

Министерство науки и высшего образования  
Российской Федерации  
ФГБОУ ВО  
«Уральский государственный горный университет»  
Кафедра информатики

УТВЕРЖДАЮ

Зав. кафедрой информатики

 А.В. Дружинин

“06” июня 2024г.

ИССЛЕДОВАНИЕ И РАЗРАБОТКА СИСТЕМЫ  
РАСПОЗНАВАНИЯ И РАЗМЕТКИ ГОЛОСОВЫХ  
ЗАПИСЕЙ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ

## Выпускная квалификационная работа магистра

### Пояснительная записка

Руководитель ВКР,  
кандидат технических наук



(подпись, дата)

Е.А. Волкова

(и. о. ф.)

Разработал студент



(подпись, дата)

А.В. Осинцев

(и. о. ф.)





Группа

ИНФ.м-22

Екатеринбург, 2024

## СОДЕРЖАНИЕ

Аннотация .....	5
Abstract .....	6
Введение .....	8
1 АНАЛИЗ ПРОБЛЕМАТИКИ.....	11
1.1. Характеристика проблемной ситуации/области.....	11
1.2. Аналитический обзор существующих решений .....	14
1.3. Основные положения, выносимые на защиту .....	19
Вывод по главе 1 .....	19
2 ПРЕДЛАГАЕМОЕ РЕШЕНИЕ .....	22
2.1 Системный анализ проблемной ситуации/области .....	22
2.1.1 Математическая модель объекта управления.....	22
2.1.2 Концептуальное моделирование .....	23
2.2 Функциональная модель.....	30
2.3 Имитационная модель.....	35
2.4 Формирование требований к проектируемой системе .....	38
Вывод по 2 главе.....	54
3 ПРОВЕДЕНИЕ ЭКСПЕРИМЕНТОВ И АНАЛИЗ РЕЗУЛЬТАТОВ.....	57
3.1 Проектирование информационной системы .....	57
3.1.1 Выбор архитектуры информационной системы .....	57
3.2 Инфологическое проектирование .....	58
3.3 Проектирование интерфейса.....	61
3.4 Программная реализация проекта.....	64
3.4.1 Выбор средств реализации .....	64
3.5 Проведение экспериментов и анализ результатов.....	67
Вывод по 3 главе.....	71
Заключение.....	73
Список использованных источников .....	75

					Д 09.04.01 2022 22286159 ВКРМ			
Должность	Фамилия	Подпись	Дата	Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей	Лит.	Масса	Масш.	
Студент	Осинцев А.В.		06.06.24					
Руководит.	Волкова Е.А.		06.06.24					
Консульт.								
Н. Контроль	Волкова Е.А.		06.06.24		Лист 2	Листов 76		
Рецензент					ИНФ.м-22			
Зав.кафедр	Дружинин А. В.		06.06.24	Содержание	УГТУ, кафедра информатики			

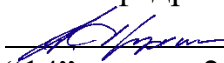
Министерство образования и науки  
Российской Федерации  
ФГБОУ ВО

«Уральский государственный горный университет»

Факультет инженерно-экономический  
Кафедра информатики  
Специальность Информатика и вычислительная техника (направление подготовки «Анализ больших данных и машинное обучение»)

УТВЕРЖДАЮ

Зав. кафедрой информатики

 А. В. Дружинин  
«14» апреля 2024 г.

**ЗАДАНИЕ  
ПО ДИПЛОМНОМУ ПРОЕКТИРОВАНИЮ**

Студенту Осинцев Антон Владимирович  
(фамилия, имя, отчество полностью)

**1 Тема ВКР** Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей

Утверждена приказом по университету № 872/3е от «14» апреля 2024г.

**2 Срок сдачи студентом ВКР:**

Сдача пояснительной записки, видео и презентации для прохождения процедуры допуска до:

«01» июня 2024г.

Сдача итоговой версии пояснительной записки (с отсканированными подписями и рецензиями), видео и презентации для процедуры дипломирования до:

«06» июня 2024г.

**Ф. И. О. руководителя дипломного проекта (работы)** Волкова Е.А.  
ученая степень кандидат технических наук, ученое звание доцент

Дата выдачи задания «14» апреля 2024г.

