

УТВЕРЖДЕНО

01.07.2024 г.

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

**ИНСТРУКЦИЯ ПО ЭКСПЛУАТАЦИИ ПРОГРАММЫ**

Листов 21

Москва

2024

## СОДЕРЖАНИЕ

<b>1 ВВЕДЕНИЕ</b> .....	3
<b>1.1 Область применения</b> .....	3
<b>1.2 Краткое описание функциональных возможностей</b> .....	3
<b>1.3 Уровень подготовки пользователя</b> .....	4
<b>1.4 Перечень эксплуатационной документации, с которыми необходимо ознакомиться пользователю</b> .....	5
<b>2 НАЗНАЧЕНИЕ И УСЛОВИЯ ПРИМЕНЕНИЯ</b> .....	5
<b>2.1 Условия применения ПО Voicee</b> .....	5
<b>2.2 Виды деятельности, функции, для автоматизации которых предназначено данное средство автоматизации</b> .....	6
<b>2.3 Условия, при соблюдении (выполнении, наступлении) которых обеспечивается применение средства автоматизации в соответствии с назначением</b> .....	7
<b>3 ПОДГОТОВКА К РАБОТЕ</b> .....	8
<b>3.1 Состав и содержание дистрибутивного носителя данных</b> .....	8
<b>3.2 Порядок загрузки данных и программ</b> .....	9
<b>3.3 Порядок проверки работоспособности</b> .....	10
<b>4 ОПИСАНИЕ ОПЕРАЦИЙ</b> .....	11
<b>4.1 Описание всех выполняемых функций, задач, комплексов задач, процедур</b> .....	11
<b>4.2 Описание операций технологического процесса обработки данных</b> .....	12
<b>5 АВАРИЙНЫЕ СИТУАЦИИ</b> .....	17
<b>5.1 Действия в случае несоблюдения условий выполнения технологического процесса, в том числе при длительных отказах технических средств</b> .....	17
<b>5.2 Действия по восстановлению программ и/или данных при отказе магнитных носителей или обнаружении ошибок в данных</b> .....	18
<b>5.3 Действия в случаях обнаружения несанкционированного вмешательства в данные</b> .....	18
<b>5.4 Действия в других аварийных ситуациях</b> .....	18
<b>6 РЕКОМЕНДАЦИИ К ОСВОЕНИЮ</b> .....	19
<b>6.1 Общая информация</b> .....	19
<b>6.2 Пошаговое руководство</b> .....	19
<b>6.3 Рекомендации по обучению и поддержке</b> .....	20
<b>6.4 Лучшие практики</b> .....	21

## 1 ВВЕДЕНИЕ

### 1.1 Область применения

Данный документ содержит описание технической архитектуры Программного комплекса распознавания и синтеза речи Voicee (далее – Voicee, Войси, программное решение, ПО), предназначенного для транскрибации аудио/видеофайлов в текстовый формат и решения обратной задачи — синтезированию речи для озвучивания текста голосом.

### 1.2 Краткое описание функциональных возможностей

Программный комплекс Voicee предлагает широкий спектр функциональных возможностей для распознавания и синтеза речи, поддерживая эффективную обработку аудио и видеоматериалов. Вот краткое описание его ключевых функций:

1. **Транскрибация аудио и видео:** преобразование устной речи из аудио- или видеофайлов в текстовый формат с высокой точностью и скоростью.
2. **Синтез речи:** озвучивание текста искусственными голосами, разработанными для реалистичного воспроизведения человеческой речи, поддерживая различные языки и акценты.
3. **Автоматическая генерация субтитров:** создание субтитров для видеоматериалов, включая автоматическую синхронизацию текста с видео.
4. **Мультиязычная поддержка:** работа с аудио и текстом на русском языке и более чем на 30 других языках, что делает программный комплекс идеальным для международного применения.
5. **Расстановка знаков препинания:** автоматизированное добавление знаков препинания в транскрибированный текст для повышения его читаемости и понимания.
6. **Разбиение на предложения и абзацы:** структурирование транскрибированного текста с четким разделением на логические разделы, предложения и абзацы.
7. **Распознавание спикеров:** идентификация и различие голосов разных спикеров в аудио- и видеозаписях, что позволяет точно отнести реплику к определённому лицу.
8. **Точные тайминги:** отметки времени для каждого слова или фразы в транскрибированном тексте, упрощающие навигацию по документу или видео.
9. **Автоматический перевод:** быстрый перевод транскрибированного текста или синтезированной речи с одного языка на другой, обеспечивая лёгкую адаптацию контента для международной аудитории.
10. **Интеграция с внешними API:** возможность улучшить и обогатить транскрибированный текст или синтезированную речь, используя различные сторонние инструменты и сервисы для пост-обработки, включая саммаризацию, автоматическое создание викторин, изменение стилистики текста и другие возможности.

