

# AUDIOGRAM

Платформа синтеза и распознавания речи



О продукте



Кейсы и решения





Технологии

# ГРУППА КОМПАНИЙ MTS AI

Один из лидеров в области  
искусственного интеллекта в России

## R&D-центр

Разрабатываем собственные  
продукты и решения  
для бизнеса на основе  
ИИ-технологий

-  Машинное обучение
-  Компьютерное зрение
-  Обработка естественного языка
-  Генеративный ИИ

## Акселератор

Помогаем превратить проекты  
в глобальный бизнес.  
Организуем акселерационные  
программы для стартапов  
и зрелых компаний



## Венчурный фонд

Инвестируем в перспективные  
компании и технологии



**INFOMATIKA**



## Исследовательский центр

Проводим научные  
и прикладные исследования,  
публикуем статьи  
в ведущих журналах



# ПРЕИМУЩЕСТВА MTS AI

## Собственная линейка продуктов

Предлагаем клиентам запустить ИИ-сервис на базе готовых решений или заказать разработку сервиса «под ключ»



## Используем последние научные достижения



### Прикладные исследования

Открытия в аналитике и синтезе речи, новых методах разметки, развитие ML/LLM



### Фундаментальные исследования

Статьи в научных журналах Tier 1

## 200+ экспертов в сфере ИИ

Объединили ученых, разработчиков технических экспертов, предпринимателей и инвесторов, чтобы вместе развивать сферу ИИ



## Технологическая инфраструктура



Самый мощный в телекоме суперкомпьютер



CPU



GPU



Доступ к большим данным корпораций для обучения моделей

# AUDIOGRAM

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

## ФУНКЦИОНАЛЬНОСТЬ

### ASR

Транскрибирование аудиозаписи в реальном времени

### TTS

Озвучивание текста синтезированным голосом



## ЭФФЕКТИВНОСТЬ

### Бизнес-эффекты

от использования сервисов синтеза и распознавания речи

### Дополнительные преимущества

Уменьшение нагрузки на контакт-центры

Экономия на услугах дикторов

**96%**

Повышение точности распознавания запросов пользователей

**20%**

Снижение затрат на оплату труда сотрудникам

**16–25%**

Увеличение продаж

**20%**

Увеличение повторных продаж

## Области применения

Голосовые боты для Call-центров

Виртуальные ассистенты (цифровые аватары)

Протоколирование записей совещаний / вебинаров

Системы речевой аналитики

# AUDIOGRAM

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



Сканируйте QR-код и протестируйте Audiogram

## Human-like синтез речи



Расстановка ударений и интонаций на основе ML позволяет быстро и качественно озвучивать голосом, который близок к естественному. Воспользуйтесь четырьмя голосами или закажите новый

## Распознавание эмоций



Audiogram определяет тональность высказывания пользователя: позитивную негативную, нейтральную

## Антиспуфинг



Платформа умеет распознавать синтезированную речь, отличая ее от живой. Данная надстройка необходима при детекции спам-звонков пользователям

## Распознавание пола и возраста



Модель способна детектировать пол говорящего: мужской или женский, а также определяет возраст спикера

## 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 бизнеса. При необходимости модель быстро дообучается на специфическую лексику клиента

# 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

### Комплексное решение

Вы можете заказать комплексное решение, в состав которого входит Audiogram, чат-боты, AI-аналитика коммуникаций WordPulse, система видеонаблюдения и аналитики TenVision, LLM для работы с текстом Cotype

### Голосовая биометрия

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

## ПРОГРАММНЫЕ КОМПОНЕНТЫ

### Два типа ML-моделей

Доменная модель эффективно распознает речь по определенной тематике (телеком, банки, ритейл). Мультиотраслевая модель распознает лексику из любой сферы без дообучения и работает в широком диапазоне шумности. При необходимости модель можно дообучить под специфическую лексику бизнес-заказчика за 14 дней на основе 300 часов аудиозаписей

### Вспомогательные сервисы

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

Трассировка - механизм, который позволяет детально анализировать информационные потоки внутри системы, что, например, позволяет точно определить элемент, который вызвал падение производительности

Мониторинг - это средства контроля работоспособности сервисов. Например, контроль нагрузки на сервисы, контроль времени ответа и т.п.

