

УТВЕРЖДЕНО

01.07.2024 г.

# **Программный комплекс распознавания и синтеза речи Voicee (Войси)**

## **ИНСТРУКЦИЯ ПО УСТАНОВКЕ ПРОГРАММЫ**

Листов 8

Москва

2024

## СОДЕРЖАНИЕ

<b>1. Назначение документа .....</b>	<b>3</b>
<b>2. Общие сведения.....</b>	<b>3</b>
<b>3. Определения, обозначения и сокращения .....</b>	<b>3</b>
<b>4. Требования к установке .....</b>	<b>3</b>
<b>5. Аппаратные требования.....</b>	<b>4</b>
<b>5.1 Десктопная часть .....</b>	<b>4</b>
<b>5.2 Серверная часть.....</b>	<b>4</b>
<b>6. Работа через Telegram-бота.....</b>	<b>4</b>
<b>7. Установка и настройка серверной версии ПО Voicee на стороне клиента .....</b>	<b>5</b>
<b>8. Настройка интеграций .....</b>	<b>6</b>
<b>9. Поддерживаемые форматы файлов .....</b>	<b>7</b>
<b>10. Скорость транскрибации .....</b>	<b>7</b>
<b>11. Поддержка и обратная связь .....</b>	<b>8</b>

## 1. Назначение документа

Настоящий документ содержит инструкцию по установке Программного комплекса распознавания и синтеза речи Voicee (далее – Voicee, Войси, программное решение, ПО).

## 2. Общие сведения

Программный комплекс предназначен для транскрибации аудио/видеофайлов в текстовый формат и решения обратной задачи — синтезированию речи для озвучивания текста голосом. Программный комплекс предназначен для работы с естественным языком в аудио форме и включает следующие функции: расстановка знаков препинания, разбиение на предложения и абзацы; распознавание и подпись спикеров; расстановка точных таймингов по словам; полностью автоматическая генерация субтитров для видео и другие функции. Программный комплекс работает с файлами на русском языке, а также поддерживает более 30 распространённых языков мира.

## 3. Определения, обозначения и сокращения

В настоящем документе применяются следующие термины с соответствующими определениями, а также обозначения и сокращения:

Сокращение	Расшифровка
ПО	Программное обеспечение
ОС	Операционная система
ПК	Персональный компьютер
API	Application Programming Interface (Интерфейс прикладного программирования)
PDF	Portable Document Format (формат переносимого документа)
TXT	Text File (текстовый файл)
DOCX	Microsoft Word Open XML Document (формат документа Microsoft Word)
SSD	Solid State Drive (твердотельный накопитель)
ОЗУ	Оперативное запоминающее устройство (RAM - Random Access Memory)
ГБ	Гигабайт
GPU	Graphics Processing Unit (графический процессор)
CUDA	Compute Unified Device Architecture (унифицированная вычислительная архитектура устройства)

## 4. Требования к установке

Для работы с Voicee через Telegram необходимо:

- устройство (смартфон, планшет, ПК) с доступом в интернет
- установленный мессенджер Telegram.

Для установки серверной версии Voicee необходимо:

- сервер с поддержкой контейнеризации Docker;
- доступ к интернету для загрузки необходимых компонентов.

## 5. Аппаратные требования

### 5.1 Десктопная часть

Тип устройства	Операционная система	Процессор	ОЗУ	Веб-браузер
Смартфоны	Android, iOS	1.5 ГГц и выше	2 ГБ	Chrome, Firefox, Safari
Планшеты	Android, iOS, Windows	1.5 ГГц и выше	2 ГБ	Chrome, Firefox, Safari
Настольные ПК и ноутбуки	Windows, macOS, Linux	2.0 ГГц и выше	4 ГБ	Chrome, Firefox, Safari
Специализированные устройства	Raspberry Pi или аналоги	1.2 ГГц и выше	1 ГБ	Chrome, Firefox, Safari

### 5.2 Серверная часть

Компонент	Рекомендуемая конфигурация
Процессор	16 ядер Intel Xeon (Haswell и выше) или аналогичный AMD
ОЗУ	Не менее 64 ГБ
Хранилище данных	SSD объемом не менее 250 ГБ

## 6. Работа через Telegram-бота

### Установка Telegram

**На смартфоне или планшете:**

1. Откройте App Store (iOS) или Google Play (Android).
2. Найдите приложение Telegram.
3. Нажмите "Установить".
4. После установки откройте приложение и зарегистрируйтесь, следуя инструкциям на экране.