В целом, Voicee является мощным и гибким инструментом для любых задач, связанных с преобразованием речи в текст и обратно, подходящим для использования в самых разных областях — от образования и юриспруденции до медиа и личного применения.

### **1.3 Уровень подготовки пользователя**

Программным комплексом Voicee можно пользоваться в трёх разных вариантах: через Telegram-бот, интеграцию с CRM или через API.

Использование программного комплекса Voicee через разные интерфейсы предполагает различный уровень навыков у пользователей. Ниже приведено описание требуемых навыков для каждого варианта использования.

#### ***1. Telegram-бот***

Целевые пользователи: частные пользователи, фрилансеры, малый бизнес и все те, кто нуждается в быстром и простом способе транскрибации аудио/видеофайлов и синтезе речи без необходимости глубоких технических знаний.

Необходимые навыки:

- основы использования Telegram: понимание того, как добавлять ботов, отправлять сообщения и файлы;
- навыки работы с интерфейсом бота: умение следовать инструкциям бота и использовать его команды для выполнения задач;
- нет необходимости в технических знаниях: процесс полностью автоматизирован и не требует понимания принципов работы технологии распознавания и синтеза речи.

#### ***2. Интеграция с CRM***

Целевые пользователи: бизнес-пользователи, включая отделы продаж, маркетинга, поддержки клиентов и другие отделы внутри компании, которым необходима интеграция функций Voicee непосредственно в их текущие системы управления взаимоотношениями с клиентами (CRM).

Необходимые навыки:

- понимание принципов CRM: Знание основных функций системы CRM и того, как она используется в вашей компании.
- базовые навыки работы с IT: Понимание того, как работать с API, может потребоваться для конфигурирования и поддержки интеграции.
- коммуникативные навыки: для взаимодействия с IT-отделом или поставщиком услуг для настройки и обслуживания интеграции.

#### ***3. Использование через API***

Целевые пользователи: разработчики программного обеспечения и IT-специалисты в компаниях, которые хотят интегрировать функционал Voicee напрямую в свои приложения, веб-сервисы или программное обеспечение.

Необходимые навыки:

- программирование: глубокие знания в области программирования, включая понимание того, как работать с REST API;
- IT-инфраструктура: понимание принципов работы сетей, протоколов обмена данными и безопасности;
- тестирование и отладка: умение тестировать интеграцию, определять и устранять возникающие проблемы;
- документация API: способность читать и понимать техническую документацию для правильной интеграции и использования API.

В зависимости от выбранного метода использования, те или иные группы пользователей должны обладать соответствующими навыками для эффективной работы с программным комплексом Voicee.

#### **1.4 Перечень эксплуатационной документации, с которыми необходимо ознакомиться пользователю**

Перед началом работы с ПО пользователю необходимо ознакомиться с данным Руководством по эксплуатации.

## **2 НАЗНАЧЕНИЕ И УСЛОВИЯ ПРИМЕНЕНИЯ**

Программный комплекс Voicee создан для обеспечения высококачественного распознавания и синтеза речи, позволяя пользователям транскрибировать аудио- и видеофайлы в текстовый формат, а также синтезировать речь из текста для озвучивания. Основное назначение этого программного обеспечения заключается в автоматизации и оптимизации процессов перевода различных видов коммуникации из устной формы в письменную и обратно, что находит широкое применение в различных сферах деятельности, включая бизнес, образование, медиа и юриспруденцию.

### **2.1 Условия применения ПО Voicee**

#### **1. Сферы применения:**

- бизнес и корпоративный сектор: для транскрипции рабочих совещаний, обучающих семинаров, маркетинговых материалов и в качестве инструмента поддержки отделов продаж и обслуживания клиентов.
- юридическая сфера: для документации заседаний, свидетельских показаний и других юридически важных событий.
- образование: для создания текстовых материалов из лекций и образовательных материалов, облегчения доступности контента для иностранных студентов через перевод.
- медиа и развлечения: для генерации субтитров, переозвучивания и доступности контента на разных языках.