**Руководитель ВКР**  Е.А. Волова

**Задание на проектирование ВКР получил**  А.В. Осинцев

## АННОТАЦИЯ

					Д 09.04.01 2022 22286159 ВКРМ	Лист
		Осинцев А.В.		06.06.24		
Студ.		Фамилия	Подп.	Дата		4

## АННОТАЦИЯ

**Название ВКР:** Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей

**ФИО автора:** Осинцев Антон Владимирович.

**Научный руководитель:** кандидат доктора наук Волкова Евгения Алексеевна.

**Ключевые слова:** транскрибация, теги, ключевые слова, анализ, speech-to-text, оптимизация, golang, kubernetes, postgresSQL

**Цель исследования:** повышение эффективности анализа записей разговоров, выявление процента соответствий заданным критериям в совершенных звонках.

**Научная новизна** работы заключается в реализации комплексной системы, которая автоматизирует разметку и распознавание голосовых данных, интегрируется с другими корпоративными системами, также вводит новые методологии оценки работы, которые ранее не применялись.

Выпускная квалификационная работа магистра состоит из аннотации, введения, трёх разделов и заключения. Работа содержит 76 страниц, на которых размещены 32 рисунка, 3 таблицы, 23 литературных источника.

В работе выполнено исследование методологии автоматизации разметки и распознавания голосовых данных, спроектирована база данных, интерфейс, прототип системы анализа голосовых записей, дано техническое обоснование использованию моделей распознавания речи, сформулированы требования к системе.

Спроектировано, разработано и представлено программное обеспечение, представляющее собой прототип автоматизированной информационной системы для анализа голосовых данных.

Внедрение предлагаемого решения позволит обеспечить повышение эффективности анализа записей разговоров, улучшение точности обработки голосовых данных, снижение количества ошибок в работе сотрудников и повышение удовлетворенности клиентов.

### **Список трудов автора по тематике работы:**

Осинцев А.В., Волкова Е.А. ИССЛЕДОВАНИЕ ТЕХНИЧЕСКИХ АСПЕКТОВ ПОИСКА КЛЮЧЕВЫХ СЛОВ В ТЕКСТОВОМ ПРЕДСТАВЛЕНИИ // Уральская горная школа регионам: сборник докладов международной научно-практической конференции – Екатеринбург: Изд-во УГГУ, 2024.

Осинцев А.В., Волкова Е.А. ИССЛЕДОВАНИЕ ВОПРОСА МОДЕЛЕЙ РАСПОЗНАВАНИЯ РЕЧИ // Уральская горная школа регионам: сборник докладов международной научно-практической конференции – Екатеринбург: Изд-во УГГУ, 2024.

## ABSTRACT

**Title of the Master's Thesis:** Research and Development of a Voice Recognition and Annotation System Using Neural Networks

**Author:** Anton Vladimirovich Osintsev

**Scientific Supervisor:** Candidate of Doctoral Sciences, Evgeniya Alekseevna Volkova

**Keywords:** transcription, tags, keywords, analysis, speech-to-text, optimization, golang, Kubernetes, PostgreSQL

**Research goal:** to enhance the efficiency of conversation record analysis, identifying the percentage of compliance with predefined criteria in completed calls.

**The scientific novelty of the work** lies in the implementation of a comprehensive system that automates the annotation and recognition of voice data, integrates with other corporate systems, and introduces new methodologies for performance evaluation that were not previously applied.

The master's thesis consists of an abstract, introduction, three chapters and a conclusion. The work comprises 76 pages, featuring 32 figures, 3 tables and 23 literary sources.

The research covers the methodology for automating the annotation and recognition of voice data, the design of a database, interface, and a prototype system for analyzing voice recordings. It provides a technical justification for using speech recognition models and formulates the system requirements.

The designed, developed, and presented software is a prototype of an automated information system for analyzing voice data.

The implementation of the proposed solution will enhance the efficiency of conversation recording analysis, improve the accuracy of voice data processing, reduce employee error rates, and increase customer satisfaction.

### **Author's work references:**

Osintsev A.V., Volkova E.A. Research of Technical Aspects of Keyword Search in Text Representation // Ural Mining School for Regions: Collection of Papers from the International Scientific and Practical Conference – Yekaterinburg: URSMU Publishing House, 2024.

Osintsev A.V., Volkova E.A. Research on Speech Recognition Models // Ural Mining School for Regions: Collection of Papers from the International Scientific and Practical Conference – Yekaterinburg: UGGU Publishing House, 2024.