**На ПК:**

1. Перейдите на официальный сайт Telegram: [telegram.org](https://telegram.org).
2. Выберите версию для вашей операционной системы (Windows, macOS, Linux).
3. Скачайте и установите Telegram, следуя инструкциям на экране.
4. Откройте приложение и зарегистрируйтесь, следуя инструкциям на экране.

## Подключение к Telegram-боту Voicee

1. Откройте Telegram.
2. В строке поиска введите "Voicee\_AI\_Bot".
3. Выберите бота из списка результатов и нажмите "Start" для начала взаимодействия.
4. Вы увидите приветственное сообщение от бота с инструкциями по дальнейшим действиям.

## Загрузка и обработка файлов через Telegram-бота

1. Отправьте боту один из следующих типов сообщений:
  - файл с видео или аудио;
  - ссылку на видео или аудио в облачном хранилище;
  - ссылку на видео в сервисе онлайн-видео (YouTube, VK Видео);
  - ссылку на аудио в сервисе подкастов или прослушивания музыки.
2. Бот автоматически распознает тип сообщения и начнет обработку файла.
3. После завершения обработки вы получите сообщение с результатами в одном из форматов:
  - PDF, TXT, DOCX;
  - текстовое сообщение с транскрипцией

## 7. Установка и настройка серверной версии ПО Voicee на стороне клиента

### Проверка соответствия системы требованиям

1. Убедитесь, что установлен Docker Engine версии 27.2 или совместимой.

### Установка Docker

#### Для Ubuntu/Debian:

```
sudo apt-get update
sudo apt-get install -y docker.io
sudo systemctl start docker
sudo systemctl enable docker
```

#### Для CentOS/RHEL:

```
sudo yum install -y yum-utils
sudo yum-config-manager --add-repo https://download.docker.com/linux/centos/docker-ce.repo
sudo yum install -y docker-ce docker-ce-cli containerd.io
sudo systemctl start docker
sudo systemctl enable docker
```

## Запуск Docker-контейнера

1. Получите docker-образ от специалистов компании «Войси» после подписания договора.
2. Выполните команду для запуска контейнера:

```
docker run -d -p 5000:5000 --gpus all voicee:latest
```

3. Это запустит веб-сервер на порту 5000, если сервер соответствует всем требованиям к аппаратному и программному обеспечению.

## Альтернативный способ: использование Docker Compose

1. Убедитесь, что у вас есть файл docker-compose.yml, полученный от компании «Войси».
2. Выполните команду для запуска контейнеров:

```
docker-compose up -d
```

3. Этот способ предпочтительнее при наличии файла docker-compose.yml.

## Проверка работоспособности

Для проверки, что вы всё сделали правильно выполните тестовую транскрибацию

```
curl -s -X POST \  
-H "Authorization: Bearer $REPLICATE_API_TOKEN" \  
-H "Content-Type: application/json" \  
-d $'{  
  "input": {  
    "audio": "https://share.voicee.ru/samples/test_sample_50sec_rus.ogg",  
    "task": "transcribe",  
    "batch_size": 2,  
    "gender": true,  
    "paragraph": true,  
    "diarization": true,  
    "debug": false  
  }  
' \  
http://localhost:5000/predictions
```

## 8. Настройка интеграций

### 8.1 HTTP API

Для интеграции с другими системами используйте HTTP API. Подробная документация доступна по адресу: <https://docs.voicee.ru/api/>

### 8.2 Настройка Telegram-бота

1. Создайте бота в Telegram с помощью @BotFather и получите токен API.

