

Математическое и компьютерное моделирование

Научная статья

Статья в открытом доступе

УДК 004.42

doi: 10.30987/2658-6436-2025-1-28-35

АНАЛИЗ ВОПРОСОВ РАЗРАБОТКИ ГОЛОСОВОГО ПОМОЩНИКА ДЛЯ СПЕЦИАЛИЗИРОВАННЫХ БИБЛИОТЕК

Наталья Александровна Борсук¹, Данила Александрович Ковров²,
Татьяна Александровна Онуфриева³

^{1, 2, 3} Калужский филиал МГТУ им. Н.Э. Баумана, г. Калуга, Россия

¹ borsuk65@yandex.ru

² KovrovDanila@qmail.com

³ onufrievata@mail.ru

Аннотация. Рассмотрен вопрос об обслуживании читателей в специальных библиотеках для людей с ограничениями по зрению и о проблемах, с которыми сталкиваются такие читатели. Существующие стандартные автоматизированные системы обслуживания посетителей оказываются не полнофункциональными из-за особенностей таких людей. Предлагается разработка голосового помощника, встроенного в функционирующую в библиотеке автоматизированную систему обслуживания. Представлена модель работы приложения, учтены особенности методов перевода текстов в числовые форматы. Разработан алгоритм функционирования голосового помощника, включающий этапы авторизации и регистрации пользователя, уведомления о поступлении новых книг, формирования рекомендаций, уведомления о задолженностях. Приведена структура базы данных библиотечного фонда для слабовидящих людей со всеми связями, которые обеспечивают удобство сбора информации. Авторы проиллюстрировали реализацию основных функций приложения. Данное приложение разработано для специализированной библиотеки Н. Островского в городе Калуге для людей с ограничениями по зрению. Это позволит улучшить организацию хранения и предоставления информации, оптимизировать рабочие процессы, повысить доступность и удобство использования для пользователей с ограниченными возможностями.

Ключевые слова: библиотека для людей с ограничениями по зрению, голосовой помощник, PyCharm, Python, цифровизация

Для цитирования: Борсук Н.А. Ковров Д.А. Онуфриева Т.А. Анализ вопросов разработки голосового помощника для специализированных библиотек // Автоматизация и моделирование в проектировании и управлении. 2025. №1 (27). С. 28-35. doi: 10.30987/2658-6436-2025-1-28-35.

Original article

Open Access Article

ANALYSIS OF DEVELOPMENT ISSUES FOR A VOICE ASSISTANT FOR SPECIALIZED LIBRARIES

Natalia A. Borsuk¹, Danila A. Kovrov², Tatyana A. Onufrieva³

^{1, 2, 3} Bauman Moscow State Technical University (Kaluga Branch), Kaluga, Russia

¹ borsuk65@yandex.ru

² KovrovDanila@qmail.com

³ onufrievata@mail.ru

Abstract. The paper examines the issue of serving readers with visual impairments in specialized libraries, and the challenges these readers face. Existing standard automated customer service systems turn out to be less than fully

functional due to the population's specific needs. The work suggests developing a voice assistant integrated into the library's existing automated service system; presents an application operation model, taking into account the specific methods for converting text into numerical formats. The authors develop an algorithm for the voice assistant's operation, including the stages of user authorization and registration, notifications of new book arrivals, recommendation generation, and notifications of overdue materials; provide the structure of the library's database for visually impaired individuals, with all the relationships that ensure ease of information retrieval. The authors illustrate the implementation of the application core functions, developed for the N. Ostrovsky specialized library in the city of Kaluga, serving people with visual impairments. This will improve the organization of information storage and delivery; optimize workflows, and increase accessibility and application for users with disabilities.

Keywords: library for visually impaired people, voice assistant, PyCharm, Python, digitalization

For citation: Borsuk N.A., Kovrov D.A., Onufrieva T.A. Analysis of Development Issues for a Voice Assistant for Specialized Libraries. Automation and modeling in design and management, 2025, no. 1 (27). pp. 28-35. doi: 10.30987/2658-6436-2025-1-28-35.