## ВВЕДЕНИЕ

					Д 09.04.01 2022 22286159 ВКРМ	Лист
		Осинцев А.В.		06.06.24		
Студ.		Фамилия	Подп.	Дата		7

## **ВВЕДЕНИЕ**

### **Актуальность проблемы**

С развитием информационных технологий и автоматизации бизнес-процессов контакт-центры сталкиваются с необходимостью эффективного анализа и обработки голосовых записей, которые являются важным инструментом для взаимодействия с клиентами. Однако, ручной анализ таких записей требует значительных временных и человеческих ресурсов, а также подвержен ошибкам. Разработка системы распознавания и разметки голосовых записей с использованием нейронных сетей представляет собой перспективное направление для контакт-центров, стремящихся улучшить качество обслуживания и эффективность работы. Несмотря на вызовы, связанные с качеством данных и сложностью моделей, преимущества автоматизации анализа голосовых записей делают этот подход крайне востребованным и актуальным в современном бизнесе. В связи с этим, актуальной становится проблема разработки и внедрения систем распознавания и разметки голосовых данных с использованием современных технологий, таких как нейронные сети.

**Объектом исследования** является транскрибация, обработка и анализ голосовых записей

**Предметом исследования** является разработка и апробация системы распознавания и разметки голосовых записей с использованием нейронных сетей.

**Цель исследования:** повышение эффективности анализа записей разговоров, выявление процента соответствий заданным критериям в совершенных звонках

### **Задачи исследования:**

- Провести анализ существующих методов распознавания и разметки голосовых записей.
- Разработать архитектуру системы распознавания и разметки голосовых данных на основе нейронных сетей.
- Реализовать и протестировать разработанную систему на реальных данных контакт-центров.
- Провести сравнительный анализ эффективности разработанной системы с существующими методами анализа голосовых записей.

**Научная новизна** работы заключается в реализации комплексной системы, которая автоматизирует разметку и распознавание голосовых данных, интегрируется с другими корпоративными системами, также вводит новые методологии оценки работы, которые ранее не применялись.



**Практические результаты** работы заключаются в разработке и апробации эффективной системы распознавания и разметки голосовых данных, которая позволит сократить время и ресурсы, затрачиваемые на анализ разговоров в контакт-центрах, а также повысить точность и качество этого процесса.

**Апробация темы** заключается в проведении двух экспериментов:

- 1) эксперимент для подтверждения эффективности внедрения автоматизированной системы
- 2) Тестирование внедренной автоматизированной системы на факт своевременного выявления и устранения недостатков в коммуникации с клиентом

**Публикации автора**

1. Осинцев А.В., Волкова Е.А. ИССЛЕДОВАНИЕ ТЕХНИЧЕСКИХ АСПЕКТОВ ПОИСКА КЛЮЧЕВЫХ СЛОВ В ТЕКСТОВОМ ПРЕДСТАВЛЕНИИ // Уральская горная школа регионам: сборник докладов международной научно-практической конференции – Екатеринбург: Изд-во УГГУ, 2024.

2. Осинцев А.В., Волкова Е.А. ИССЛЕДОВАНИЕ ВОПРОСА МОДЕЛЕЙ РАСПОЗНАВАНИЯ РЕЧИ // Уральская горная школа регионам: сборник докладов международной научно-практической конференции – Екатеринбург: Изд-во УГГУ, 2024.

## 1 АНАЛИЗ ПРОБЛЕМАТИКИ

					Д 09.04.01 2022 22286159 ВКРМ	Лист
		Осинцев А.В.		06.06.24		
Студ.	Фамилия	Подп.	Дата			10

## **1 АНАЛИЗ ПРОБЛЕМАТИКИ**

### **1.1. Характеристика проблемной ситуации/области**

В современном бизнесе колл-центры стали неотъемлемой частью работы практически любой крупной компании. Это связано с тем, что компании стремятся оставаться на связи со своими клиентами и быть в курсе их потребностей и проблем.[1]

