



ООО «ЗИАКС»
410028, Россия, г. Саратов, ул. Вольская, 28, пом. 2А
+7 (495) 108-70-01
info@ziax.ru; ziax.ru

**Описание функциональных характеристик модуля
распознавания речи
«ZIAХ ASR»**

СОДЕРЖАНИЕ

1. ВВЕДЕНИЕ	3
1.1 Описание решения.....	3
2. ФУНКЦИОНАЛЬНЫЕ ХАРАКТЕРИСТИКИ ПО	3
2.1 Цели и назначения.....	3
2.2 Ключевые функции.....	4
2.3 Область применения.....	4
2.4 Ключевые преимущества.....	5
3. ОПИСАНИЕ СИСТЕМНЫХ ТРЕБОВАНИЙ	6

1. ВВЕДЕНИЕ

Настоящий документ содержит описание ключевых функциональных характеристик решения ZIAX ASR, которое является модулем потокового распознавания речи, ориентированного на телефонный канал связи.

1.1 Описание решение

ZIAX ASR - высокоэффективный голосовой модуль, предназначенный для распознавания речи в телефонном канале. Решение обладает знанием характеристик передающей среды и помогает решать задачи шумоочистки и улучшения качества речевых сигналов, а также оценивать их пригодность для последующего использования в системах автоматического распознавания речи и голоса. Точность распознавания составляет 85%.

Голосовой модуль ZIAX ASR предназначен для работы следующих систем: голосовое управление, голосовые команды, голосовой ввод текста, голосовой поиск. Кроме того, разработанное решение может быть применимо для IVR, голосовых чат-ботов, написания стенограмм, диктовки текста, речевой аналитики.

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 распознавания речи, которая работает через API. На втором контейнере располагается MRCP сервер с плагином для работы в режиме реального времени в телефонном канале.

Есть возможность работать по 2 протоколам:

- gRPC - это система удалённого вызова процедур (RPC) с открытым исходным кодом. В качестве транспорта используется HTTP/2, в качестве языка описания интерфейса — Protocol Buffers.
- MRCP v2 - работает в связке с протоколом SIP, используется для потокового распознавания для АТС разного уровня (Avaya, Cisco, Asterisk, FreeSwitch и тд.)

2.2 Ключевые функции

Голосовой модуль распознавания речи ZIAX ASR реализует следующие функции для конечного пользователя:

- Распознавание речи по заданным характеристикам
- Указание таймеров для отслеживания состояния сессии
- Определение конца фразы
- Определение голосовой активности
- Постобработка распознанного текста

2.3 Область применения

Основные области применения:

- Голосовое управление
- Голосовые команды
- Голосовой ввод текста
- Голосовой поиск

Модуль распознавания речи ZIAX ASR предназначен для обеспечения точности распознавания речи, необходимой для работы вышеуказанных систем в реальном времени. Модуль учитывает следующие факторы: уровень окружающего шума, канал передачи речевой информации (GSM, ISDN, VoIP, аналоговые линии связи), темы распознаваемого разговора и лингвистические особенности языка.

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

ZIAX ASR может быть использован в области IVR систем с голосовым управлением для замены кнопок на голосовые команды. Сегодня широко распространены кнопочные IVR, и многие компании стремятся перейти на взаимодействие с клиентами с помощью голоса, а не с помощью DTMF сигналов. Кроме того, системы голосового самообслуживания обеспечат большее количество услуг и справочной информации, чем существующие IVR-меню на DTMF.

Модуль ZIAX ASR 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.

В речевой аналитике используется распознавание речи телефонных разговоров, поэтому ZIAX ASR может быть использован и для данного направления деятельности с целью осуществления более быстрого анализа качества работы операторов с клиентами.

Все больше сервис идентификации человеческой речи применяется в технических устройствах или для усовершенствования работы социальных инфраструктур. В качестве интеллектуального агента система оказывает неоценимую помощь, выполняя задачи голосового робота, помощника людей с ограниченными возможностями или офисного ассистента различного направления.

Последним новшеством в мире технологий «умных систем» выступает управление голосом, которое позволяет системе мгновенно реагировать на заранее записанную команду. Таким образом, модуль распознавания речи ZIAX ASR может стать хорошим и практичным решением данного вопроса.

Кроме того, все чаще используется голосовой поиск или диктовка текстов. Например, популярные поисковые системы интегрировали голосовых помощников, распознающих речевой запрос человека. Для этого достаточно внятно проговорить интересующий запрос, и система тут же покажет перечень релевантных сайтов. Внедрение модуля ZIAX ASR поможет пользователям значительно упростить использование различных гаджетов.

ZIAX ASR будет полезен голосовым ассистентам, которые в последнее время стали более активно применяться в повседневной жизни. Для работы подобных голосовых роботов требуется качественная система распознавания речи клиента, чтобы переводить текст в аудиоданные, особенно когда речь идет о персональных данных. Разрабатываемый в рамках Проекта модуль подходит для решения данной задачи, поскольку его установка осуществляется в защищенном периметре заказчика.

2.4 Ключевые преимущества

Среди ключевых преимуществ ZIAX ASR можно выделить следующее:

- **Работа в сложных акустических условиях**
Технологичность разработанного продукта ZIAX ASR позволяет осуществлять распознавание речи в самых сложных акустических условиях.
- **Высокая точность распознавания**
ZIAX ASR - высокоэффективный модуль распознавания речи, построенный с использованием нейронных сетей и глубокого обучения на массивах записей различного типа.
- **Настраиваемые речевые модели**
Возможность существенно повысить качество распознавания за счет адаптации языковой модели под предметную область,

специализированную лексику и терминологию, особенности дикции или акценты.

- **В облаке или локально**

При наличии определенных требований к безопасности поддерживается гибкое развертывание системы на сторонних серверах.

- **Простая интеграция**

ZIAX ASR API работает на основе HTTP-запросов, что позволяет быстро встраивать функции распознавания 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 распознавания нужно 1 CPU и 250 МБ оперативной памяти.