#### **2. Технические требования:**

- ПО Voicee доступно для использования на любых ПК, мобильных телефонах и планшетах с подключением к интернету;
- для работы через Telegram-бот необходимо иметь учетную запись в Telegram;



- для интеграции с CRM или использования API требуются навыки в области IT и программирования.

### 3. Условия лицензирования:

- необходимо ознакомиться с условиями лицензионного соглашения, предоставляемого разработчиками Voicee, включая политику конфиденциальности и обработки данных.

### 4. Поддержка языков:

- поддержка распознавания и синтеза речи на русском языке и более чем на 30 других языках.

### 5. Доступность функций:

- доступность различных функций может варьироваться в зависимости от выбора способа использования (Telegram-бот, интеграция с CRM, API) и может требовать настройки или интеграции со стороны пользователя.

## **2.2 Виды деятельности, функции, для автоматизации которых предназначено данное средство автоматизации**

Программный комплекс Voicee является многофункциональным инструментом, предназначенным для автоматизации широкого спектра деятельности, связанной с обработкой устной и письменной информации. Основные виды деятельности и функции, автоматизация которых возможна с помощью данного средства:

### ***1. Деловая коммуникация и корпоративные мероприятия***

- транскрибация деловых встреч и переговоров: автоматическое преобразование устного обсуждения в текст для последующего анализа и архивирования.
- создание протоколов собраний: автоматическая запись обсуждений на заседаниях и совещаниях в текстовом формате.
- обработка корпоративных обучающих материалов: транскрибация и озвучивание учебных видео и аудиоматериалов.

### ***2. Образование и научные исследования***

- адаптация образовательного контента: перевод и синтез речи для создания мультязычных учебных материалов.
- транскрибация лекций и семинаров: преобразование устных лекций преподавателей в текстовую форму для студентов.
- документирование научных исследований: фиксация интервью, экспериментов и научных дискуссий в виде текста.

### ***3. Юридическая деятельность***

- документация судебных заседаний: транскрибация заседаний и свидетельских показаний для создания официальной документации.

- обработка юридических документов: распознавание и озвучивание юридических текстов для различных целей ведения дел.

#### ***4. Маркетинг и реклама***

- создание рекламного и маркетингового контента: автоматическая генерация субтитров для видеороликов, перевод и озвучивание рекламных текстов.
- анализ потребительских отзывов: транскрибация устных отзывов клиентов с последующим анализом полученной информации.

#### ***5. Медиа и развлекательная индустрия***

- генерация субтитров для фильмов и видео: создание и редактирование субтитров, а также автоматический перевод для доступности на разных языках.
- создание аудиокниг и мультимедийного контента: синтез речи из текстов для создания аудиокниг или озвучивания видео.

#### ***6. Персональное использование***

- фиксация личных заметок и дневников: преобразование устных заметок в письменный текст.
- обучение иностранным языкам: использование функций перевода и синтеза речи для обучения и практики произношения.

### **2.3 Условия, при соблюдении (выполнении, наступлении) которых обеспечивается применение средства автоматизации в соответствии с назначением**

Для обеспечения эффективного применения программного комплекса Voicee в соответствии с его назначением, необходимо соблюдение ряда условий и требований. Эти условия касаются как технической инфраструктуры, так и квалификации пользователей.

#### ***Вид ЭВМ и конфигурация технических средств***

- **устройства:** программный комплекс Voicee может быть использован на любых типах электронно-вычислительных машин, включая настольные компьютеры, ноутбуки, мобильные телефоны и планшеты.
- **конфигурация:** для обеспечения стабильной работы программы требуется подключение к интернету, наличие современного браузера для доступа к веб-версии или поддерживаемой операционной системы для мобильного приложения или десктопного клиента.

#### ***Операционная среда и общесистемные программные средства***

- **операционные системы:** Windows, macOS, Linux для ПК и Android, iOS для мобильных устройств;
- **дополнительное программное обеспечение:** не требуется специфического дополнительного ПО за исключением браузера или приложений, предоставляемых разработчиком Voicee.