Введение

Современный век – это век стремительной цифровизации всех сфер человеческой деятельности. Согласно указу Президента РФ от 21 июля 2020 г. № 474 определены главные целевые показатели цифровой информации в России. А именно:

- достижение цифровой зрелости ключевых отраслей экономики и социальной сферы, в том числе здравоохранения и образования, а также государственного управления;
- доля массовых социально значимых услуг, доступных в электронном виде – 95 %;
- доля домохозяйств, имеющих широкополосный доступ к Интернету – 97 %;
- увеличение вложений в российские IT-решения в четыре раза по сравнению с их размером в 2019 году [1].

Из представленных цифр видно, что значительная доля IT выросла в секторе образования и в секторе социальных услуг. Рассмотрим более подробно сферу образования, а именно получение новых знаний посредством библиотек. Люди всегда тянулись к получению какой-либо информации из различных источников. В современном мире почти не осталось населенных пунктов, которые не имели бы библиотеки. По данным Главного информационно-вычислительного центра (ГИВЦ) на 1 января 2023 года количество публичных библиотек МК РФ составило 40 906 единиц.

Изначально работа с читателем в библиотеке была рутинной и довольно кропотливой. Сначала необходимо было завести карточку читателя, присвоить ей порядковый номер, выдать посетителю читательский билет, найти в довольно большой картотеке информацию об издании, которое хотел бы взять читатель, отметить в карточке читателя взятую книгу и т.д. В таком виде библиотечная система работала довольно длительное время.

Все поменялось с общей цифровизацией человеческой деятельности. В библиотеках используются информационные системы учета и работы с читателями, осуществляется оцифровка наиболее ценных книг, создается единая база электронных книг, доступ к которой может быть осуществлен в режиме 24/7 из любой точки земного шара и многое другое.

Цифровизация в библиотеках – это процесс оцифровки книг, журналов и других источников информации, а также создание мультимедийных ресурсов. Это позволяет: улучшить организацию хранения и предоставления информации, сохранить культурное наследие, многократно использовать цифровые копии, оптимизировать рабочие процессы.

Однако, когда речь идет о специальных библиотеках для людей с ограничениями по зрению и о проблемах, с которыми сталкиваются такие читатели, стандартные, автоматизированные существующие системы обслуживания посетителей оказываются не полнофункциональными из-за уникальных особенностей данного рода людей. По данным аналитического бюро информационного портала *GxP News* исследователи считают, что к 2025 году число слабовидящих в России увеличится как минимум на 5,1 %, до 6,2 млн человек [2]. Доступ к любой информации данным людям крайне необходим. Поэтому возникает задача облегчить доступ слабовидящих к библиотечному фонду.

В результате плодотворной длительной работы по автоматизации работы библиотеки для людей с ограничениями по зрению им. Н. Островского в городе Калуге создана база данных

читателей, работает каталогизация библиотечного фонда, реализована система сбора и анализа статистических данных, облегчающая труд сотрудников библиотеки, автоматизирован вход слабовидящих читателей в библиотечную систему путем считывания штрихкода на читательском билете, разработан сайт самой библиотеки.

На данном этапе возникла задача разработки голосового помощника, облегчающего вход в библиотечную систему по *ID* читателя, голосового вывода имеющихся у него на руках книг, выдачи рекомендаций по жанрам читаемых книг.

Проектирование объекта

Для достижения поставленной цели предлагается интегрировать в существующую автоматизированную систему обслуживания пользователей библиотеки голосового помощника, используя технологии распознавания и синтеза речи, а также диалоговые модели. Это позволит упростить поиск информации, предоставлять справочную поддержку и рекомендации пользователям, а также автоматизировать некоторые процессы, такие как бронирование книг и регистрация.

Разработано приложение голосового ассистента – программный комплекс, в котором основным объектом управления является информация о читателе. Модель работы приложения представляет собой сочетание существующих методов рельефно-графического управления и взаимодействия с речевыми технологиями - системами синтеза и распознавания речи.