Центры обработки звонков играют ключевую роль в обслуживании клиентов компании. Они предоставляют возможность клиентам обращаться за помощью, получать информацию о продуктах или услугах, а также разрешать возникшие проблемы и задавать вопросы. Отсутствие контакт-центра может негативно сказаться на удовлетворенности и лояльности клиентов, так как это лишает компанию важного канала коммуникации с ними [2,3]. Кроме того, колл-центры принимают и обрабатывают заявки, жалобы и отзывы от клиентов. Они выполняют роль посредника между клиентами и различными отделами компании, обеспечивая своевременное решение возникающих проблем. Но роль центра обработки звонков не ограничивается только обслуживанием клиентов. Они также могут использоваться для проведения маркетинговых исследований, сбора обратной связи от клиентов и реализации маркетинговых кампаний. Предоставление информации о новых продуктах, акциях и специальных предложениях также является важной функцией. Кроме того, точки приема и обработки звонков собирают ценные данные о клиентах, их запросах, предпочтениях и поведении. Эти данные могут быть использованы для анализа эффективности работы колл-центра, улучшения обслуживания клиентов и разработки более эффективных стратегий продаж. Таким образом, контакт-центры играют важную роль в формировании репутации компании. Качественное обслуживание клиентов, быстрое реагирование на их запросы и эффективное решение проблем способствуют созданию положительного имиджа компании.

Исходя из всего вышесказанного, колл-центры не только обеспечивают важный канал коммуникации с клиентами, но и являются ключевым элементом стратегии взаимодействия компании с рынком и обеспечения ее успешной деятельности.

Отдел колл-центра играет ключевую роль во взаимодействии с клиентами. Огромный объем обработанных звонков становится неотъемлемой частью их работы, подчеркивая важность и значимость. Каждый звонок от клиента представляет собой возможность для сбора ценной информации о его потребностях, запросах и жалобах [4]. Анализ этой информации дает компании понимание того, как улучшить продукты или услуги, а также оптимизировать свои процессы.

Большой объем звонков также позволяет выявлять тренды и паттерны в поведении клиентов. Это помогает компании адаптировать свои стратегии обслуживания и разработки продуктов под изменяющиеся потребности рынка. Кроме того, количество обработанных

звонков является показателем эффективности работы колл-центра, включая скорость ответа на звонки, уровень удовлетворенности клиентов и успешное разрешение их проблем.

Таким образом, большой объем обработанных звонков является не только индикатором активности бизнеса, но и ключевым элементом успешной стратегии обслуживания клиентов и развития компании в целом.[5]

Хотя отделы колл-центров существуют в компаниях уже давно, проблема большого объема информации становится все более острой для руководителей этих отделов и других высокопоставленных сотрудников компании [6]. Одной из наиболее сложных задач является оценка качества работы конкретного сотрудника, его компетентности и вежливости в общении с клиентами. Этот аспект требует внимательного анализа, так как качество обслуживания клиентов напрямую влияет на репутацию компании и ее отношения с клиентами. Руководители отделов колл-центра и вышестоящие по должности сотрудники сталкиваются с необходимостью оценивать множество параметров в работе своих сотрудников, таких как скорость ответа на звонки, профессионализм в решении проблем, эмпатия и вежливость в общении.

Кроме того, развитие технологий, таких как системы аналитики данных и искусственный интеллект, предоставляет новые возможности для автоматизации процессов оценки качества обслуживания клиентов и анализа работы сотрудников [7]. Это позволяет более точно выявлять сильные и слабые стороны каждого сотрудника, а также предлагать индивидуальные программы обучения и развития для повышения их профессиональных навыков и компетентности [8].

Исходя из предоставленной выше информации, видно, что центры обработки звонков являются неотъемлемой частью бизнеса в современном мире и требуют эффективных систем управления и анализа данных, включая системы распознавания и разметки голосовых записей с применением нейронных сетей.

Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей – это область, в которой используются методы машинного обучения для автоматического анализа и обработки аудиоданных, а также создаются алгоритмы и модели для автоматического анализа и классификации речевых данных [9]. В данной сфере требуется создание автоматизированной системы, способной обрабатывать миллионы голосовых записей без значительного вмешательства человека, а после выводить данные о разговоре и качестве работы сотрудника в воспринимаемом для руководителей виде.

Основные термины, важные для понимания данной темы, включают в себя:

Распознавание речи и разметка голосовых записей: Процесс анализа и классификации

аудиоданных, который может включать в себя определение ключевых слов и фраз, эмоционального окраса и других аспектов речи. [10,11]

Нейронные сети для обработки речи: Использование различных архитектур нейронных сетей, таких как рекуррентные нейронные сети (RNN) или трансформеры, для извлечения признаков из аудиоданных и обучения моделей распознавания речи.

Применение техник машинного обучения: Обучение моделей на больших объемах данных с использованием методов машинного обучения, таких как обратное распространение ошибки и градиентный спуск.