### ***Входная информация и носители данных***

- **входная информация:** аудио или видео файлы в поддерживаемых форматах (MP3, WAV, MP4 и т.п.), текст для синтеза речи.
- **носители данных:** можно использовать как локальные, так и облачные хранилища данных. Программное средство может работать с файлами, загруженными напрямую или через API.

### ***Требования к подготовке специалистов***

- **пользователи:** для использования через Telegram-бот или веб-интерфейс специальной подготовки не требуется.
- **специалисты IT:** для интеграции с CRM или использования API необходимы знания в области программирования и работы с API. Должны быть разбираться в основах безопасности данных.
- **подготовка:** рекомендуется прохождение обучения или изучение документации, предоставляемой разработчиком Voicee, для специалистов, занимающихся интеграцией и технической поддержкой программного комплекса.

### ***Другие условия***

- **безопасность и конфиденциальность:** соблюдение принципов и норм защиты персональных данных и информационной безопасности.
- **лицензионные соглашения:** соблюдение условий лицензионного соглашения, предоставляемого разработчиком Voicee.

Соблюдение вышеуказанных условий обеспечит оптимальное использование программного комплекса Voicee для автоматизации распознавания и синтеза речи в соответствии с разнообразными профессиональными и персональными потребностями пользователей.

## **3 ПОДГОТОВКА К РАБОТЕ**

### **3.1 Состав и содержание дистрибутивного носителя данных**

#### ***Клиентская часть***

Для клиентской части программного комплекса Voicee предоставляются следующие варианты подключения:

- **Telegram-бот:** пользователи могут взаимодействовать с продуктом через Telegram. Поддерживаются все операционные системы, на которых существует Telegram-клиент.
- **Интеграция с CRM:** возможность подключения к различным системам управления взаимоотношениями с клиентами.
- **API:** возможность отправлять запросы к API для гибкой интеграции с другими системами.

В данном случае отсутствует физический или программный дистрибутив; вся необходимая документация и ключи предоставляются через соответствующие каналы.



### **Серверная часть (коробочная версия)**

Для серверной части (коробочной версии) дистрибутив включает:

- Docker-образ, содержащий всю необходимую программную инфраструктуру.
- набор инструкций по установке и использованию данного образа.

## **3.2 Порядок загрузки данных и программ**

### **Клиентская часть**

#### **1. Telegram-бот:**

- установите Telegram-клиент для вашей операционной системы;
- добавьте бота Voicee в Telegram;
- следуйте инструкциям, предоставленным ботом, для настройки и начала работы.

#### **2. Интеграция с CRM:**

- обратитесь к документации вашей CRM системы для получения инструкции по интеграции с Voicee.
- используйте предоставленный API ключ и следуйте шагам для завершения настройки.

#### **3. Запросы к API:**

- получите API ключ от поставщика голосового робота Voicee;
- следуйте документации API для правильной интеграции с вашей системой
- для работы с Replicate совместимым API выполните следующие шаги:
  - установите библиотеку: `pip install replicate`
  - установите токен и базовый URL в переменные окружения или добавьте их в .env файл:

```
shell
export REPLICATE_API_TOKEN=<ваш ключ>
export REPLICATE_BASE_URL=https://api.voicee.ru/
```

- используйте полученные настройки в коде приложения:

```
python
from replicate import Client

client = Client(base_url='https://api.voicee.ru/', api_token='Ваш ключ')

# Пример для транскрибации аудиофайла
audio_file_url = "https://example.com/audio.mp3"

result = client.run(
    "voicee.ru/all:transcribe",
    input={"url": audio_file_url},
)
```

```
print(result)
```

### **Серверная часть (коробочная версия)**

1. Убедитесь, что у вас установлены Docker Engine версии 27.2 или совместимой.
2. Получите Docker-образ Voicee и распакуйте набор инструкций, если это ещё не сделано.
3. Откройте командную строку или терминал.
4. Перейдите в каталог, содержащий Docker-compose файл.
5. Введите команду: ``docker compose up -d``
6. Дождитесь завершения запуска всех служб.

