MTS AI
ONE OF THE
LEADING COMPANIES
IN AI DEVELOPMENT

in Russia

MTS AI

Developing AI solutions and products

Based on CV, NLP and Edge Computing

Developing DeepTech and AI market

Investing in promising projects, helping start ups turn their ideas into profitable business

R&D center

More than 200 experts and technological infrastructure: the most powerful supercomputer in telecommunications

20+ projects implemented

- Introduced chat bot to automate MTS customer service
- Using AI to predict profitability of investment in telecom infrastructure
- Taught KION to skip credits and locate commercials

MTS AI ASSETS: VisionLabs MACHINES CAN SEE   INFOMATIKA   Kneron   Just AI
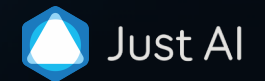
Audiogram

Neural network and machine learning-based speech recognition and synthesis platform.

+20%

FUNCTIONS

ASR
Audio recording transcription in real time and receipt of results in the text format

TTS
Text voicing with the use of a synthesized voice

EFFICIENCY

Business effects
from the use of speech synthesis and recognition services

Additional advantages

Saving on the narrator's services

Reduction of call center loading

96%
Increase in the accuracy of recognition of user requests

20%
Reduction of expenses on employees' labor remuneration

16–25%
Increase in sales

20%
Increase in repeated sales

TARGET AUDIENCE

Call centers

Developers of bots and virtual assistants

Publishing houses, social networks and messengers

E-commerce

Medicine

Audiogram

Neural network and machine learning-based speech recognition and synthesis platform.

Artistic speech synthesis

Indication of accentuation and intonationon the basis of ML makes it possible to provide a voice for the books and videos quickly and to a high standard with a voice that is close to natural

Soon

Emotion recognition

Audiogram determines the tonality of the user's statement: positive, negative, or neutral

Soon

Voice biometry

Isolation of the voice of a specific speaker on the background of several voices in a single channel, detection of the fact whether or not a certain voice belongs to a specific speaker

Soon

Voice cloning technology

Create voices on the basis of several minutes of the narrator's speech or create new voices that do not exist in the real world on the basis of preset parameters, for instance, male, low, calm

Work in various noise conditions

Audiogram successfully recognizes the voice when the speaker is talking quietly or is in a place where there are some extraneous noises

Multi sector model

Speech recognition model does not require additional training and can be used in various fields: in the retail sector, telecommunications companies, banks, etc.

Audiogram

Easy integration with the customer's systems

Audiogram supports interaction with external applications by means of gRC and REST API, as well as UniMRCP and SIP protocols to ensure integration with telephony

Additional training of domain models

Opportunity to adapt the speech recognition model to the specific vocabulary of the business customers during 14 days on the basis of 300 hours of audio recordings

Complex solution

You can order a complex solution, the contents of which include Audiogram, NLP Platform intended for bot creation, AI Speech Analytics, and video surveillance and analytics system TenVision

Automatic Speech Recognition

Stream and file-based speech recognition and transcription into text

Text-to-Speech

For text voicing, four voices are preset, and there is SSML support for natural sound and the control of intonation, speed, and accentuation

Two types of ML models

The domain model efficiently recognizes the speech related to a specific topic (telecommunications, banks, retail sector). The general model recognizes vocabulary from any field without additional training and operates in a wide range of noise emission levels

Auxiliary services

Collection of statistics on the use of the platform and services of connectors intended for the maintenance of interaction with external applications
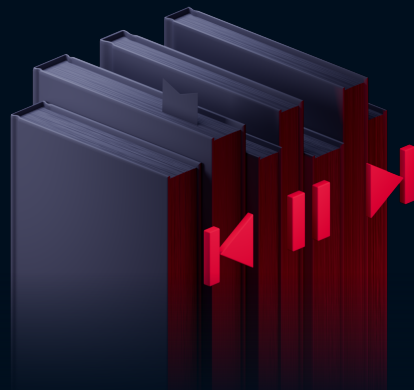
Applying audiogram in various spheres

IT, DEVELOPERS OF BOTS AND SMART ASSISTANTS

- Voicing of the bot remarks

- Use of the multi sector speech recognition model to support a dialog on any topic

PUBLISHING HOUSES AND ELECTRONIC LIBRARIES

- Voicing of fiction, sci-fi, and other types of literature with a view to create audio books

- Deciphering of podcasts and authors' lectures

EDTECH COMPANIES

- Deciphering of audio and video recordings of lectures

- Creation of subtitles to training videos, voicing of videos for training courses

- Voicing of articles on the website

MASS MEDIA AND OTHER CONTENT CREATORS

- Automatic generation of subtitles for videos

- Voicing of articles and videos, website navigation for weak-sighted people

- Conversion into text of interviews and video conferences

Applying audiogram in various spheres

SOCIAL NETWORKS AND MESSENGERS	GAMEDEV	BANKS, TELECOMMUNICATIONS COMPANIES, RETAIL SECTOR	TRANSPORT INDUSTRY, RETAIL SECTOR, MANUFACTURING INDUSTRY