В целом, целью разработки данной системы является создание универсального инструмента для распознавания и разметки голосовых записей с использованием передовых методов машинного обучения, что позволит пользователям эффективно анализировать и обрабатывать аудиоданные для различных целей, таких как автоматическая транскрипция, анализ тональности речи, распознавание команд и многое другое.

В рамках исследования и разработки системы распознавания и разметки голосовых записей с применением нейронных сетей возникает важная потребность в создании специализированного сервиса, который будет предоставлять инструменты для эффективного анализа аудиоданных. Основным фокусом данного сервиса будет являться предоставление информации в доступной и удобной форме, а также обеспечение возможности анализа полной картины совершенного диалога.

Прежде всего, сервис должен быть способен обрабатывать голосовые записи с высокой точностью, используя для этого передовые алгоритмы машинного обучения, включая нейронные сети. Это позволит распознавать речь и выделять ключевые фрагменты диалога для последующего анализа.

Одним из ключевых преимуществ предлагаемого сервиса будет его способность представлять информацию в понятной и удобной форме для пользователя. Это может включать в себя текстовые транскрипты голосовых записей, а также визуализацию ключевых моментов диалога. Помимо этого, сервис будет выводить тип собеседника, а также оценку сотрудника за проделанную работу, а именно аудио контакт с клиентом. Такой подход значительно упростит процесс анализа для начальников отдела колл-центра, позволяя им быстро и точно оценить качество работы и выявить области для улучшения.

Важным аспектом создания сервиса будет также возможность анализа полной картины диалога. Это включает в себя не только распознавание речи, но и анализ тональности, эмоциональной окраски и других аспектов коммуникации. Такой подход позволит компаниям получить более глубокое понимание взаимодействия с клиентами и выявить потенциальные проблемы или возможности для улучшения обслуживания.

Итак, создание специализированного сервиса по распознаванию и разметке

голосовых записей с использованием нейронных сетей представляет собой значительный шаг в направлении автоматизации процессов анализа аудиоданных. Предоставление информации в удобной форме и анализ полной картины диалога поможет компаниям улучшить качество обслуживания клиентов и повысить эффективность работы колл-центров.

## **1.2. Аналитический обзор существующих решений**

В современном мире, в контексте исследования и разработки системы распознавания и разметки голосовых записей с применением нейронных сетей, существует стремление к созданию инновационных решений, позволяющих компаниям автоматизировать работу с клиентами.

Способы применения разрабатываемой системы:

Речевые помощники и умные ассистенты: Такие системы используют распознавание и разметку голосовых записей для понимания запросов пользователей и предоставления им соответствующих ответов или услуг.

Мониторинг и анализ аудиоконтента: это может включать в себя мониторинг радио- и телепередач, анализ телефонных разговоров в контактных центрах для оценки качества обслуживания клиентов и др.

Автоматическая транскрибация аудио и видео: Нейронные сети позволяют автоматически преобразовывать речь из аудио и видеофайлов в текстовый формат, что полезно для создания субтитров, анализа контента и т.д.

Безопасность и наблюдение: Распознавание и разметка голосовых записей могут использоваться для обнаружения определенных звуков (например, выстрелов) или голосовых паттернов (например, признаки стресса в голосе), что может быть полезно для систем видеонаблюдения и безопасности.

Проблемы, которые необходимо преодолеть при создании системы:

Качество распознавания речи: несмотря на значительные успехи, нейронные сети все еще могут допускать ошибки при распознавании речи, особенно в условиях с шумом или с различными акцентами.

Необходимость больших объемов данных: для обучения нейронных сетей требуются большие объемы размеченных аудиоданных, что может быть трудоемким и дорогостоящим процессом.

Конфиденциальность и безопасность данных: поскольку аудиоданные могут содержать конфиденциальную информацию (например, персональные разговоры), важно обеспечить их защиту и соблюдение приватности.

Исследование и разработка системы распознавания и разметки голосовых записей с

применением нейронных сетей представляет собой важное направление в области обработки аудиоданных, обладающее широким спектром потенциальных применений и вызовов.

В связи с этим очень немногие компании занимаются решением данной проблемы, в связи с объемностью работы и множеством подводных камней.