## **3.3 Порядок проверки работоспособности**

### **Клиентская часть**

1. Telegram-бот:
  - отправьте тестовое аудиосообщение боту и убедитесь, что получаете корректный ответ.
2. Интеграция с CRM:
  - совершите тестовый звонок через телефонию и привяжите его к сделке. Убедитесь, что данные корректно передаются и отображаются.
3. Запросы к API:
  - отправьте тестовый запрос к API Voicee:

```
shell
curl -X POST https://api.voicee.ru/v1/transcribe \
-H "Authorization: Bearer <ваш ключ>" \
-H "Content-Type: application/json" \
-d '{"url": "https://example.com/audio.mp3"}'
```
  - проверьте правильность ответа.

### **Серверная часть (коробочная версия)**

1. Введите команду ``docker ps`` в командной строке, чтобы убедиться, что все контейнеры запущены.
2. Отправьте тестовые запросы к локальному API и убедитесь, что ответы корректны.

Например:

```
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>

Если программа не проходит проверку работоспособности:

- проверьте лог-файлы Docker-контейнеров с помощью команды `docker logs [container\_id];
- перепроверьте, что все зависимости и условия, указанные в инструкциях, соблюдены;
- обратитесь к поддержке с предоставлением логов и описанием возникшей проблемы.

## 4 ОПИСАНИЕ ОПЕРАЦИЙ

### 4.1 Описание всех выполняемых функций, задач, комплексов задач, процедур

Программный комплекс Voicee выполняет следующие основные функции:

- преобразование речи на записи аудио/видео в текстовый формат;
- синтез речи на естественном языке из текста;
- анализ и пост-обработка текста, полученного из речи, для выделения ключевых мыслей, перевода на другие языки и других задач;
- определение характеристик спикера на аудио/видео записи для разделения речи разных спикеров и/или узнавания спикеров на разных записях, включая голосовое подтверждение личности;
- комбинация обозначенных выше задач для создания новых функциональных возможностей, таких как автоматический перевод и озвучивание видеороликов, сокращение длинных аудиозаписей с сохранением голоса спикера посредством его клонирования, и другие.

Программный комплекс Voicee решает следующие задачи:

- бизнес и корпоративный сектор: транскрибация рабочих совещаний, обучающих семинаров, маркетинговых материалов, поддержка отделов продаж и обслуживания клиентов;
- юридическая сфера: документация заседаний, свидетельских показаний и других юридически важных событий;
- образование: создание текстовых материалов из лекций и образовательных материалов, облегчение доступности контента для иностранных студентов через перевод;
- медиа и развлечения: генерация субтитров, переозвучивание и доступность контента на разных языках.

## 4.2 Описание операций технологического процесса обработки данных

### 4.2.1 Наименование операции: Распознавание речи

#### *Условия выполнения:*

- запись с речью на одном из поддерживаемых языков.
- подключение к интернету для отправки записи на серверы Voicee или к серверу компании, на которой развёрнут Docker-контейнер Voicee.

#### *Подготовительные действия:*

- загрузить аудиофайл или получить ссылку на файл;
- убедиться в наличии подключения к интернету;
- получить API ключ для доступа к сервису.

#### *Основные действия:*

- отправить запрос к API с ссылкой или файлом аудиозаписи;
- дождаться ответа от сервера с распознанным текстом;
- обработать полученные данные (например, для дальнейшего анализа или сохранения).

#### *Заключительные действия:*

- сохранить распознанный текст;
- отправить результат пользователю или интеграционной системе.

#### *Ресурсы:*

- процессорное время для выполнения обработки;
- память для хранения промежуточных и окончательных данных;
- API запросы для передачи и получения данных.

### 4.2.2 Наименование операции: Создание поста для блога

#### *Условия выполнения:*

- запись с речью на одном из поддерживаемых языков;
- подключение к интернету;
- подключение к серверу компании, на котором развёрнут Docker-контейнер Voicee, в случае коробочной версии.

#### *Подготовительные действия:*

- загрузить аудиофайл или получить ссылку на файл;
- убедиться в наличии подключения к интернету;
- получить API ключ для доступа к сервису.

#### *Основные действия:*

- отправить запрос к API для распознавания речи;
- дождаться ответа с расшифровкой текста;

- отправить полученный текст в LLM модель для создания поста;
- получить готовый текст поста.

***Заключительные действия:***

- сохранить созданный пост;
- отправить текст пользователю или интеграционной системе.

***Ресурсы:***

- процессорное время для выполнения обработки;
