

УДК 371.31:376.33

Кинтонова Алия Жексембаевна, к.т.н., доцент,
Евразийского национального университета им.Л.Н.Гумилева,
Республика Казахстан, Астана

ЦИФРОВЫЕ АССИСТИВНЫЕ ТЕХНОЛОГИИ ДЛЯ АДАПТИРОВАНИЯ ЧАТА ДЛЯ СЛЕПЫХ И СЛАБОВИДЯЩИХ ОБУЧАЮЩИХСЯ В ВУЗАХ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ GPT

Аннотация. В статье рассматривается вопрос о необходимости использования цифровых технологий для адаптации их для слепых и слабовидящих людей в образовательных процессах.

В статье дается анализ технологий, позволяющих перевести аудио сообщения в текст и обратно. В статье говорится о необходимости интеграции существующих технологий для создания мощных и интерактивных голосовых приложений. Также в статье говорится о необходимости использовать международные стандарты, позволяющие сделать веб-содержание более доступным для людей с ограниченными возможностями.

Ключевые слова: стандарт Web Content Accessibility Guidelines, адаптированный чат с использованием технологии GPT, голосовой интерфейс, инклюзивное образование, ассистивные технологии, лица с ограниченными возможностями здоровья, цифровые ассистивные технологии для обучения.

Инклюзивная образовательная среда - это образовательная среда, позволяющая использовать специализированные образовательные ресурсы для передачи знаний и умений обучающимся с индивидуальными особенностями и возможностями. Сегодня важным вопросом инклюзивного образования является создание условий для обучения людей с ограниченными возможностями. Для повышения качества обучения детей с ограниченными возможностями, необходимо внедрять в процесс обучения методы обучения, технические, учебные и иные средства[1].

Цифровые ассистивные технологии включают в себя разнообразные средства, устройства, программные обеспечения и услуги для обучения людей с индивидуальными особенностями и возможностями.

Технологии перевода аудио сообщений в текст находятся в активной разработке и совершенствовании. Существуют различные методы и алгоритмы для этой задачи, включая:

Распознавание речи (Speech Recognition): Это технология, которая использует алгоритмы и нейронные сети для преобразования аудио записей в текст. Такие системы могут быть обучены на большом объеме данных, чтобы лучше понимать различные голоса и акценты[2].

Машинное обучение (Machine Learning): Модели машинного обучения, такие как рекуррентные нейронные сети (RNN) и трансформеры, используются для улучшения точности перевода аудио в текст. Google's Speech-to-Text и Amazon Transcribe - примеры таких сервисов.

Облачные API: Крупные компании предоставляют облачные сервисы для перевода аудио в текст, что делает эту технологию доступной разработчикам и компаниям.

Голосовые ассистенты: Голосовые ассистенты, такие как Siri, Google Assistant и Alexa, также используют подобные технологии для распознавания и обработки аудио команд.



Автоматические системы транскрипции: Существуют специализированные программы и сервисы, такие как Otter.ai и Rev.com, которые предоставляют услуги по автоматической транскрипции аудио файлов.

Эти технологии помогают сделать аудио информацию более доступной и обеспечивают широкий спектр применений, от субтитров в видео до создания записей звонков в текстовом виде[3].

Интеграция с API Alexa и OpenAI. Необходимо использовать международные стандарты, позволяющие сделать веб-содержание более доступным для людей с ограниченными возможностями. Стандарт Web Content Accessibility Guidelines (WCAG) - это набор мировых стандартов и рекомендаций, разработанных консорциумом W3C (World Wide Web Consortium), с целью сделать веб-содержание более доступным для людей с ограниченными возможностями. WCAG определяет требования к веб-сайтам, чтобы они были воспринимаемы, используемы и понимаемы людьми с разными видами инвалидности, включая слабовидящих, незрячих, глухих, глухослышащих и тех, у кого есть моторные ограничения.

WCAG состоит из четырех основных принципов, которые описывают общие цели доступности, и каждый принцип имеет свои рекомендации и проверяемые критерии:

Воспринимаемость (Perceivable): Цель - сделать информацию и пользовательский интерфейс воспринимаемыми для всех пользователями, независимо от их способностей восприятия. Примеры включают использование альтернативного текста для изображений и предоставление видео с субтитрами.

Используемость (Operable): Цель - обеспечить навигацию и интеракцию, которые доступны для всех. Это включает в себя использование клавиатуры вместо мыши, предоставление ясной структуры страницы и времени для взаимодействия.

Понимаемость (Understandable): Цель - сделать контент и управление понятными для всех пользователей. Рекомендации включают в себя предоставление четких инструкций и обеспечение предсказуемого поведения элементов управления.