Модель управления представлена на рис. 1, где *id* – идентификатор читателя, читательский билет, который авторизует пользователя, *u* – речевой запрос поиска, с учетом коррекции, построенной на обратной связи и аудиодискрипции.

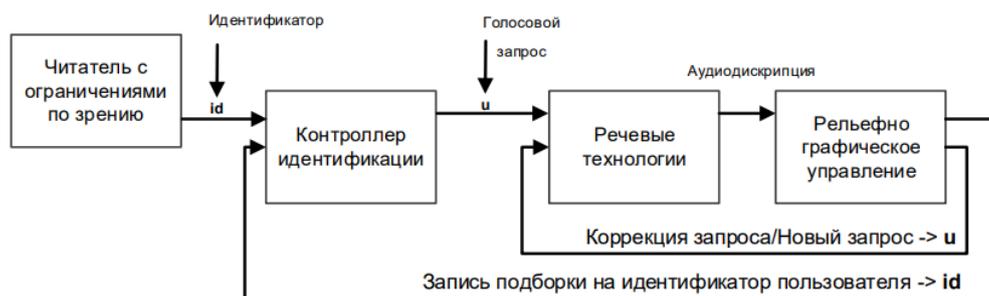


Рис. 1. Модель управления
Fig. 1. Management model

Для разработки приложения голосового ассистента (помощника) необходимо исследовать способы обеспечения доступа к информации для пользователей с ограниченными возможностями зрения, такие как технологии речи (включая синтез и распознавание) [3] и рельефное графическое управление [4]. Синтез речи позволяет преобразовывать текстовую информацию в аудиоформат, обеспечивая альтернативу тексту и предоставляя звуковую обратную связь. Технологии распознавания речи позволяют пользователю управлять приложением и выполнять поиск с помощью голосовых команд. При разработке модели распознавания человеческой речи разумнее использовать методы векторизации или обработки естественного языка *NLP*.

Обработка естественного языка, также известная как *NLP* (*Natural Language Processing*), занимается применением алгоритмов машинного обучения для работы с текстовыми данными. Как правило, модели машинного обучения преобразуют текстовую информацию в числовые представления, которые могут быть обработаны и проанализированы. [5, 6].

В процессе обработки естественного языка текстовая информация подвергается сегментации на отдельные смысловые единицы, именуемые токенами. Это могут быть буквы, слова, словосочетания, предложения, абзацы и иные текстовые элементы. Наиболее распространенным подходом является разделение текста на отдельные лексические единицы – слова. Сово-

купность всех уникальных токенов формирует словарный запас, который может быть упорядочен в алфавитном порядке. В области обработки естественного языка также используются понятия «документ» и «корпус». Документ представляет собой набор токенов, относящихся к одной смысловой единице, например, предложение, комментарий или сообщение пользователя. Корпус, в свою очередь, является совокупностью всех документов, рассматриваемых в рамках анализа.

На практике используются 4 наиболее популярных метода для перевода текстов в числовые форматы.

Прямое кодирование. Прямое кодирование (*One-Hot Encoding*) является простейшим подходом, при котором каждому уникальному слову в словаре сопоставляется бинарный вектор, где единица соответствует данному слову, а остальные элементы равны нулю. Такое представление не учитывает взаимосвязи между словами.

Bag of words. Метод мешка слов (*Bag of Words*) основан на подсчете частоты встречаемости каждого слова в документе, формируя вектор, где каждый элемент соответствует количеству вхождений конкретного слова. Данный подход игнорирует порядок слов и их семантические связи.

TF-IDF. *TF-IDF* (*Term Frequency – Inverse Document Frequency*) является развитием мешка слов, где каждому слову присваивается вес, отражающий его важность в документе с учетом распространенности в корпусе. Это позволяет выделять наиболее значимые слова.