# СЦЕНАРИИ ПРИМЕНЕНИЯ AUDIOGRAM ПО ОТРАСЛЯМ

## ИТ, РАЗРАБОТЧИКИ БОТОВ И УМНЫХ АССИСТЕНТОВ



- Интерактивное взаимодействие с клиентом благодаря использованию синтеза и распознавания речи
- Создание цифровых аватаров

## 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

## 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
- Транскрибация и протоколирование видеоконференций

## GAMEDEV



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

## ИЗДАТЕЛЬСТВА И ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ



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

## 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


# «СЕКРЕТАРЬ» С ИИ


Audiogram


JAICP

ИИ-помощник, который принимает звонки, если абонент не отвечает

## ФУНКЦИОНАЛЬНОСТЬ

 По окончании разговора абонент получает по SMS или в Telegram текст и аудиозапись диалога

 Бот определяет по ключевым словам, кто именно звонил: служба доставки, коллега, родственник или администратор салона

 ИИ различает спам и полезные звонки

## РЕАЛИЗАЦИЯ ПРОЕКТА

01 Разработчики MTS AI создали бота с помощью платформы JAICP компании Just AI

- JAICP и Audiogram интегрировали с системами MTC

02 Функции синтеза и распознавания речи реализованы на базе платформы синтеза и распознавания речи Audiogram

- Разработчики подготовили для «Секретаря» различные голоса, в том числе дерзкий мужской, веселый женский
- Подготовлены различные сценарии диалога и «револьвер» из фраз — набор шуток, позволяющий сделать общение более неформальным.

03 Внедрены антимат и денормализатор, маскирующие спецсимволами нецензурные слова

## РЕЗУЛЬТАТ

9

Кастомных голосов подготовлены для проекта

100

Сценариев диалога поддерживает бот


# ГОЛОСОВОЙ ПОМОЩНИК ДЛЯ КОНТАКТ-ЦЕНТРА МТС


Audiogram


JAICP

ИИ-оператор, обслуживающий абонентов параллельно с IVR

## ФУНКЦИОНАЛЬНОСТЬ

 Голосовой бот принимает входящие звонки и синтезирует устный ответ

 Бот помогает абонентам подключить новые тарифы и услуги, информирует о балансе

 Эксперимент проходит в шести регионах России

## РЕАЛИЗАЦИЯ ПРОЕКТА

01 Разработчики MTS AI подключили чат-бот, созданный на платформе JAICP, к АТС

02 Модули распознавания и синтеза речи Audiogram интегрировали с ПО и чат-ботом контакт-центра МТС

03 Настроен детектор активности речи (Voice Activity Detector – VAD) – алгоритм, предназначенный для различения интервалов активной речи и пауз

## РЕЗУЛЬТАТ

**+17–20%**

Повышение лояльности клиентов благодаря улучшению сервисного обслуживания\*

**+50%**

Уровень GROSS автоматизации\*

\*Данные MTS AI, МТС

# ИИ-ФУНКЦИИ ДЛЯ «ЗАЩИТНИКА»

Audiogram

Защита клиентов от спама и нежелательных звонков

## ФУНКЦИОНАЛЬНОСТЬ



ИИ расшифровывает диалог со спамером, расставляет числительные и запятые, маскирует нецензурную лексику



Абонент получает СМС с расшифровкой диалога и категорией звонка



Функционал реализован на базе платформы синтеза и распознавания речи Audiogram от MTS AI

## РЕАЛИЗАЦИЯ ПРОЕКТА

01

Audiogram был интегрирован с внутренними системами МТС

- Платформу синтеза и распознавания речи подключили к ПО контакт-центра

02

Функции Audiogram были кастомизированы

- Добавлен модуль расстановки знаков препинания
- Платформу научили переводить числительные в цифровые значения;
- Реализован сервис «Антимат»
- по преобразованию нецензурной лексики в спецсимволы.

