

УТВЕРЖДЕНО
28.11.2025 г.

**Программный комплекс
распознавания и синтеза речи «Войси 2.0»
(Voicee)**

ИНСТРУКЦИЯ ПО ЭКСПЛУАТАЦИИ ПРОГРАММЫ

Листов 35

Москва
2025



СОДЕРЖАНИЕ

1 Введение	3
1.1 Область применения	3
1.2 Краткое описание функциональных возможностей	3
1.3 Уровень подготовки пользователя.....	5
1.4 Перечень эксплуатационной документации, с которыми необходимо ознакомиться пользователю	7
2 Назначение и условия применения.....	7
2.1 Виды деятельности, функции, для автоматизации которых предназначено данное средство автоматизации.....	9
2.2 Условия, при соблюдении (выполнении, наступлении) которых обеспечивается применение средства автоматизации в соответствии с назначением	11
3 Подготовка к работе	13
4 Описание операций	17
5 Аварийные ситуации.....	28
6 Рекомендации к освоению.....	32

1 Введение

1.1 Область применения

Данный документ содержит описание технической архитектуры Программного комплекса распознавания и синтеза речи «Войси 2.0» (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 55 языках, включая русский, делая 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, позволяя 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.

11. Распознавание эмоционального окраса речи: Автоматическое определение эмоционального тона предложений (грустный, злобный, весёлый, нейтральный) для анализа тональности разговоров и встреч.

12. Оптимизация для специфичных доменов: Специализированные модели для телефонных разговоров, видеоконференций, студийных записей, диктофонных записей и новостных выпусков с автоматическим определением типа аудиозаписи.

13. Расширенная пост-обработка текста: Генерация конспекта, структурированных итогов встречи (summary, action items), таймкодов ключевых моментов, а также выполнение индивидуальных инструкций пользователя.

14. Интерактивный веб-редактор расшифровок: Веб-интерфейс с возможностью редактирования текста, переименования спикеров, синхронного воспроизведения аудио/видео по клику, экспорта изменений.

15. Корпоративные аккаунты: Создание аккаунтов для юридических лиц с привязкой сотрудников, системой ролей (администратор, сотрудник), отдельным учётом использования ресурсов и выгрузкой отчётов.

16. Управление API-ключами: Генерация, отключение и удаление ключей, установка лимитов, просмотр статистики и протоколирование действий для корпоративных клиентов.

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. Давайте разберемся, какие именно навыки требуются для каждого из вариантов использования:

1. Telegram-бот

Целевые пользователи: Этот вариант идеально подходит для частных пользователей, фрилансеров, малого бизнеса и всех тех, 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: Этот подход наиболее подходит для бизнес-пользователей, включая отделы продаж, маркетинга, поддержки клиентов и других отделов 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: Этот метод предназначен для разработчиков программного обеспечения и IT-специалистов 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.

Условия применения ПО Voicee

1. Сферы применения:

- Бизнес и корпоративный сектор: Для транскрибации 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 55 языках, включая русский.

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.

2.1 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. Вот 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.

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

Описывается вид ЭВМ и конфигурация технических средств, операционная среда и общесистемные программные средства, входная информация, носители данных, база данных, требования к подготовке специалистов и т. П

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 разбираются 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 система удовлетворяет необходимым требованиям:

— Установленный 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.

3. 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.
- Распознавание эмоционального окраса речи для анализа тональности разговоров.
- Автоматическая оптимизация параметров обработки в зависимости от типа аудиозаписи (телефония, видеоконференции, студийные записи и др.).
- Расширенная ИИ-пост-обработка текста: конспект, итоги встречи, таймкоды, пользовательские инструкции.
- Управление корпоративными аккаунтами: роли, разграничение доступа, учёт ресурсов, отчётность.
- 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

4.2.2 Условия:

- 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.

4.2.3 Подготовительные действия:

- Загрузить аудиофайл или получить ссылку на файл.
- Убедиться в наличии подключения к интернету.
- Получить API ключ для доступа к сервису.