- память для хранения промежуточных и окончательных данных;
- API запросы для передачи и получения данных.

**4.2.3 Наименование операции: Создание викторины**

***Условия выполнения:***

- запись с речью на одном из поддерживаемых языков;
- подключение к интернету;
- подключение к серверу компании, на котором развёрнут Docker-контейнер Voicee, в случае коробочной версии.

***Подготовительные действия:***

- загрузить аудиофайл или получить ссылку на файл;
- убедиться в наличии подключения к интернету;
- получить API ключ для доступа к сервису.

***Основные действия:***

- отправить запрос к API для распознавания речи;
- дождаться ответа с расшифровкой текста;
- отправить расшифрованный текст на обработку для создания вопросов и ответов;
- получить готовую викторину.

***Заключительные действия:***

- сохранить полученную викторину;
- отправить результат пользователю или интеграционной системе.

***Ресурсы:***

- процессорное время для выполнения обработки;
- память для хранения промежуточных и окончательных данных;
- API запросы для передачи и получения данных.

**4.2.4 Наименование операции: Саммари**

***Условия выполнения:***

- запись с речью на одном из поддерживаемых языков;

- подключение к интернету;
- подключение к серверу компании, на котором развёрнут Docker-контейнер Voicee, в случае коробочной версии.

**Подготовительные действия:**

- загрузить аудиофайл или получить ссылку на файл;
- убедиться в наличии подключения к интернету;
- получить API ключ для доступа к сервису.

**Основные действия:**

- отправить запрос к API для распознавания речи;
- дождаться ответа с расшифровкой текста;
- отправить полученный текст для анализа и выделения ключевых мыслей;
- получить текстовое саммари.

**Заключительные действия:**

- сохранить текстовое саммари;
- отправить результат пользователю или интеграционной системе.

**Ресурсы:**

- процессорное время для выполнения обработки;
- память для хранения промежуточных и окончательных данных;
- API запросы для передачи и получения данных.

#### **4.2.5 Наименование операции: Субтитры**

**Условия выполнения:**

- видео файл с речью на одном из поддерживаемых языков;
- подключение к интернету;
- подключение к серверу компании, на котором развёрнут Docker-контейнер Voicee, в случае коробочной версии.

**Подготовительные действия:**

- загрузить видеофайл или получить ссылку на файл;
- убедиться в наличии подключения к интернету;
- получить API ключ для доступа к сервису.

**Основные действия:**

- отправить запрос к API для распознавания речи;
- дождаться ответа с расшифровкой текста;
- отправить полученный текст для синхронизации с видео;
- получить текстовые субтитры.

**Заключительные действия:**

- сохранить текстовые субтитры;
- интегрировать субтитры с видео и отправить результат пользователю.

**Ресурсы:**

- процессорное время для выполнения обработки;
- память для хранения промежуточных и окончательных данных;
- API запросы для передачи и получения данных.

#### **4.2.6 Наименование операции: Перевод**

**Условия выполнения:**

- запись с речью на одном из поддерживаемых языков;
- подключение к интернету;
- подключение к серверу компании, на котором развёрнут Docker-контейнер Voicee, в случае коробочной версии.

**Подготовительные действия:**

- загрузить аудиофайл или получить ссылку на файл;
- убедиться в наличии подключения к интернету;
- получить API ключ для доступа к сервису.

**Основные действия:**

- отправить запрос к API для распознавания речи;
- дождаться ответа с расшифровкой текста;
- отправить полученный текст на перевод;
- получить переведённый текст.

**Заключительные действия:**

- сохранить переведённый текст;
- отправить результат пользователю или интеграционной системе.

**Ресурсы:**

- процессорное время для выполнения обработки;
- память для хранения промежуточных и окончательных данных;
- API запросы для передачи и получения данных.

#### **4.2.7 Наименование операции: Транскрибация аудио и видео**

**Условия выполнения:**

- наличие аудио/видео файла в поддерживаемом формате.

**Подготовительные действия:**

- загрузка файла в систему Voicee.

**Основные действия:**

- анализ аудио/видео данных алгоритмами распознавания речи;
- распознавание и преобразование устной речи в текст;
- обработка текста для улучшения читаемости (расстановка знаков препинания, разбиение на абзацы).

**Заключительные действия:**

- сохранение транскрибированного текста в выбранном формате.

**Ресурсы:**

- программное обеспечение Voicee, вычислительные мощности для обработки данных.

