

Учреждение образования
“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ”
Кафедра интеллектуальных информационных технологий

Отчёт по практической работе №8-9
по дисциплине “Естественно-языковой интерфейс
интеллектуальных систем”

Выполнил: ст. группы 121701
Пашин Н. А.
Маевский В. Г.

Проверил: Крапивин Ю.Б.

Минск 2024

Цель работы: освоить на практике основные принципы создания систем анализа и распознавания речи. .

Задание:

Язык текста: английский, русский.

Предметная область: Сочинения по литературе

Ход работы

Библиотека Speech Recognition — это технология, которая позволяет преобразовывать аудио в текст (Automatic Speech Recognition, ASR). Она используется в различных приложениях, таких как голосовые помощники, системы автоматического ответа, инструменты для транскрипции и другие.

Современные веб-браузеры предоставляют встроенную поддержку для распознавания речи через Web Speech API, позволяя разработчикам интегрировать голосовые команды и взаимодействие с пользователями. Эта технология была представлена в сообществе W3C и продолжает развиваться для повышения точности и доступности.

Web Speech API Specification

Поддержка ASR определяется через ключевой объект `SpeechRecognition`. Как и в случае с синтезом речи, поддержка этого API существует в основном в браузерах на основе движка WebKit (например, Google Chrome, Opera), а также в других современных браузерах.

Основные возможности библиотеки Speech Recognition

- Распознавание речи: преобразование аудиопотока в текст.
- Поддержка различных языков: возможность распознавать и транскрибировать речь на нескольких языках.
- Обработка потокового аудио: работа с аудиофайлами и потоками в реальном времени.
- События управления: возможность отслеживать начало, завершение и ошибки распознавания.
- Кроссплатформенность: поддерживается в большинстве современных браузеров.

Преимущества использования Speech Recognition

- Простота интеграции: Web Speech API встроен в большинство современных браузеров, что упрощает его использование.
- Многоязычность: поддержка множества языков и акцентов, что расширяет аудиторию пользователей.
- Интерактивность: возможность создания интерактивных приложений с голосовым управлением.
- Улучшение доступности: помогает пользователям с ограничениями по зрению и моторными нарушениями взаимодействовать с устройствами.

Примеры использования

- Голосовые помощники: такие как Google Assistant, Siri, Alexa.
- Автоматические системы ответа: распознавание голосовых команд для обработки запросов пользователей.
- Образовательные приложения: инструменты для изучения языков, которые распознают речь пользователей.
- Транскрипция: автоматическое создание текстовых версий аудиозаписей в реальном времени.

Библиотека `pyttsx3` — это библиотека для синтеза речи в Python, которая позволяет преобразовывать текст в речь (Text-to-Speech, TTS). Она используется в различных приложениях, таких как голосовые помощники, обучающие программы и инструменты для людей с нарушениями зрения.

`pyttsx3` является кроссплатформенной библиотекой, которая не зависит от интернет-соединения и использует встроенные возможности синтеза речи операционной системы.

Основные возможности библиотеки `pyttsx3`

- Генерация речи: преобразование текстовых данных в аудиопоток, который можно воспроизводить.
- Поддержка различных голосов: возможность выбора между разными голосами, акцентами и языками, установленными на системе.

- Управление скоростью и высотой голоса: регулирование темпа речи и изменение тона для улучшения восприятия.
- События управления: отслеживание событий, таких как начало, пауза и завершение речи.
- Кроссплатформенность: поддерживается на Windows, macOS и Linux.

Преимущества использования pyttsx3

- Простота интеграции: легко интегрируется в различные Python-приложения.
- Многоязычность: поддерживает множество языков и акцентов, что делает его универсальным инструментом.
- Настраиваемость: возможность изменять параметры синтеза для улучшения пользовательского опыта.
- Работа без интернета: не требует интернет-соединения, так как использует локальные возможности синтеза речи.

Примеры использования

- Голосовые помощники: создание простых голосовых интерфейсов для взаимодействия с пользователем.
- Образовательные приложения: чтение вслух текстов для детей или изучающих иностранные языки.
- Доступность: помощь пользователям с ограничениями по зрению, предоставляя возможность прослушивания текстов.
- Автоматизация: использование в скриптах и инструментах для автоматизированного чтения сообщений, уведомлений и отчетов.