Стоит отметить, что *TF* считается для токенов документа, тогда как *IDF* – токенов всего корпуса. Кроме того, *IDF* можно считать и другими способами, например, в *Python*-библиотеке *Scikit-learn* этот параметр гибко регулируется.

Word embeddings. Четвертый метод кодирования *word embeddings* – векторное представление слов. Векторы можно складывать, вычитать, сравнивать. Например, при сравнении близости слов «мужчина», «мальчик», «девочка», очевидно, что «мужчина» и «мальчик» стоят ближе друг к другу. Именно так и работают *word embeddings*, представляя слова (токены) в векторы. Самой распространенной реализацией векторного представления слов является *Word2vec*.

Исходя из анализа, было решено использовать метод *Word embeddings*. Метод облегчает извлечение смысла из текста и обработку семантических отношений, а также снизить размерность данных, что ускоряет обучение моделей и уменьшает потребность в вычислительных ресурсах по сравнению с другими методами обработки текста. Более того *Word embeddings* способен улавливать семантический контекст слов, что делает их более информативными и эффективными для различных задач, включая классификацию, кластеризацию и перевод текста.

Разработанная система обладает функционалом как для читателей, так и для сотрудников. Посредством голоса пользователь запрашивает выполнение программой заданных функций, программа озвучивает результаты своей работы и всю информацию, которая отображается на экране.

Помимо распознавания человеческой речи голосовой помощник для библиотечной системы имеет возможность сбора информации и предоставления её пользователю [7]. Используя подключение к базе данных, ассистент, по запросу пользователя, выводит информации о посетителе. На рис. 2 приведена структура базы данных библиотечного фонда для слабовидящих людей со всеми связями, которые обеспечивают удобство сбора информации.

Голосовой ассистент запрашивает номер читательского билета посетителя, который соответствует *id* пользователя. На основе обработанного запроса программа начинает сбор информации о пользователе из базы данных, которая будет отображена на терминале и также озвучена. При этом предоставляется только нужная информация (ФИО, дата регистрации в системе, а также последние взятые книги), ненужная информация игнорируется. Более того, при анализе формируется информация о задолженностях пользователя с соответствующим напоминанием. Такой подход облегчает работу в библиотечной системе путем оперативного сбора и предоставления необходимой информации.