4.2.4 Основные действия:

1. Отправить запрос к API с ссылкой или файлом аудиозаписи.
2. Дождаться ответа от сервера с распознанным текстом.
3. Обработать полученные данные (например, для дальнейшего анализа или сохранения).

4.2.5 Заключительные действия:

- Сохранить распознанный текст.
- Отправить результат пользователю или интеграционной системе.

4.2.6 Ресурсы:

- Процессорное время для выполнения обработки.
- Память для хранения промежуточных и окончательных данных.
- API запросы для передачи и получения данных.

4.2.1 Наименование операции: Создание поста для блога

4.2.2 Условия:

- 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.

4.2.3 Подготовительные действия:

- Загрузить аудиофайл или получить ссылку на файл.
- Убедиться в наличии подключения к интернету.
- Получить API ключ для доступа к сервису.

4.2.4 Основные действия:

1. Отправить запрос к API для распознавания речи.
2. Дождаться ответа с расшифровкой текста.
3. Отправить полученный текст в LLM модель для создания поста.
4. Получить готовый текст поста.

4.2.5 Заключительные действия:

- Сохранить созданный пост.
- Отправить текст пользователю или интеграционной системе.

4.2.6 Ресурсы:

- Процессорное время для выполнения обработки.
- Память для хранения промежуточных и окончательных данных.
- API запросы для передачи и получения данных.

4.2.1 Наименование операции: Создание викторины

4.2.2 Условия:

- 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.

4.2.3 Подготовительные действия:

- Загрузить аудиофайл или получить ссылку на файл.
- Убедиться в наличии подключения к интернету.
- Получить API ключ для доступа к сервису.

4.2.4 Основные действия:



1. Отправить запрос к API для распознавания речи.
2. Дождаться ответа с расшифровкой текста.
3. Отправить расшифрованный текст на обработку для создания вопросов и ответов.
4. Получить готовую викторину.

4.2.5 Заключительные действия:

- Сохранить полученную викторину.
- Отправить результат пользователю или интеграционной системе.

4.2.6 Ресурсы:

- Процессорное время для выполнения обработки.
- Память для хранения промежуточных и окончательных данных.
- API запросы для передачи и получения данных.

4.2.1 Наименование операции: Саммари

4.2.2 Условия:

- 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.

4.2.3 Подготовительные действия:

- Загрузить аудиофайл или получить ссылку на файл.

- Убедиться в наличии подключения к интернету.
- Получить API ключ для доступа к сервису.

4.2.4 Основные действия:

1. Отправить запрос к API для распознавания речи.
2. Дождаться ответа с расшифровкой текста.
3. Отправить полученный текст для анализа и выделения ключевых мыслей.
4. Получить текстовое саммари.

4.2.5 Заключительные действия:

- Сохранить текстовое саммари.
- Отправить результат пользователю или интеграционной системе.

4.2.6 Ресурсы:

- Процессорное время для выполнения обработки.
- Память для хранения промежуточных и окончательных данных.
- API запросы для передачи и получения данных.

4.2.1 Наименование операции: Субтитры

4.2.2 Условия:

- 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.

4.2.3 Подготовительные действия:

- Загрузить видеофайл или получить ссылку на файл.
- Убедиться в наличии подключения к интернету.
- Получить API ключ для доступа к сервису.

4.2.4 Основные действия:

1. Отправить запрос к API для распознавания речи.
2. Дождаться ответа с расшифровкой текста.
3. Отправить полученный текст для синхронизации с видео.
4. Получить текстовые субтитры.

4.2.5 Заключительные действия:

- Сохранить текстовые субтитры.
- Интегрировать субтитры с видео и отправить результат пользователю.

4.2.6 Ресурсы:

- Процессорное время для выполнения обработки.
- Память для хранения промежуточных и окончательных данных.
- API запросы для передачи и получения данных.

4.2.1 Наименование операции: Перевод



4.2.2 Условия:

- 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.