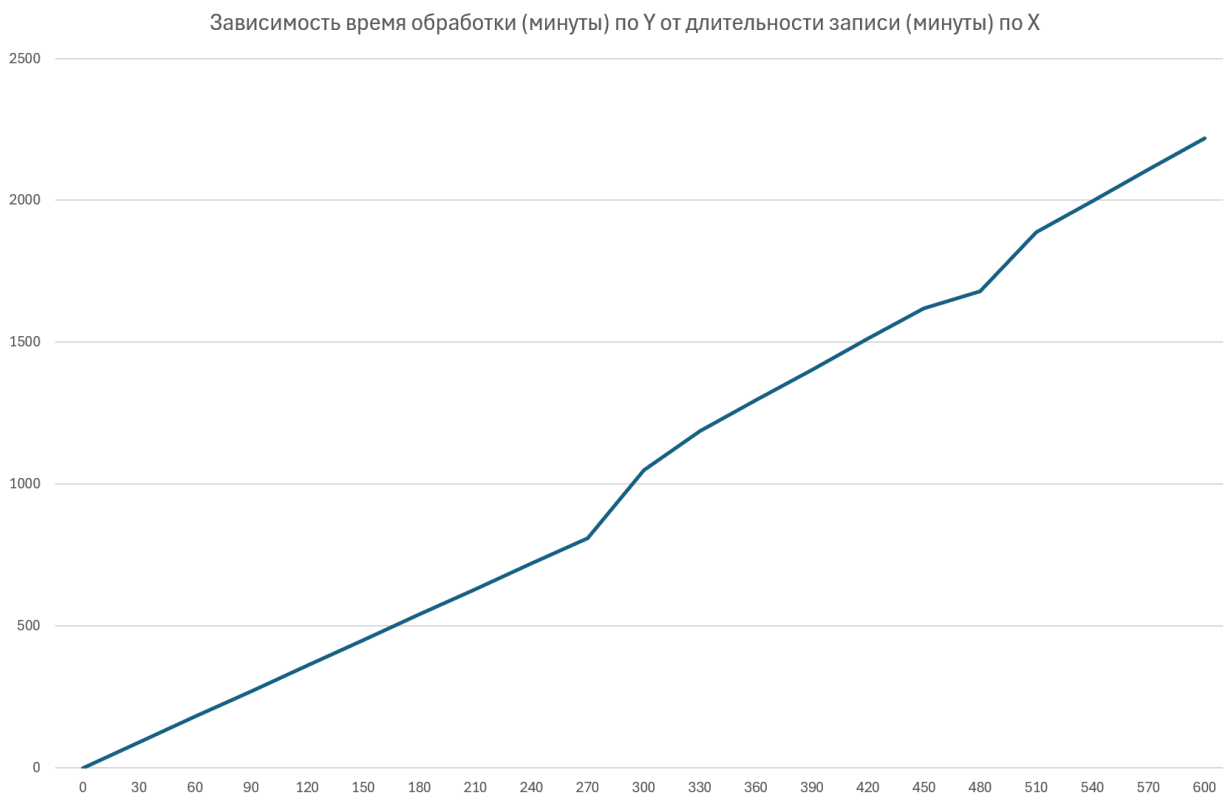
2. Настройте бота, указав адрес вашего сервера и порт.
3. Используйте методы API для настройки сценариев обработки и управления ботом.

## 9. Поддерживаемые форматы файлов

Аудио форматы	Видео форматы
MP3 (.mp3)	MP4 (.mp4)
AAC (.aac, .m4a)	AVI (.avi)
WAV (.wav)	MKV (.mkv)
FLAC (.flac)	MOV (.mov)
OGG Vorbis (.ogg)	WMV (.wmv)
ALAC (.m4a)	FLV (.flv)
AIFF (.aiff, .aif)	WebM (.webm)
WMA (.wma)	MPEG (.mpg, .mpeg)
Opus (.opus)	3GP (.3gp)
AMR (.amr)	OGV (.ogv)

## 10. Скорость транскрибации

Ниже приведен график, демонстрирующий примерную скорость транскрибации аудио и видео файлов с помощью Voicee:



## **11. Поддержка и обратная связь**

Если у вас возникли вопросы или проблемы, вы можете обратиться в службу поддержки через:

- Электронную почту: [support@voicee.com](mailto:support@voicee.com)
- Telegram: [@VoiceeSupport](https://t.me/VoiceeSupport)