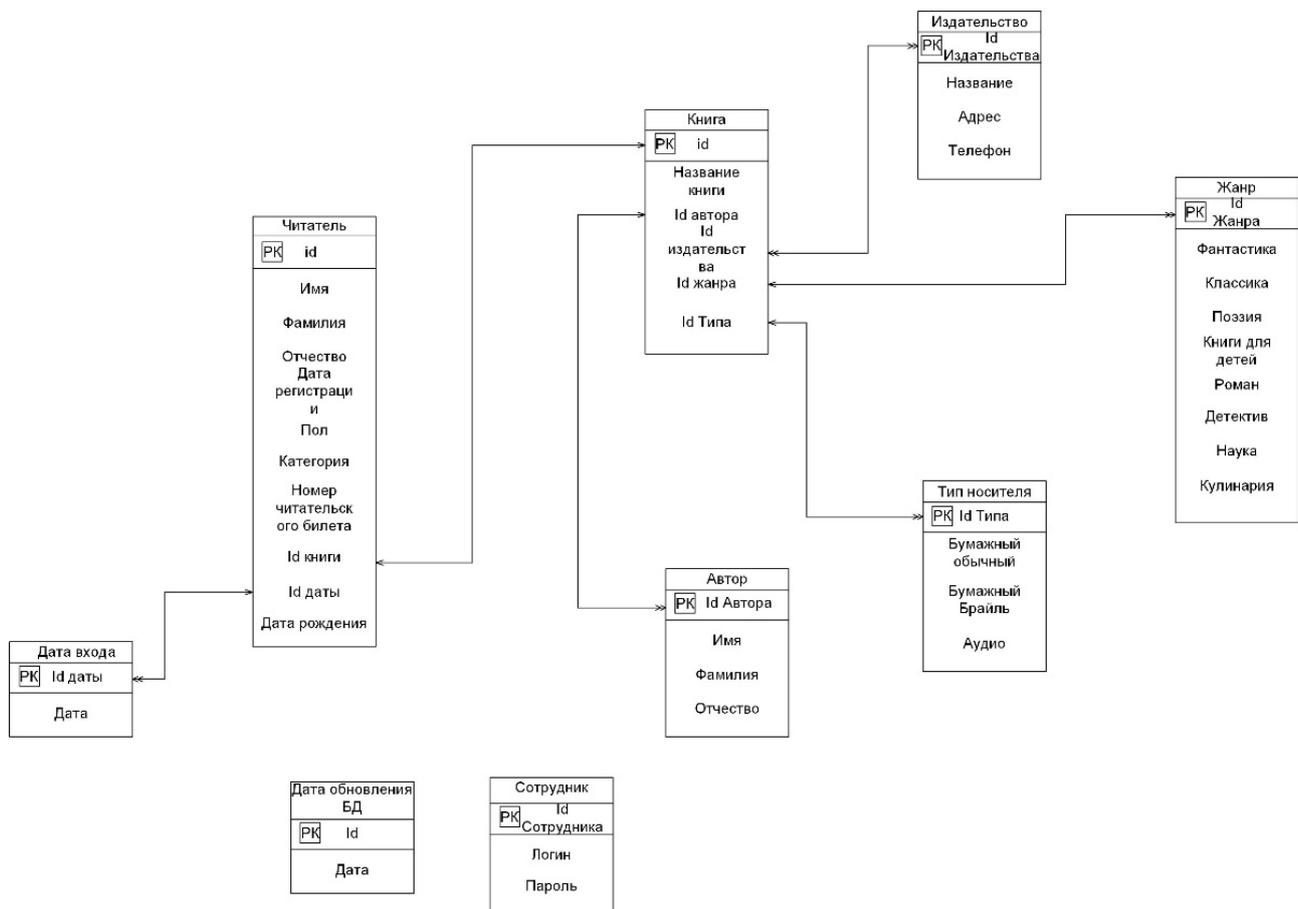


Рис. 2. Структура базы данных библиотеки для слабовидящих людей
Fig. 2. The structure of the library database for visually impaired people

Однако если по введенной информации о номере читательского билета не было ничего найдено, или же пользователь еще не зарегистрирован в системе, программа предлагает сделать это. Запрашивая все необходимые данные, голосовой помощник уже не собирает информацию о пользователе, а заносит её в базу данных с автоматическим присвоением индивидуального *id* и номера читательского билета. Данный процесс обеспечивает быстрдействие в рабочем процессе, исключая работу с картотекой.

В дополнении к сбору информации, функцией, обеспечивающей удобную работу с библиотечным фондом, является функция создания рекомендаций. Для реализации этого процесса на основе базы данных библиотеки проходит обучение нейронной сети.

Нейронной сети предоставляется информация о всех книгах, взятых пользователем за весь период с момента его регистрации в библиотечной системе. В основном для формирования рекомендаций берется информация о часто встречающихся авторах и жанрах прочитанного. На основе полученных данных нейронная сеть формирует список по популярности книг у пользователя и формирует новый список книг, которые могут заинтересовать посетителя. При создании данного списка игнорируются уже прочитанные произведения и произведения, которые меньше всего берутся другими пользователями. Кроме того, рекомендации создаются на основе типа носителя информации, который подходит посетителю. Алгоритм функционирования системы представлен на рис. 3.

Блоки 1...15: функции авторизации и регистрации пользователя, блоки 16...19: функция уведомления о поступлении новых книг, блоки 20...25: функция формирования рекомендаций, блоки 26, 27: функция уведомления о задолженностях, блок 29: конец сеанса с пользователем.