Одна из компаний, которая занимается решением задачи распознавания речи, является Twin Speech Analytics [12]

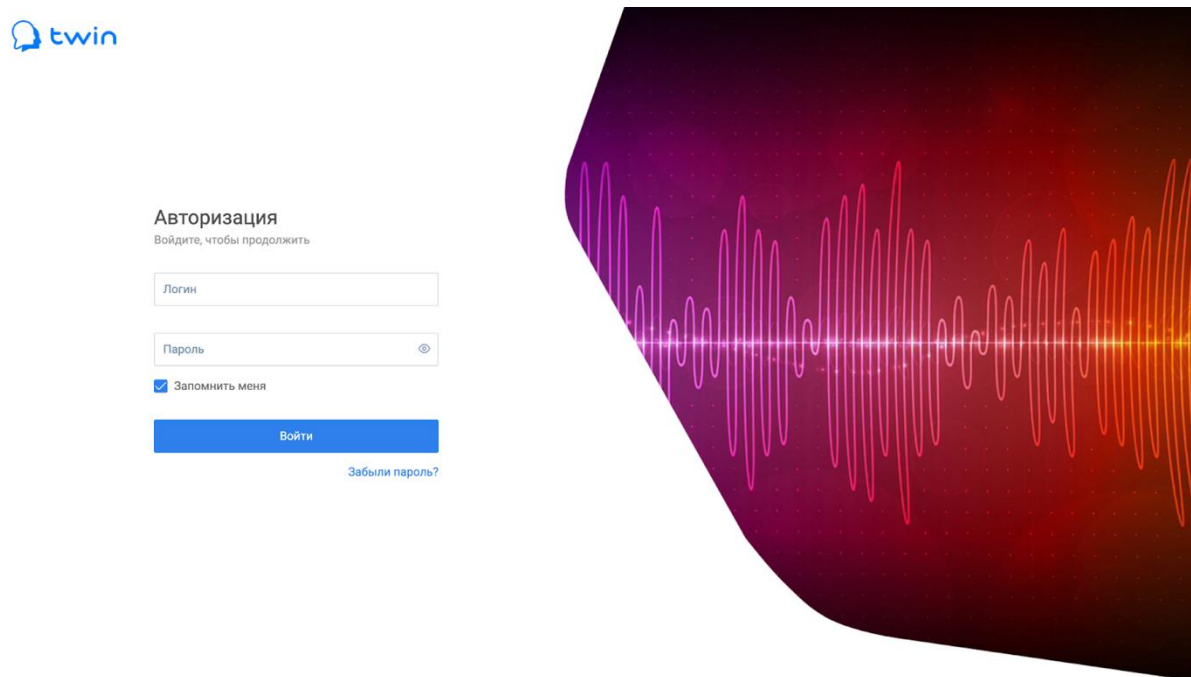


Рисунок 1 - сервис аналитики речи Twin Speech Analytics

Twin Speech Analytics — это сервис аналитики речи, который обычно используется для анализа разговоров, проводимых через телефонные линии, или другие аудиоканалы. Этот сервис может быть полезен для компаний, которые хотят оценить качество обслуживания клиентов, понять потребности своей аудитории, а также выявить тенденции и проблемы в коммуникации. Преимущества и недостатки описаны в таблице 1.

Таблица 1 – преимущества и недостатки сервиса Twin Speech Analytics

Преимущества	Недостатки
<b>Транскрибирование речи:</b> Преобразование аудиозаписей разговоров в текстовый формат для дальнейшего анализа.	<b>Точность распознавания речи:</b> несмотря на значительные усовершенствования в технологиях распознавания речи, ошибки все еще могут возникать, особенно при обработке сложных диалектов, акцентов или разговоров с плохим качеством аудио.
<b>Анализ тональности и настроения:</b> Определение тональности и эмоционального окраса речи участников разговора. Например, определение уровня	<b>Неоднозначность тонов и эмоций:</b> Анализ тональности и настроения в разговорах может быть непростым из-за неоднозначности интонаций и эмоций,