4.2.3 Подготовительные действия:

- Загрузить аудиофайл или получить ссылку на файл.
- Убедиться в наличии подключения к интернету.
- Получить API ключ для доступа к сервису.

4.2.4 Основные действия:

1. Отправить запрос к API для распознавания речи.
2. Дождаться ответа с расшифровкой текста.
3. Отправить полученный текст на перевод.
4. Получить переведённый текст.

4.2.5 Заключительные действия:

- Сохранить переведённый текст.
- Отправить результат пользователю или интеграционной системе.

4.2.6 Ресурсы:

- Процессорное время для выполнения обработки.
- Память для хранения промежуточных и окончательных данных.

- API запросы для передачи и получения данных.

4.2.1 Наименование операции: Распознавание эмоций

4.2.2 Условия:

- Запись с речью на одном из поддерживаемых языков (русский, английский, французский).

- Подключение к интернету или к серверу с Docker-контейнером Voicee.

4.2.3 Подготовительные действия:

- Загрузить аудиофайл или получить ссылку на файл.

4.2.4 Основные действия:

1. Отправить запрос к API с параметром `emotion=true`.
2. Дождаться ответа с размеченными эмоциями по предложениям.
3. Обработать результаты для анализа тональности разговора.

4.2.5 Заключительные действия:

- Сохранить результаты с эмоциональной разметкой.

4.2.6 Ресурсы:

- Процессорное время для выполнения обработки.
- API запросы для передачи и получения данных.

Этот раздел документации описывает основные операции и технологические процессы, связанные с функциональностью программного комплекса Voicee. Если у вас есть какие-либо комментарии или дополнения, пожалуйста, дайте знать.

Блок 4 заполняется для всех функций/задач/процедур/модулей ПО

Транскрибация аудио и видео



1. Наименование операции: Транскрибация аудио и видео.
2. Условия выполнения: Наличие аудио/видео файла в поддерживаемом формате.
3. Подготовительные действия: Загрузка файла в систему Voicee.
4. 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).
5. Заключительные действия: Сохранение транскрибированного текста в выбранном формате.
6. Ресурсы: Программное обеспечение Voicee, вычислительные мощности для обработки данных.

Синтез речи

1. Наименование операции: Синтез речи из текста.
2. Условия выполнения: Наличие текста для синтеза.
3. Подготовительные действия: Ввод текста в программный комплекс.
4. Основные действия:
 - Анализ текста и его преобразование в фонетическое представление.
 - Генерация аудиофайла с искусственным воспроизведением текста.
5. Заключительные действия: Сохранение сгенерированного аудиофайла.
6. Ресурсы: Программное обеспечение Voicee, текст для синтеза, вычислительные мощности.

Мультиязычный перевод и переозвучивание

1. Наименование операции: Перевод и переозвучивание контента.
2. Условия выполнения: Наличие исходного текста или аудио/видеоматериала.
3. Подготовительные действия: Выбор исходного и целевого языков, загрузка материала.
4. Основные действия:
 - Перевод текста с исходного на целевой язык.
 - Синтез речи на целевом языке для озвучивания переведенного текста.
5. Заключительные действия: Сохранение переведенного и озвученного контента.
6. 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 распознал 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 обрабатывает и вытаскивает 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 Voicee 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:

1. Немедленная изоляция системы для предотвращения дальнейшего повреждения данных или хищения информации.
2. Уведомление службы безопасности для оценки ущерба и разработки плана дальнейших действий.
3. Проведение анализа логов для определения источника и масштаба вмешательства.
4. Восстановление систем из последних безопасных резервных копий.
5. 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 Voicее.

6.2.3 Интеграция с другими системами

1. Интеграция с CRM:

- Ознакомьтесь с инструкциями по интеграции Voicее с вашей CRM системой. Настройте API и проверьте корректность передаваемых данных.

2. Использование API:

- Следуйте документации для интеграции Voicее 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 Voicее. 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.