- Transcription of voice messages
- Conversion of text messages into audio format
- Automatic generation of subtitles for videos

- Implementation of the speech recognition and synthesis function into software and applications for users
- Audio navigation
- Generation of subtitles for videos, voicing of computer game characters

Use of Audiogram in the call centers of companies:

- user speech recognition and bot answer generation
- prompt change of the intelligent voice menu (IVR) without engaging a narrator
- call transcription

- Voicing of advertising and information messages to attract the passengers' and customers' attention
- Audio cue generation for comfortable visitor navigation

Audiogram operation principle
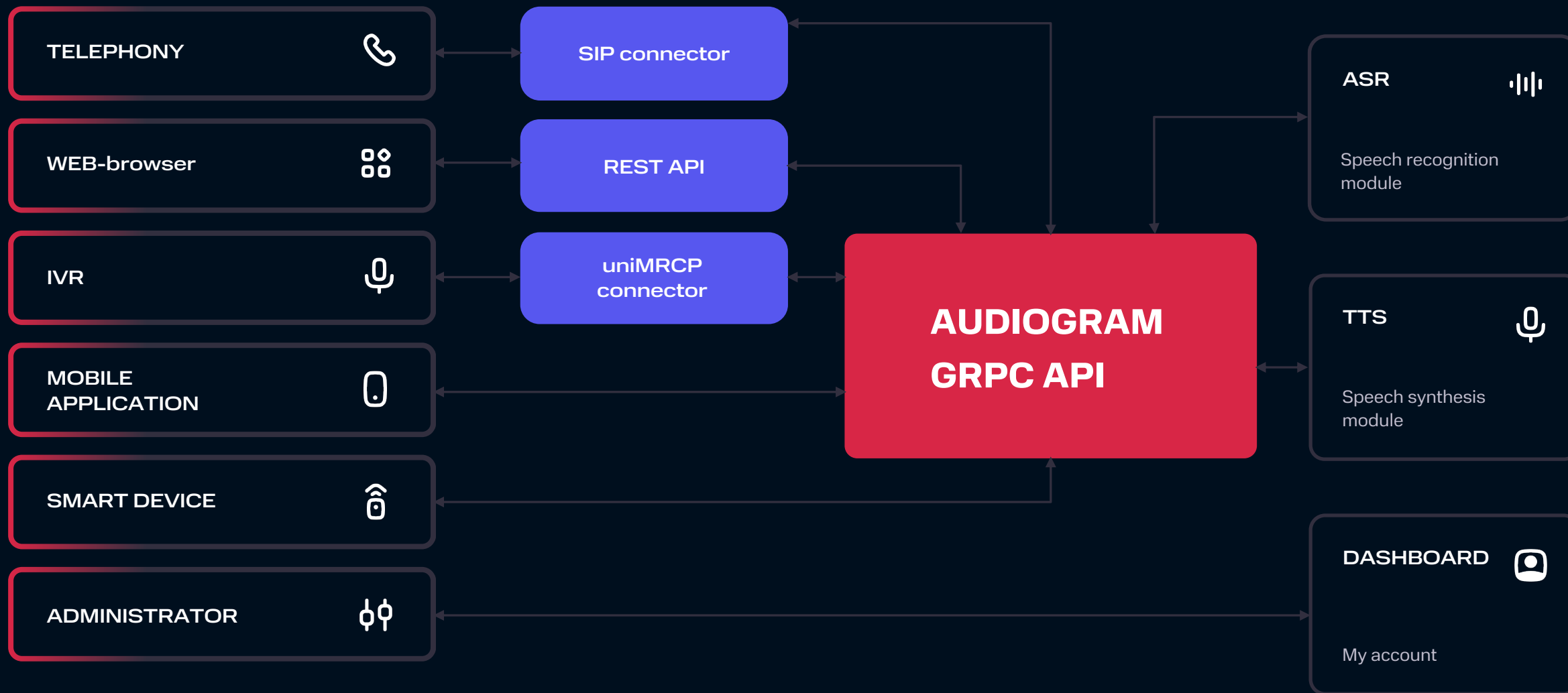
AUDIOGRAM SUPPLY MODELS

ON PREMISE

The customer receives a distribution package of software and license for its installation on his/her servers

CLOUD

Audiogram will be deployed in the cloud service; to ensure proper operation, the customer gains access to API and receives a link to his/her account

COMPATABILITY WITH OTHER MTS AI PRODUCTS

Use Audiogram opportunities in combination with the NLP Platform, AI Communication Analytics, and TenVision in order to increase the quality of customer service and ensure safety

AI SPEECH ANALYTICS

Service intended for the analysis of voice and text interaction with customers on the basis of ML

