

Платформа для распознавания, синтеза и анализа речи «SOVA Speech»

**Документация, содержащая описание
функциональных характеристик экземпляра
программного комплекса, предоставленного
для проведения экспертной проверки.**

Оглавление

| | |
|---|---|
| 1. ОБЩИЕ СВЕДЕНИЯ | 3 |
| 1.1. Обозначение и наименование программы..... | 3 |
| 1.2. Языки программирования, на которых написана программа | 3 |
| 1.3. Функциональное назначение | 3 |
| 1.4. Функциональные характеристики программного обеспечения | 3 |

1. ОБЩИЕ СВЕДЕНИЯ

1.1. Обозначение и наименование программы

Наименование программного обеспечения: Платформа для распознавания, синтеза и анализа речи SOVA Speech.

1.2. Языки программирования, на которых написана программа

Программный код написан на следующих языках: Python, JavaScript (React).

1.3. Функциональное назначение

ПО SOVA Speech содержит в себе модули распознавания (ASR – automatic speech recognition) и синтеза (TTS – text to speech) речи, позволяет выполнять полный спектр задач, связанных с обработкой и генерацией человеческой речи: распознавание речи в реальном времени с последующей обработкой полученного текста диалоговым процессором чат-бота, транскрибирование записей конференций, обработка типовых телефонных звонков в колл-центр, синтез ответных сообщений голосами выбранных дикторов, а также модификация сгенерированных записей путём контроля высоты тона и скорости.

ПО SOVA Speech может использоваться в системах автоматизации диалога – чат-ботах и голосовых ассистентах для предоставления необходимой информации клиентам контактных центров.

Конечными пользователями ПО могут являться как сотрудники предприятий или контактных центров (полное право доступа к информации), так и лица, обратившиеся на предприятие по каналам связи.

Возможна интеграция ПО в сторонние системы при помощи интерфейса API или применение в качестве его локального или сетевого модуля для использования в сетевой инфраструктуре.

1.4. Функциональные характеристики программного обеспечения

- Распознавание речи на русском языке для решения задачи перевода аудиозаписей с речью в текст.
- Генерация аудио с речью на русском языке дублирующего содержание поступившего на вход текста.
- Возможность установки на как серверное и локальное ПО (для возможности работы оффлайн без передачи чувствительных данных).
- Возможность оптимизации нейронных сетей для каждого случая под конкретные конечные устройства.
- Возможность применения распределённой серверной системы

вычислений при работе на слабых устройствах.

- API для интеграции ПО.
- Возможность регулирования высоты тона (увеличение на 50% и уменьшение на 25%) и скорости воспроизведения (ускорение/замедление до двух раз) сгенерированной речи.
- Возможность контроля ударений в сгенерированной речи на базе словарей.
- Выбор синтезируемого голоса: мужской и женский.