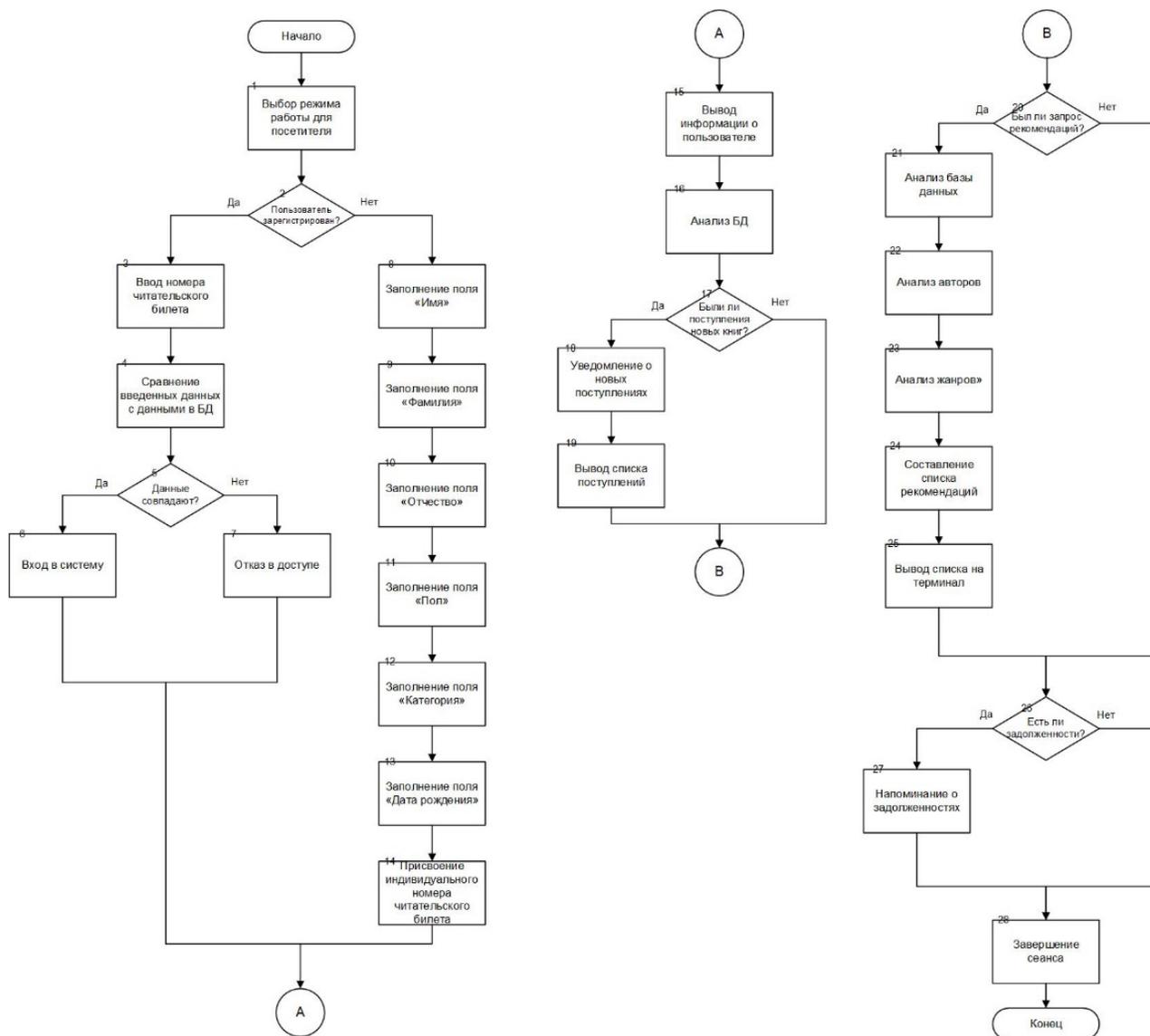


Рис. 3. The algorithm of the system functioning
 Fig. 3. The structure of the library database for visually impaired people