Надежность (Robust): Цель - обеспечить, чтобы контент мог быть надежно интерпретирован различными технологиями, включая адаптивные и будущие технологии[4].

WCAG также имеет три уровня соответствия: А, АА и ААА, где уровень АА считается базовым уровнем доступности, а уровень ААА - более высоким.

Эти стандарты помогают разработчикам и владельцам веб-сайтов создавать более инклюзивное и доступное веб-содержание, что позволяет людям с разными специальными потребностями получать доступ к информации и сервисам в Интернете.

В частности, многие проблемы слепых людей можно решить с помощью адаптированного чата GPT

Да, адаптированный чат с использованием технологии GPT может быть очень полезным для слепых и людей с ограниченными зрительными возможностями. Вот несколько способов, как это может помочь:

Голосовой интерфейс: Специализированный чат GPT может быть интегрирован в голосовые ассистенты или приложения для чтения текста вслух. Это позволяет пользователям слушать и взаимодействовать с текстовой информацией, даже если они не могут видеть экран.

Автоматическая транскрипция: Чат GPT может автоматически транскрибировать аудио сообщения, записи разговоров или медиафайлы, делая их доступными для прослушивания.



Ответы на вопросы: Пользователи могут задавать вопросы в текстовом или голосовом формате, а адаптированный чат GPT может предоставлять информацию и отвечать на запросы в понятной форме.

Помощь в чтении и понимании текста: Чат GPT может помочь пользователям читать и интерпретировать текст, предоставляя разъяснения, переводы или резюме содержания.

Навигация по веб-сайтам и приложениям: Чат GPT может помогать пользователям перемещаться по веб-сайтам и приложениям, предоставляя инструкции и описания элементов интерфейса.

Адаптированный чат GPT с учетом потребностей пользователей с ограниченными зрительными возможностями может значительно улучшить их доступность к информации и коммуникации в цифровом мире[5].

Этапы разработки приложения «Голосовой помощник»:

Исследование и планирование: Исследование потребностей пользователей и определение ключевых функциональных требований к приложению. Этот этап также включает выбор подходящих технологий и инструментов разработки.

Проектирование: На этом этапе происходит разработка архитектуры приложения и проектирование пользовательского интерфейса, учитывающего потребности слепых и слабовидящих пользователей.

Разработка: Непосредственное программирование приложения. Это включает интеграцию с API Alexa и OpenAI, реализацию функций распознавания голоса и обработки естественного языка.

Тестирование: Проверка работоспособности приложения, исправление обнаруженных ошибок и недоработок.

Развертывание и поддержка: Запуск приложения на целевых устройствах, мониторинг его работы, исправление возникающих проблем и внесение улучшений на основе отзывов пользователей.

Для реализации приложения «Голосовой помощник» подключаем необходимые библиотеки:

```
// Подключаем необходимые библиотеки
const Alexa = require('ask-sdk-core'); // библиотека для работы с Alexa SDK
const { Configuration, OpenAIApi } = require('openai'); // библиотека для работы с OpenAI
const keys = require('./keys'); // файл, содержащий ключи API
```

Рисунок 1 – Подключение необходимых библиотек

Конфигурируем подключение к Open AI, рисунок 2:

```
// Конфигурируем подключение к OpenAI
const config = new Configuration({
  apiKey: keys.OPEN_AI_KEY // используем ключ API из файла keys
});
```

Рисунок 2 – Конфигурируем Подключение необходимых библиотек



Визуализация документа отражена на рисунке 3:

```
// Создаем экземпляр API OpenAI
const openai = new OpenAIApi(config);

// Функция для создания команды визуализации документа для Alexa
const createDirectivePayload = (aplDocumentId, tokenId = "documentToken") =>
  return {
    type: "Alexa.Presentation.APL.RenderDocument",
    token: tokenId,
    document: {
      type: "Link",
      src: "doc://alexa/apl/documents/" + aplDocumentId
    }
  }
};
```

Рисунок 3 – Визуализация документа

Инструменты для сбора и анализа обратной связи от пользователей с нарушениями зрения, такие как опросники и интервью, которые помогают определить, какие функции и улучшения наиболее важны для пользователей.

Создание нейронной сети для поиска информации в интернете - это достаточно сложный процесс, который включает в себя несколько этапов. Создание нейронной сети для поиска информации в интернете - это достаточно сложный процесс, который включает в себя несколько этапов.

1. Сбор данных: для начала необходимо собрать данные, которые будут использоваться для обучения нейронной сети. Эти данные могут быть представлены в виде текстовых документов, изображений, аудиофайлов или видео.

