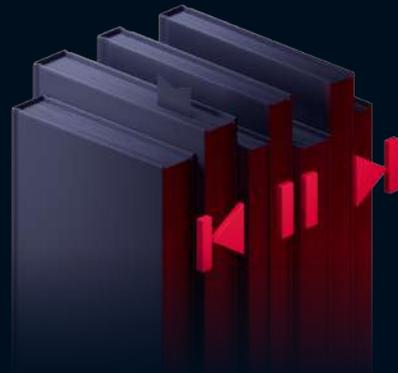
# Сценарии применения Audiogram по отраслям

## ИТ, РАЗРАБОТЧИКИ БОТОВ И УМНЫХ АССИСТЕНТОВ



- Озвучивание реплик бота
- Использование мультиотраслевой модели распознавания речи для поддержания диалога на любую тему

## ИЗДАТЕЛЬСТВА И ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ



- Озвучивание художественной, научно-популярной и другой литературы для создания аудиокниг
- Расшифровка подкастов и лекций писателей

## EDTECH-КОМПАНИИ



- Расшифровка аудио- и видеозаписей лекций
- Создание субтитров к обучающим видео, озвучивание роликов для курсов
- Озвучивание статей на сайте

## СМИ И ДРУГИЕ СОЗДАТЕЛИ КОНТЕНТА



- Автоматическая генерация субтитров для видео
- Озвучивание статей и видео, навигации по сайту для людей с ослабленным зрением
- Перевод в текст интервью и видеоконференций

# Сценарии применения Audiogram по отраслям

## СОЦИАЛЬНЫЕ СЕТИ И МЕССЕНДЖЕРЫ



- Транскрибация голосовых сообщений
- Перевод текстовых сообщений в звуковые
- Автоматическая генерация субтитров для видео

## GAMEDEV



- Внедрение функции распознавания и синтеза речи в программы и приложения для пользователей
- Аудионавигация
- Генерация субтитров для видео, озвучивание персонажей компьютерных игр

## БАНКИ, ТЕЛЕКОМ, РИТЕЙЛ



Использование Audiogram в кол-центрах компаний:

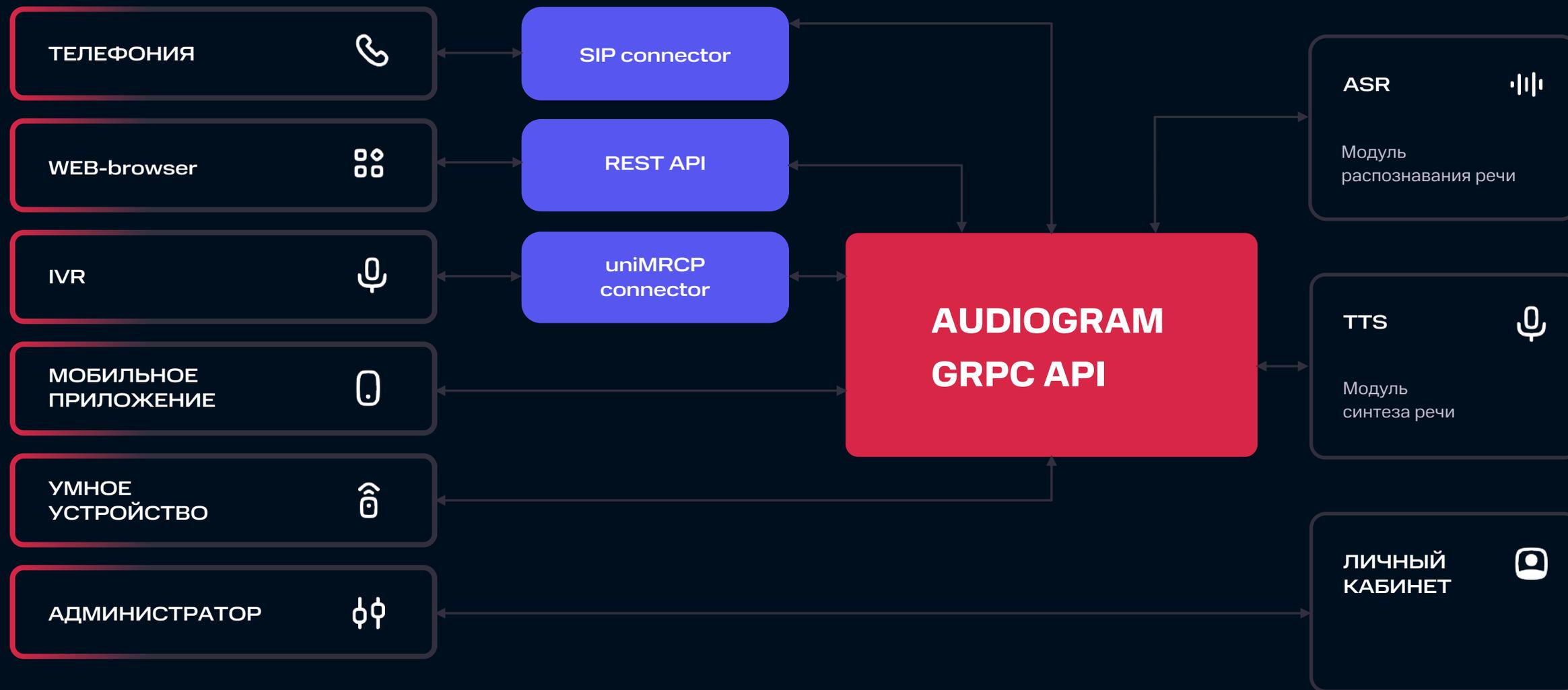
- распознавание речи пользователей и генерация ответов бота
- оперативное изменение умного голосового меню (IVR) без привлечения диктора,
- транскрибация звонков