Примеры реализации функции распознавания и обработки речи и авторизации представлены на рис. 4, 5.

```
# Реализация ввода запросов от пользователя голосом
7 usages
def get_audio():
    r = sr.Recognizer()
    with sr.Microphone() as source:
        audio = r.listen(source)
        said = ""
        try:
            said = r.recognize_google(audio, language="ru-RU")
            print("Вы сказали:", said)
        except Exception as e:
            print("Извините, не могу распознать вашу речь")
    return said.lower()
```

Рис. 4. Код функции распознавания речи
 Fig. 4. The code of the speech recognition function

```

# Основная логика программы
speak("Привет! Зарегистрированы ли вы?")
print("Программа: Привет! Зарегистрированы ли вы?")
response = get_audio()
if "да" in response:
    speak("Назовите номер читательского билета")
    print("Назовите номер читательского билета")
    user_ticket_number = get_audio()
    sql = "SELECT * FROM users WHERE ticket_number = %s"
    mycursor.execute(sql, (user_ticket_number,))
    user = mycursor.fetchone()
    # Проверка наличия пользователя в базе данных
    if user:
        user_info = f"Имя: {user[1]}, Фамилия: {user[2]}, Дополнительная информация: {user[3]}"
        speak(user_info)
        print("Программа:", user_info)
    else:
        speak("Пользователь с таким номером читательского билета не найден.")
        print("Программа: Пользователь с таким номером читательского билета не найден.")
else:
    speak("Введите ваше имя, пожалуйста.")
    print("Программа: Введите ваше имя, пожалуйста.")
    user_name = get_audio()
    speak("Введите вашу фамилию, пожалуйста.")
    print("Программа: Введите вашу фамилию, пожалуйста.")
    user_surname = get_audio()
    speak("Введите ваше отчество, пожалуйста.")
    print("Программа: Введите ваше отчество, пожалуйста.")
    user_patronymic = get_audio()

```

Рис. 5. Код функции авторизации
Fig. 5. The code of authorization function

Помимо читателя программой могут пользоваться также сотрудники библиотеки. Благодаря функциям голосового ассистента сотрудники будут иметь возможность собирать статистику о половой принадлежности читателей, о наиболее популярных жанрах или авторах. По собранной статистике сотрудники смогут спланировать работу библиотечного фонда, исходя из предпочтений читателей.

Данная система была реализована с использованием высокоуровневого языка программирования *Python*. *Python* – это язык программирования, который отличается простотой в освоении и обширным набором библиотек, охватывающих широкий спектр задач, включая распознавание речи, синтез речи, машинное обучение, искусственный интеллект и обработку естественного языка. Более того, *Python* предоставляет богатый выбор библиотек, которые позволяют выполнять распознавание и синтез речи, а также проводить машинное обучение.

В качестве среды разработки использовался *PyCharm*. *PyCharm* – это интегрированная среда разработки (*IDE*) для языка программирования *Python*, разработанная компанией *JetBrains*. *PyCharm* является удобным редактором кода с подсветкой синтаксиса, мощными инструментами отладки, системой автодополнения кода, интеграцией с системами контроля версий, а также поддержкой различных фреймворков, что облегчает процесс разработки программы.

Как и любая информационная система, голосовой ассистент обладает перспективами развития. К таким перспективам можно отнести расширение функционала, чтобы не только посетитель, но сотрудник библиотеки мог работать с программой. Используя функционал голосового помощника, сотрудник может анализировать не только список пользователей, зарегистрированных в системе, но и формировать статистику о наиболее популярных книгах и типах носителей информации среди посетителей. Данное расширение функционала значительно оптимизирует работоспособность персонала, снижая время и силы, необходимые для выполнения даже самой простой задачи.

Заключение

Голосовой помощник – удобный инструмент для автоматизации производства. Его интеграция в библиотечные системы требует комплексного подхода, включающего разработку голосового интерфейса, оптимизацию доступности и удобства использования для пользователей с ограниченными возможностями по зрению, а также интеграцию с существующими информационными ресурсами и каталогами. Важно также уделить внимание аспектам конфиденциальности и безопасности данных при использовании голосового помощника в библиотечной среде.