- Integration of AI Speech Analytics with Audiogram provides the opportunity to analyze up to 100% of dialogs with customers and get business insights

- Furnishing of information on pauses, duration, and time of a call

- Report generation on the basis of various aspects of the customer's statement

TENVISION

Intelligent video surveillance and analytics system

Configure sound detection on the frames received from video surveillance cameras and decipher with the use of Audiogram what the people are saying in the video

NLP PLATFORM

Tool used to create voice and text bots and control them

Addition of Audiogram function to the NLP Platform will provide the opportunity to teach the bots how to talk to customers and decipher telephone conversations

ANTI-SWEARING

Service determining and masking the taboo vocabulary with special symbols

DETOXIFIER

Service used to replace abusive and insulting terms in the chats with more neutral equivalents

By integrating these services into Audiogram, you can provide your customers with the versions of deciphering of audio recordings cleared of abusive language

Audiogram Platform Business Cases

МТС ЗАЩИТНИК

Audiogram for customer
spam protection

СТРОКИ

Audiogram
in book voice-over

МТС КОНТАКТ-ЦЕНТР

Audiogram for creating
an AI-operator

MTS "Zaschitnik" ("Defender")

TASK

Implement Audiogram
in MTS "Zaschitnik"
("Defender"): for customer
protection
from phone spammers and
unwanted calls

The solution comprises
functions for:

- recording and transcribing calls,
- keeping the dialogue
- going with spammers.

SOLUTION

Audiogram was integrated with internal MTS systems.

1. The speech recognition and synthesis platform was deployed on CPUs within the customer's loop.

2. Audiogram was connected to MTS contact center software using gRPC protocol that facilitates the message exchange between the clients and the internal services.

The functions of the platform were customized:

- the punctuation module was added;
- the platform was trained to convert numbers
- spoken in words into numbers written in digits;
- the Antimat service was implemented to convert obscene language into special characters.

RESULT

MTS has received the features necessary for the product "Zaschitnik" ("Defender"): recording and accurate deciphering of spam calls, listening to a message in full, until the end.

Business effect:

- the number of service users has increased by 3 times on a year-on-year basis and reached 7.2 mln*;
- the interconnection time on spam calls has grown by 2.3 times;
- increase in user loyalty by reducing the number of undesirable conversations.

* Data as on March 2023

MTS "Stroki" (Lines)

TASK

Implement Audiogram to create voice narrations without involving a voice talent, renting a studio, audio processing costs

СТРОКИ

SOLUTION

Audiogram was integrated with "Stroki" (Lines) .

1. The process was automated for creating audiobooks from electronic publications in the common e-book format with .epub extension

2. The speech synthesis model was enhanced: improved the intonations that are typical of literary texts, including questions, and stress placement*

* According to a study by MTS AI, Audiogram performs better than the industry leaders: Yandex.Reader and spichki.org in placing stresses and intonations.

** MTS AI conducted the survey among the MTS Library users

RESULT

From October 2022, MTS AI has been providing the service on a monthly basis with the books voiced with the use of Audiogram. According to data as on March 2023, more than 1500 books have been voiced.

Audio books are prepared for publishing in the free catalog.

As part of arrangements, MTS AI is planning to transfer to "Stroki" (Lines) more than 10000 works voiced with the use of Audiogram.

According to the market data, book voicing with a synthesized voice is cheaper by 100-300 times than the narrator's services*.

Voicing with a synthesized voice is significantly faster: it takes from 30 minutes to several hours, by the narrator — from several days to several weeks.

MTS Contact Center

TASK

Create an AI-operator featuring speech synthesis and recognition function for contact center customer service, along with IVR that is currently in use

SOLUTION

1. A chatbox that was developed on the JAICP platform was connected to PBX

2. UniMRCP was used to connect the Audiogram speech recognition and synthesis modules to the software and to the MTS Contact Center chatbot

3. The Voice Activity Detector (VAD),

4. a feature that detects active speech intervals and pauses, was configured

RESULT

Within the framework of the pilot testing, AI operator is receiving incoming calls, transcribing them, sending them to the bot, and synthesizing a spoken reply.

• The experiment is taking place in six Russian regions.

• Service improvement has led to increase in customer loyalty by 17%-20%.

Plans are underway to engage in the experiment other regions and directions: bank, digital products, service of the landline network subscribers, and putting of the project into commercial use.

Required resources for deploying Audiogram on premise

No.	CONFIGURATION	CPU (pcs.)	RAM (Gb)	GPU (pcs.)	Disk (Gb)
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. Calculation has been performed for 200 RTFx ASR and 200 RTFx TTS
2. Audiogram can be scaled horizontally, on the basis of the expected load

DMITRY CHERNOUS

MTS AI Business Development Manager

Scan the QR code, make an appointment and ask all your questions