удовлетворенности клиента или эмоционального состояния оператора.	особенно в случае использования сарказма, иронии или других нюансов в речи.
<b>Идентификация ключевых слов и фраз:</b> Поиск определенных ключевых слов или фраз, которые могут свидетельствовать о конкретных потребностях клиента или проблемах в обслуживании.	<b>Необходимость правильной интерпретации данных:</b> даже если алгоритмы сервиса успешно распознают ключевые слова или фразы, интерпретация этих данных может требовать экспертного анализа, чтобы понять их реальное значение и контекст.
<b>Анализ эффективности общения:</b> Оценка эффективности коммуникации оператора с клиентом, включая скорость реакции, ясность выражения и т. д.	<b>Конфиденциальность данных:</b> поскольку Twin Speech Analytics работает с аудиозаписями разговоров, необходимо обеспечить высокий уровень конфиденциальности и защиты данных клиентов, чтобы избежать возможных нарушений приватности.
<b>Выявление тенденций и проблем:</b> Анализ данных на предмет выявления повторяющихся проблем или тенденций в разговорах, что позволяет компаниям принимать меры для улучшения качества обслуживания.	<b>Сложность внедрения и настройки:</b> для полноценного использования сервиса может потребоваться интеграция с существующими системами компании, а также настройка алгоритмов анализа в соответствии с конкретными потребностями и целями бизнеса.
<b>Мониторинг соответствия стандартам:</b> Оценка того, насколько разговоры соответствуют установленным стандартам обслуживания или предоставления информации.	<b>Затраты на обучение персонала:</b> для эффективного использования Twin Speech Analytics сотрудники компании могут нуждаться в обучении, чтобы понимать, как правильно интерпретировать данные и принимать на их основе обоснованные бизнес-решения.

Исходя из вышеизложенного и факта отсутствия пробного периода, а также обработанная информация выводится в виде отчета, но при этом процесс транскрибации скрыт от пользователя, то есть при возникновении неполадок в использовании, пользователь не сможет самостоятельно разобраться в первоисточнике проблемы

В заключении важно отметить, что хотя Twin Speech Analytics обладает определенными преимуществами в анализе речи и предоставлении ценных инсайтов для бизнеса, у него также есть значительные недостатки. Учитывая изложенную выше информацию, компании должны тщательно подумать о том, стоит ли взаимодействовать с данной компанией.





Рисунок 2 - сервис аналитики речи Salute Speech

SaluteSpeech, предоставляемый Сбером, представляет собой передовую систему синтеза и распознавания речи. Он может преобразовывать текст в естественную речь с правильными ударениями и распознавать речь, игнорируя фоновые шумы и анализируя эмоции говорящего (позитив, нейтрал, негатив). Сервис используется для автоматизации бизнес-процессов, таких как интерактивные голосовые меню (IVR), голосовые рассылки и анализ качества обслуживания. Поддержка SSML позволяет детально настраивать синтез речи для более естественного звучания. Таблица 2 отображает преимущества и недостатки данного сервиса.

Таблица 2 – преимущества и недостатки сервиса Salute Speech

Преимущества	Недостатки
<b>Синтез речи и правильные ударения:</b> Технология генерирует естественную речь, корректно расставляя ударения и понимая контекст (например, "звонишь").	<b>Точность распознавания:</b> как и у большинства систем, могут возникать ошибки при сложных диалектах и акцентах.
<b>Распознавание речи и снижение шума:</b> Сервис игнорирует фоновые шумы, определяет завершение высказываний и распознает три эмоции (позитив, нейтрал, негатив).	<b>Интерпретация данных:</b> несмотря на распознавание ключевых слов, требуется экспертный анализ для точного понимания контекста.

<b>Автоматизация бизнес-процессов:</b> IVR, голосовые рассылки, анализ качества обслуживания и озвучка контента.	<b>Конфиденциальность данных:</b> Обработка аудиозаписей требует высоких стандартов защиты данных.
<b>Гибкость и поддержка SSML:</b> позволяет детально настраивать синтез речи для более естественного звучания.	<b>Стоимость и настройка:</b> Полноценное использование требует интеграции с существующими системами компании и может быть затратным.
	<b>Обучение персонала:</b> Сотрудникам может потребоваться обучение для эффективного использования данных и принятия обоснованных решений.

SaluteSpeech предлагает значительные преимущества для бизнеса в области синтеза и распознавания речи, однако требует тщательного рассмотрения и подготовки для внедрения. В свою очередь, количества и значимость недостатков перевешивает преимущества, что ставит целесообразность использования данного сервиса под сомнение.

В то же время на мировом рынке представлены и другие сервисы, которые предлагают аналитику речи и обработку аудиоданных для бизнеса. Вот некоторые из них:

Google Cloud Speech-to-Text и Natural Language Processing (NLP) [13]: Google Cloud предоставляет инструменты для распознавания речи и анализа текста, которые могут использоваться для создания собственных систем аналитики речи. Эти инструменты позволяют транскрибировать разговоры в текст, а также проводить анализ тональности, выделение ключевых фраз и многие другие задачи.