Список источников:

1. <https://www.garant.ru/article/1605871>
2. <https://asi.org.ru/news/2022/11/28/issledovately-pochitali-skolko-slabovidyashhih-i-slaboslyshashhih-lyudej-budet-v-rossii-k-2025-godu>
3. Крысин И.А., Онуфриева Т.А. Структура электронного (тиффо) каталога для библиотек, в том числе специализированных // Международный научный журнал «Символ науки». – 2016. – №2-2. – 58 с.
4. Web Content Accessibility Guidelines (WCAG) 2.0 [Электронный ресурс] URL: <https://www.w3.org/Translations/WCAG20-ru/>
5. Паттерсон Д., Гибсон А. Глубокое обучение с точки зрения практика. – М.: ДМК Пресс, 2018. – 418 с.
6. Цитильский А.М., Иванников А.В., Рогов И.С. NLP. Обработка естественных языков // Научно-образовательный журнал для студентов и преподавателей «StudNet». – 2020. – №6. – С. 467-474.
7. Каримов А.Ю. Анализ алгоритмов голосовых помощников // Международный журнал «Вестник науки». – 2019. – №5 (14). – Т. 4. – С. 492-495.

Информация об авторах:

Борсук Наталья Александровна

кандидат технических наук, доцент, кафедра «Информационные системы и сети», Калужский филиал МГТУ им. Н.Э. Баумана

Ковров Данила Александрович

студент 4 курса, кафедра «Информационные системы и сети», Калужский филиал МГТУ им. Н.Э. Баумана

Онуфриева Татьяна Александровна

кандидат технических наук, доцент, кафедра «Информационные системы и сети», Калужский филиал МГТУ им. Н.Э. Баумана

References:

1. [Internet]. Available from: <https://www.garant.ru/article/1605871>
2. [Internet]. Available from: <https://asi.org.ru/news/2022/11/28/issledovately-pochitali-skolko-slabovidyashhih-i-slaboslyshashhih-lyudej-budet-v-rossii-k-2025-godu>
3. Krysin I.A., Onufrieva T.A. Structure of the Electronic (Typhlo) Catalog for Libraries, Including Specialized Ones. Symbol of Science: International Scientific Journal. 2016;2-2:58.
4. Web Content Accessibility Guidelines (WCAG) 2.0 [Internet]. Available from: <https://www.w3.org/Translations/WCAG20-ru/>
5. Patterson D., Gibson A. Deep Learning. A Practitioner's Approach. Moscow: DMK Press; 2018.
6. Tsitulytsky A.M., Ivannikov A.V., Rogov I.S. NLP. Natural Language Processing. StudNet. 2020;6:467-474.
7. Karimov A.Yu. Analysis of Voice Assistant Algorithms. Science Bulletin. 2019;5(14)-4:492-495.

Information about the authors:

Borsuk Natalia Alexandrovna

Candidate of Technical Sciences, Associate Professor at the Department of Information Systems and Networks of Bauman Moscow State Technical University (Kaluga Branch)

Kovrov Danila Alexandrovich

4th year student at the Department of Information Systems and Networks of Bauman Moscow State Technical University (Kaluga Branch)

Onufrieva Tatyana Alexandrovna

Candidate of Technical Sciences, Associate Professor at the Department of Information Systems and Networks of Bauman Moscow State Technical University (Kaluga Branch)

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации.

Contribution of the authors: the authors contributed equally to this article.

Авторы заявляют об отсутствии конфликта интересов.

The authors declare no conflicts of interests.

Статья поступила в редакцию 26.04.2024; одобрена после рецензирования 18.12.2024; принята к публикации 19.02.2024.

The article was submitted 26.04.2024; approved after reviewing 18.12.2024; accepted for publication 19.02.2024.

Рецензент – Малаханов А.А., кандидат технических наук, доцент, Брянский государственный технический университет.

Reviewer – Malakhanov A.A., Candidate of Technical Sciences, Associate Professor, Bryansk State Technical University.