## РЕЗУЛЬТАТ

7 млн

Пользователей сервиса. Их число стабильно растет\*

33–54 млн

Звонков ежемесячно принимает «Защитник»

+130%

Время интерконнекта на спам-звонках

Повышение лояльности пользователей за счет снижения количества нежелательных разговоров

\*Данные на сентябрь 2023 года

# АВТОМАТИЧЕСКОЕ СОСТАВЛЕНИЕ ОТЧЁТОВ ПО ИТОГАМ СОВЕЩАНИЙ

Cotype

Audiogram

## ЗАДАЧА

Пользователям корпоративного мессенджера и системы конференций нужна была возможность записывать основные аспекты коммуникаций: договоренности, инструкции, важные обсуждения и выводы

## РЕШЕНИЕ

- 01 Транскрипция звонков
- 02 Интеграция LLM для извлечения более 20 семантических аспектов разговоров
- 03 Использование большой языковой модели для выявления важных аспектов, договоренностей, выводов, поручений

## РЕЗУЛЬТАТ

Автоматизировано заполнение CRM по итогам конференций

Сотрудники получают резюме проведенных встреч, все поручения и задачи в текстовом виде

Снижение ошибок из-за человеческого фактора и технических сбоев

# ПРЕДОТВРАЩЕНИЕ МОШЕННИЧЕСТВА С ПОМОЩЬЮ ГЕНЕРАТИВНОГО ИИ

Cotype

Audiogram

## ЗАДАЧА

Один из крупнейших сервисов объявлений обратился в MTS AI с задачей обеспечить защиту своих клиентов от нежелательных звонков конкурирующих сервисов и мошенников.

Чтобы избежать возможного оттока продавцов, заказчик планирует мониторить подобные звонки и оперативно реагировать на них.

## РЕШЕНИЕ

- 01 Транскрипция звонков с помощью сервиса Audiogram
- 02 Обнаружение признаков нежелательной коммуникации с помощью Cotype — большой языковой модели MTS AI
- 03 Сбор и передача информации заказчику: название объявления, категория звонка — реклама, спам или полезный звонок

## РЕЗУЛЬТАТ

**3 секунды**

на обработку звонка и классификацию

**98%**

точность распознавания тематики звонков

Увеличение лояльности пользователя к площадке

Сокращение оттока клиентов

# СОВМЕСТИМОСТЬ С ДРУГИМИ ПРОДУКТАМИ MTS AI

Воспользуйтесь возможностями Audiogram в сочетании с чат-ботами, WordPulse, TenVision и Cotype для повышения качества обслуживания клиентов и обеспечения безопасности

## WordPulse

AI 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

## ЧАТ-БОТЫ

AI-ассистенты для автоматизации коммуникации бизнеса с клиентами

## Cotype

LLM для генерации и редактирования текстов, суммирования и анализа информации

## АНТИМАТ

сервис, определяющий и маскирующий специальными символами табуированную лексику

## ДЕТОКСИФАЕР

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

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

# ТРЕБУЕМЫЕ РЕСУРСЫ ДЛЯ РАЗВЕРТЫВАНИЯ AUDIOGRAM ON-PREMISE

№	КОНФИГУРАЦИЯ	CPU (шт.)	RAM (Гб)	GPU V100 16 Gb (шт.)	Disk (Гб)	Storage (Гб)
1	ASR Only	56	46	5	750	1000
2	TTS Only	40	60	3	750	1000
3	ASR + TTS	96	106	8	1000	1000

1. Расчет выполнен для 50 одновременных потоков ASR и 50 одновременных потоков TTS

2. Audiogram можно горизонтально масштабировать, исходя из ожидаемой нагрузки

3. Объем хранилища (Storage) зависит от профиля использования продукта

# СПАСИБО ЗА ВНИМАНИЕ!

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