## ТРАНСПОРТ, РИТЕЙЛ, ПРОМЫШЛЕННОСТЬ



- Озвучивание рекламных и информационных сообщений для привлечения внимания пассажиров и клиентов
- Генерация аудиоподсказок для комфортной навигации посетителей

# Схема работы Audiogram



# МОДЕЛИ ПОСТАВКИ AUDIOGRAM



## ON PREMISE

Заказчик получает дистрибутив ПО и лицензию для установки на своих серверах



## CLOUD

Audiogram будет развернут в облаке, заказчик получает для работы доступ к API и ссылку на личный кабинет



## СОВМЕСТИМОСТЬ С ДРУГИМИ ПРОДУКТАМИ MTS AI

Воспользуйтесь возможностями Audiogram в сочетании с NLP-платформой, AI-аналитикой коммуникаций и TenVision для повышения качества обслуживания клиентов и обеспечения безопасности

## AI-АНАЛИТИКА КОММУНИКАЦИЙ —

сервис для анализа голосового и текстового взаимодействия с клиентами на основе ML

- Интеграция AI-аналитики коммуникаций с Audiogram позволяет проанализировать до 100% диалогов с клиентами и получить бизнес-инсайты
- Предоставление информации о паузах, длительности и времени звонка
- Формирование отчета по различным аспектам высказывания клиента

## TENVISION —

интеллектуальная система видеонаблюдения и аналитики

Настройте детекцию звуков на кадрах с камер наблюдения, и с помощью Audiogram расшифровывайте, что говорят люди на видео

## NLP PLATFORM —

инструмент для создания голосовых и текстовых ботов и управления ими

Добавление функций Audiogram к NLP Platform позволит научить ботов разговаривать с клиентами и расшифровывать телефонные разговоры

## АНТИМАТ —

сервис, определяющий и маскирующий специальными символами табуированную лексику

Интегрировав эти сервисы в Audiogram, вы сможете предоставлять своим клиентам Очищенные от нецензурной лексики версии расшифровки аудио

## ДЕТОКСИФАЕР —

сервис для замены нецензурных и оскорбительных выражений в чатах на более нейтральный эквивалент

# Бизнес-кейсы применения Платформы Audiogram



## ЗАЩИТНИК

Внедрение Audiogram  
для защиты клиентов  
от телефонного спама



Использование  
Audiogram для  
озвучивания книг



## КОНТАКТ-ЦЕНТР

Создание  
AI-оператора  
с помощью Audiogram

# МТС «Защитник»

## ЗАДАЧА

Внедрить Audiogram в продукт МТС «Защитник» для защиты клиентов от телефонного спама и нежелательных звонков.

В составе решения:

- запись и транскрибация звонков,
- поддержание диалога со спамерами.



## РЕШЕНИЕ

Audiogram интегрировали с внутренними системами МТС.