2. Предобработка данных: полученные данные необходимо предобработать, чтобы их можно было использовать для обучения нейронной сети. Это может включать в себя такие этапы, как очистка данных от шума, нормализация данных и т.д.

3. Разработка модели нейронной сети: затем нужно разработать модель нейронной сети, которая будет обучаться на предварительно подготовленных данных. Модель может состоять из различных слоев, включая входной слой, скрытые слои и выходной слой.

4. Обучение модели: после того как модель нейронной сети разработана, ее необходимо обучить на собранных и предобработанных данных. Обучение модели может занять значительное время и может потребовать мощных вычислительных ресурсов.

5. Тестирование и оптимизация модели: после того, как модель обучена, ее необходимо протестировать на новых данных, чтобы убедиться в ее эффективности. Если модель не проявляет достаточно хорошую точность, необходимо провести оптимизацию модели, что может включать в себя изменение параметров модели, изменение данных и т.д.

Создание чата для слепых может быть сложной задачей, но есть несколько ключевых моментов, которые нужно учесть:

1. Доступность интерфейса: чат должен быть доступен для использования с помощью специальных устройств для слепых, таких как Брайлевские дисплеи, голосовые синтезаторы, аудиоинтерфейсы и другие. Интерфейс должен быть разработан таким образом, чтобы пользователь мог легко получить доступ ко всем функциям чата.

2. Функциональность: чат для слепых должен иметь все необходимые функции, такие как отправка и получение сообщений, просмотр списка контактов, добавление новых контактов и т.д. Важно также предусмотреть возможность использования голосовых команд и других ассистентов, чтобы пользователь мог легко управлять чатом.



3. **Безопасность:** так как слепые люди могут иметь проблемы с визуальным восприятием, особое внимание следует уделить безопасности. Важно предусмотреть защиту от спама, фишинга и других атак на пользователей.

4. **Доступность на разных устройствах:** чат должен быть доступен для использования на разных устройствах, таких как персональные компьютеры, смартфоны, планшеты и другие. Это позволит пользователю выбрать устройство, которое наиболее удобно для него.

Кроме того, для создания чата для слепых необходимо учесть определенные особенности целевой аудитории и провести тестирование интерфейса с помощью группы слепых пользователей. Важно также учитывать международные стандарты в области доступности, такие как Web Content Accessibility Guidelines (WCAG), чтобы обеспечить максимально возможный уровень доступности и удобства использования для пользователей со слабым зрением[6].

Список литературы:

1. Кинтонова А.Ж., Габдрешов Г.Е., Кульбасов С. SMART технологии для слабовидящих. Сборник Международной научно-практической конференции «Актуальные проблемы теории, методологии и практики научной деятельности», 2022г., С.42.РИНЦ

2. Кинтонова А.Ж., Габдрешов Г.Е., Кульбасов С., Енсебаев Н.А., Енсебаев Т.М. ИННОВАЦИИ ДЛЯ ИНКЛЮЗИВНОГО ОБРАЗОВАНИЯ. V Международная научно-практическая конференция: «Научные дискуссии в условиях мирового кризиса: новые вызовы, взгляд в будущее».2022г. С. 92-95

3. Кинтонова А.Ж., Габдрешов Г.Е., Кульбасов С., Енсебаев Н.А., Енсебаев Т.М. ИССЛЕДОВАНИЯ В ОБЛАСТИ АССИСТИВНЫХ ТЕХНОЛОГИЙ. V Международная научно-практическая конференция: «Научные дискуссии в условиях мирового кризиса: новые вызовы, взгляд в будущее». 2022г. С.89-91

4. Кинтонова А.Ж., Габдрешов Г.Е., Кульбасов С., Енсебаев Н., Енсебаев Т. ТЕХНОЛОГИИ РАСПОЗНАВАНИЯ ЖЕСТОВ. Вестник научных конференций 2022 · N 9-2 (85), С.121-123. ISSN 2412-8988

5. Kintonova A., Gabdrashov G., Abyzbek A., Musayf M., Khaimulina D. ONLINE LEARNING TECHNOLOGIES. Norwegian Journal of development of the International Science ISSN 3453-9875. №108/2023. P.84-86 What Is Braille? | American Foundation for the Blind. afb.org

6. Maraim Alnefaie, Taif University, Srinivas Sampalli. Braille Enter: A Touch Screen Braille Text Entry Method for the Blind. December 2017. Procedia Computer Science 109:257-264. DOI:10.1016/j.procs.2017.05.349