**4.2.8 Наименование операции: Синтез речи из текста**

**Условия выполнения:**

- наличие текста для синтеза.

**Подготовительные действия:**

- ввод текста в программный комплекс.

**Основные действия:**

- анализ текста и его преобразование в фонетическое представление;
- генерация аудиофайла с искусственным воспроизведением текста.

**Заключительные действия:**

- сохранение сгенерированного аудиофайла.

**Ресурсы:**

- программное обеспечение Voicee, текст для синтеза, вычислительные мощности.

**4.2.9 Наименование операции: Перевод и переозвучивание контента**

**Условия выполнения:**

- наличие исходного текста или аудио/видеоматериала.

**Подготовительные действия:**

- выбор исходного и целевого языков, загрузка материала.

**Основные действия:**

- перевод текста с исходного на целевой язык;
- синтез речи на целевом языке для озвучивания переведенного текста.

**Заключительные действия:**

- сохранение переведенного и озвученного контента.



**Ресурсы:**

- программное обеспечение Voicee, исходный материал, выбранные языки для перевода, вычислительные мощности.

## **5 АВАРИЙНЫЕ СИТУАЦИИ**

### **5.1 Действия в случае несоблюдения условий выполнения технологического процесса, в том числе при длительных отказах технических средств**

При возникновении ошибок и сбоев при выполнении технологического процесса, следует предпринимать следующие шаги:

**Пропал интернет:**

- восстановите подключение к интернету;
- если файл не загрузился полностью, отправьте его на обработку повторно;
- если файл уже загружен, получите результат через API или веб-интерфейс.

**В файле нет речи:**

- проверьте запись вручную; если нет слышимой речи, проблема в исходной записи.
- если речь присутствует, но система не распознаёт её, обратитесь в службу поддержки ООО «Войси».

**Неправильно распознался язык:**

- укажите язык вручную при следующем запросе для улучшения точности распознавания.

**Некорректно разместились спикеры:**

- укажите число спикеров вручную для повышения качества разметки.

**Некоторые слова распознались некорректно:**

- проверьте качество записи.
- используйте режим корректора для нахождения и исправления ошибочных слов.

**Пропущены отрывки в распознанной речи:**

- используйте режим корректора для выявления пропущенных отрывков, обозначенных красными паузами.
- допишите пропущенные отрывки вручную и проверьте исходное качество записи.

Общий совет: всегда отправляйте на обработку оригинальную запись в виде файла без дополнительных пережатий и обработок. Voicee самостоятельно обработает и вытащит максимум качества из доступного материала.

## **5.2 Действия по восстановлению программ и/или данных при отказе магнитных носителей или обнаружении ошибок в данных**

На стороне ООО «Войси» для обеспечения надёжности данных реализованы следующие процедуры:

- реплицирование баз данных;
- регулярное резервное копирование виртуальных машин и баз данных;
- хранение файлов на закрытых S3-хранилищах с избыточностью и географическим распределением.

В случае восстановления:

- используйте последние доступные резервные копии для восстановления систем и данных;
- восстановите базу данных из реплик;
- при обнаружении повреждений данных выполните проверку консистентности и повторите восстановление из резервной копии.

На стороне коробочной версии:

- ответственность за резервное копирование данных лежит на администраторах компании-заказчика;
- коробочная версия Voicее не хранит пользовательские данные;
- для восстановления системы используйте оригинальный дистрибутив и повторите шаги установки, описанные в инструкции.

## **5.3 Действия в случаях обнаружения несанкционированного вмешательства в данные** ***Защита и мониторинг:***

- настройка фаерволов для защиты от несанкционированного доступа.
- постоянный мониторинг логов и системы на предмет подозрительной активности.

### ***Действия при обнаружении вмешательства:***

- немедленная изоляция системы для предотвращения дальнейшего повреждения данных или хищения информации;
- уведомление службы безопасности для оценки ущерба и разработки плана дальнейших действий;
- проведение анализа логов для определения источника и масштаба вмешательства.
- восстановление систем из последних безопасных резервных копий;
- усиление мер безопасности, включающих обновление фаерволов, смену паролей и обновление программного обеспечения до последних версий.

## **5.4 Действия в других аварийных ситуациях**

### ***Несанкционированный доступ или утечка данных:***

- немедленное уведомление пользователей и заинтересованных сторон о происшествии;
