



ООО «ЗИАКС»
410056, Саратовская область, г. Саратов, ул. Советская, дом 61
+7 (495) 108-70-01
info@ziah.ru; ziah.ru

ZIAX TTS: Синтез речи

ОПИСАНИЕ ФУНКЦИОНАЛЬНЫХ ХАРАКТЕРИСТИК

СОДЕРЖАНИЕ

1. ВВЕДЕНИЕ.....	3
1.1 Описание решения.....	3
2. ФУНКЦИОНАЛЬНЫЕ ХАРАКТЕРИСТИКИ ПО.....	3
2.1 Цели и назначения.....	3
2.2 Ключевые функции.....	3
2.3 Область применения.....	3
2.4 Ключевые преимущества.....	4
3. ОПИСАНИЕ СИСТЕМНЫХ ТРЕБОВАНИЙ.....	5

1. ВВЕДЕНИЕ

Настоящий документ содержит описание ключевых функциональных характеристик решения ZIAX TTS: Синтез речи.

1.1 Описание решение

ZIAX TTS - высокоэффективный голосовой модуль, работающий на технологиях искусственного интеллекта, а именно - нейронной сети, предназначенный для синтеза речи в телефонном канале. Решение позволяет синтезировать голос в потоковом режиме на лету с частотой дискретизации 8000 Гц. Также система позволяет генерировать аудиофайлы в режиме офлайн.

ZIAX TTS: Синтез речи предназначен для работы следующих систем: голосовое управление, голосовой ввод текста, голосовой поиск. Кроме того, разработанное решение может быть применимо для IVR, голосовых чат-ботов, озвучивание стенограмм, диктовки текста.

2. ФУНКЦИОНАЛЬНЫЕ ХАРАКТЕРИСТИКИ ПО

2.1 Цели и назначения

Модуль представляет собой набор виртуальных контейнеров для Docker. На первом контейнере располагается загруженная обученная голосовая модель для синтеза речи, которая работает через API. На втором контейнере располагается сервер нормализации текста.

Есть возможность работать по протоколу:

- gRPC - система удалённого вызова процедур (RPC) с открытым исходным кодом. В качестве транспорта используется HTTP/2, в качестве языка описания интерфейса — Protocol Buffers.

2.2 Ключевые функции

ZIAX TTS: Синтез речи реализует следующие функции для конечного пользователя:

- Нормализация поступившего текст (предобработка)
- Синтез речи по заданным характеристикам

2.3 Область применения

Основные области применения:

- Синтез ответных фраз в голосовых ассистентах

- Синтез файлов для систем IVR
- Озвучивание текста

ZIAX TTS: Синтез речи предназначен для обеспечения синтеза речи, необходимым для работы вышеуказанных систем в реальном времени.

Речевые технологии - неотъемлемая составляющая современной цифровизации всех сфер деятельности компаний.

ZIAX TTS может быть использован в области IVR систем с голосовым управлением для быстрой замены приветствий и других аудиоданных.

ZIAX TTS будет демонстрировать эффективный результат при проведении массового обзвона партнеров, контрагентов или потребителей. Модуль позволяет оперативно донести идентичную информацию до большого количества адресатов. Также может выполняться обратная задача по сбору информации от неограниченного количества источников и анализ сведений в бизнес-структуре заказчика.

Система позволяет слабовидящим гражданам озвучивать текст голосом, что существенно упрощает данным гражданам жизнь.

ZIAX TTS будет полезен голосовым ассистентам, которые в последнее время стали более активно применяться в повседневной жизни. Для работы подобных голосовых роботов требуется качественная система синтеза речи робота, чтобы переводить текст в аудиоданные, особенно когда речь идет о персональных данных. Система подходит для решения данной задачи, поскольку ее установка осуществляется в защищенном периметре заказчика.

2.4 Ключевые преимущества

Среди ключевых преимуществ ZIAX TTS можно выделить следующее:

- **Работа на CPU**
Данная система адаптирована под работу на обычных ядрах вычислительного процессора, что существенно сокращает стоимость внедрения.
- **Высокое качество произношения речи**
Благодаря нормализации и предобработки текста все аббревиатуры, сокращения и числа переводятся в прописной вид с анализом контекста фразы для того, чтобы синтез был максимально логичным и человечным.
- **Брендирование голоса заказчика**
Есть возможность дообучать модели новым голосам.
- **В облаке или локально**
При наличии определенных требований к безопасности поддерживается гибкое развертывание системы на сторонних серверах.
- **Простая интеграция**
ZIAX TTS API работает на основе gRPC протокола, что позволяет быстро встраивать функции синтеза речи в любые кроссплатформенные приложения и сервисы.

3. Описание системных требований

Среди основных системных требований:

- Операционная система - Linux
- Контейнеризация – Docker
- Управление контейнерами - Docker-compose

Минимальные аппаратные требования включают в себя:

- CPU - 8
- Оперативная память - 8 Гб
- Размер жесткого диска - 40 Гб.

В среднем для одного потока синтеза нужно 1 CPU и 250 МБ оперативной памяти.