1. Платформу синтеза и распознавания речи развернули на CPU (центральных процессорах) в контуре заказчика;
2. Audiogram подключили к ПО контакт-центра МТС с помощью протокола gRPC, который упрощает обмен сообщениями между клиентами и внутренними службами.

**Функции платформы были кастомизированы:**

- добавлен модуль расстановки знаков препинания;
- платформу научили переводить числительные в цифровые значения;
- реализован сервис «Антимат» по преобразованию нецензурной лексики в специальные символы.

## РЕЗУЛЬТАТ

МТС получила для продукта «Защитник» необходимые опции: запись и точную расшифровку спам-звонков, дослушивание сообщения до конца.

**Бизнес-эффект:**

- количество пользователей сервиса выросло в 3 раза год к году и достигло 7,2 млн\*;
- в 2,3 раза выросло время интерконнекта на спам-звонках;
- повышение лояльности пользователей за счет снижения количества нежелательных разговоров.

\* Данные на март 2023 года

# МТС «Строки»

## ЗАДАЧА

Внедрить Audiogram, чтобы озвучивать художественные тексты без привлечения диктора, аренды студии, затрат на обработку звука.



## РЕШЕНИЕ

Audiogram интегрировали с МТС «Строками».

1. Пилотируется автоматизированный процесс по созданию аудиокниг из электронных версий изданий в распространенном формате EPUB.
2. Усовершенствовали модель синтеза речи: улучшены интонации, характерные для литературных текстов, в том числе вопросительные, расстановка ударения\*.

\* Согласно исследованию, проведенному MTS AI, интонации и ударения Audiogram расставляет лучше, чем лидеры отрасли: Яндекс.Читалка и spichki.org.

## РЕЗУЛЬТАТ

С октября 2022 года MTS AI ежемесячно предоставляет сервису книги, озвученные с помощью сервиса Audiogram. По данным на март 2023 года, было озвучено более 1500 книг.

Аудиокниги готовятся для публикации в бесплатном каталоге.

В рамках договоренностей MTS AI планирует передать «Строкам» более 10 000 произведений, озвученных с помощью Audiogram.

Согласно данным по рынку, озвучивание книг синтезированным голосом дешевле в 100-300 раз услуг диктора\*.

Озвучка синтезированным голосом значительно быстрее: она занимает от 30 минут до нескольких часов, диктором — от нескольких дней до нескольких недель.

# Контакт-центр МТС

## ЗАДАЧА

Создать AI-оператора с функцией синтеза и распознавания речи для обслуживания клиентов в контакт-центрах параллельно с текущим IVR



## РЕШЕНИЕ

1. Подключили чат-бот, созданный на платформе JAICP, к АТС.
2. Модули распознавания и синтеза речи Audiogram подключили к программному обеспечению и к чат-боту контакт-центра МТС с помощью UniMRCP.
3. Настроили детектор активности речи (voice activity detector – VAD) – алгоритм, предназначенный для различения интервалов активной речи и пауз.

## РЕЗУЛЬТАТ

AI-оператор в рамках пилота принимает входящие звонки, транскрибирует их, отправляет в бот и синтезирует устный ответ.

- Эксперимент проходит в шести регионах России.
- Улучшение сервисного обслуживания привело к повышению лояльности клиентов на 17-20%.

Планируется подключить к эксперименту другие регионы и направления: банк, цифровые продукты, обслуживание абонентов стационарной телефонной сети, и перевод проекта в коммерческую эксплуатацию.

# Требуемые ресурсы для развертывания Audiogram on premise

№	КОНФИГУРАЦИЯ	CPU (шт)	RAM (Гб)	GPU (шт)	Disk (Гб)
1	ASR (CPU) + TTS (CPU)	1093	172	-	54
2	ASR (GPU) + TTS (GPU)	71	66	6	54
3	ASR (CPU)	52	11	-	22
4	ASR (GPU)	31	14	2	22
5	ASR Lite (GPU) + TTS (GPU)	63	129	6	52
6	ASR Lite (GPU)	24	82	2	28
7	TTS (GPU)	45	55	4	37

1. Расчет выполнен для 200 RTFх ASR и 200 RTFх TTS

2. Audiogram можно горизонтально масштабировать, исходя из ожидаемой нагрузки

# СПАСИБО ЗА ВНИМАНИЕ!

Задавайте  
все интересующие  
вас вопросы по адресу  
[sales@mts.ai](mailto:sales@mts.ai)