- анализ утечки и определение её масштабов;
- принятие мер по устранению уязвимости и предотвращению повторения инцидента.

***Сбой в облачной инфраструктуре:***

- обратитесь к провайдеру облачных услуг для получения информации о причине сбоя и сроках его устранения;
- переключитесь на резервные серверы или инфраструктуру, если это предусмотрено вашим планом аварийного восстановления;
- уведомите пользователей о проводимых работах и возможных перерывах в работе.

***Плановые и внеплановые обновления:***

- уведомление пользователей о предстоящих обновлениях и возможных перерывах в работе;
- выполнение резервного копирования данных перед началом обновления;
- запуск и проверка системы после завершения обновления.

***Ошибки конфигурации:***

- проверка и корректировка конфигурационных файлов и настроек;
- перезапуск сервисов после внесения изменений;
- выполнение тестов для подтверждения корректной работы системы после исправлений.

## **6 РЕКОМЕНДАЦИИ К ОСВОЕНИЮ**

### **6.1 Общая информация**

Программный комплекс Voicee является мощным и универсальным инструментом для распознавания и синтеза речи, анализа текстов и выполнения множества других задач, связанных с обработкой голосовых данных. Для эффективного использования комплекса рекомендуется следовать приведённым ниже рекомендациям.

### **6.2 Пошаговое руководство**

#### **6.2.1 Первый запуск**

1. Изучение документации:

Прежде чем начать работу с Voicee, внимательно ознакомьтесь с официальной документацией, особенно с разделами об установке, настройке и использовании различных функций.

2. Установка клиента и/или серверной части:

Следуйте инструкциям по установке, описанным в разделах документации, и убедитесь, что все зависимости и требования системы выполнены.

3. Регистрация и получение доступа:

Получите у поставщика необходимые ключи API для доступа к сервисам.

#### **6.2.2 Работа с основными функциями**

#### 1. Распознавание речи:

- начните с загрузки тестового аудиофайла и отправки запроса на распознавание речи через API или веб-интерфейс.
- ознакомьтесь с результатами и осознайте возможности настройки параметров для улучшения точности распознавания, таких как указание языка и количество спикеров.

#### 2. Синтез речи:

- попробуйте преобразовать текст в речь, используя различные настройки голоса, темпа и других параметров. Экспериментируйте с разными текстами, чтобы изучить возможности синтеза.

#### 3. Анализ текста и пост-обработка:

- используйте функции анализа текста для выделения ключевых мыслей, перевода на другие языки и других задач. Понимание этого процесса поможет вам применять Voicee для более сложных сценариев.

#### 4. Дополнительные функции:

- изучите возможности автоматического создания постов для блогов, создания викторин, генерации субтитров и других функций, предложенных Voicee.

### **6.2.3 Интеграция с другими системами**

#### 1. Интеграция с CRM:

- ознакомьтесь с инструкциями по интеграции Voicee с вашей CRM системой;
- настройте API и проверьте корректность передаваемых данных.

#### 2. Использование API:

- следуйте документации для интеграции Voicee через API в ваши текущие системы;
- используйте предоставленные примеры кода и шага за шагом встраивайте функции в свои процессы.

### **6.3 Рекомендации по обучению и поддержке**

#### 1. Вебинары и обучающие видео:

- посещайте вебинары и смотрите обучающие видео, которые предоставляются командой Voicee. Это поможет вам быстрее освоить систему и узнать о новых функциях и возможностях.

#### 2. Обратная связь и поддержка:

- в случае возникновения вопросов или проблем не стесняйтесь обращаться в службу поддержки ООО «Войси». Регулярно проверяйте обновления документации и релиз-ноты для получения информации о новых функциях и улучшениях.

## 6.4 Лучшие практики

### 1. Качество исходных данных:

- всегда старайтесь использовать аудиофайлы высокого качества без лишнего сжатия и обработки. Оригинальные записи обеспечивают наилучшие результаты распознавания и анализа.

### 2. Оптимизация параметров:

- экспериментируйте с настройками и параметрами различных функций Voicee для достижения оптимальных результатов. Указывайте язык, количество спикеров и другие параметры вручную, если это необходимо.

### 3. Резервное копирование:

- регулярно создавайте резервные копии данных и настроек ваших систем, особенно если вы используете коробочную версию Voicee. Это поможет избежать потери данных и быстрого восстановления системы в случае сбоев.