IBM Watson Speech to Text и Watson Tone Analyzer: IBM Watson [14] предоставляет собственные инструменты для распознавания речи и анализа тональности. С их помощью можно транскрибировать разговоры, определять тональность речи и эмоциональный окрас, а также анализировать текст на предмет ключевых слов и тем.

Amazon Transcribe [15] и Amazon Comprehend [16]: Amazon Web Services (AWS) предлагает сервисы для транскрибации аудио и анализа текста. Amazon Transcribe обеспечивает точное распознавание речи, а Amazon Comprehend позволяет проводить анализ тональности, выявлять ключевые фразы и темы, а также определять эмоциональный окрас текста.

Microsoft Azure Speech to Text и Text Analytics [17]: Microsoft Azure предоставляет инструменты для распознавания речи и анализа текста, которые могут быть использованы для создания систем аналитики речи. Azure Speech to Text позволяет транскрибировать аудио, а Text Analytics предоставляет функционал для анализа тональности, выделения

ключевых фраз и многих других задач.

Эти сервисы предоставляют различные функциональные возможности и уровни интеграции с другими системами. Но в связи с текущими обстоятельствами, предприятия стремятся минимизировать связь с зарубежными компаниями из-за национальных вопросов и сложностей в ведении переговоров и оплаты услуг.

Таким образом, исходя из вышеизложенной информации, что компании не хотят зависеть от иностранных технологий, на отечественном рынке представлена всего одна компания, предоставляющая сервис для решения данного бизнес-процесса. В этом контексте создание такого проекта становится актуальным и важным для предприятий, предоставляя им возможность решать задачи без зависимости от зарубежных контрагентов. Это также способствует импортозамещению и облегчает компаниям решение проблем без дополнительных сложностей.

### **1.3 Основные положения выносимые на защиту**

1. Внедрение системы помогает своевременно выявлять и устранять недостатки в коммуникации, улучшая общую эффективность взаимодействия с клиентами и повышая удовлетворенность клиентов.

2. Целесообразно использование автоматизированных систем для обеспечения непрерывного процесса разметки и распознавания голосовых данных с целью повышения эффективности анализа записей разговоров.

3. Постоянное обновление и оптимизация моделей распознавания речи на основе анализа данных и актуальных тенденций помогает уменьшить ошибки и улучшить точность обработки.

### **Вывод по главе 1**

Цель данной работы состоит в создании системы, способной эффективно распознавать, классифицировать и анализировать голосовые записи с минимальным участием человека, а также дальнейшим выводом информации для анализа.

Проведенный анализ конкурентов выявил значительные недостатки в существующих системах для распознавания и разметки голосовых записей с применением нейронных сетей. Существующие платформы не соответствуют всем требованиям качества и не решают все обнаруженные проблемы. В свете этого предлагается разработать новую платформу, охватывающую широкий спектр функций обучения и тестирования, чтобы обеспечить более эффективное решение задач с учетом всех обнаруженных проблем. Twin Speech Analytics — это сервис для анализа речевых взаимодействий через телефонные


линии и другие аудиоканалы. Он помогает компаниям оценивать качество обслуживания, выявлять потребности клиентов и определять тенденции в коммуникации. Из основных недостатков можно выделить: возможные ошибки распознавания речи при сложных акцентах и плохом качестве аудио, неоднозначность интерпретации тонов и эмоций и сложность внедрения и настройки. Сервис имеет значительные преимущества, но требует внимательного подхода к интеграции и защите данных. Salute Speech от Сбера — это высокотехнологичный сервис для синтеза и распознавания речи, обеспечивающий преобразование текста в естественную речь и точное распознавание аудиозаписей с фильтрацией фоновых шумов. Он поддерживает автоматизацию бизнес-процессов, включая IVR, голосовые рассылки и анализ качества обслуживания. Из основных недостатков можно выделить: возможные ошибки распознавания сложных диалектов, требования к защите данных и затратам на интеграцию и обучение. Сервис предлагает значительные возможности для бизнеса, но требует тщательного подхода к внедрению и эксплуатации.

Предлагаемый сервис будет включать в себя инновационные методы обучения и тестирования, а также обеспечивать высокую точность и надежность распознавания и разметки голосовых данных.

Кроме того, в ходе разработки платформы будут проведены эксперименты, направленные на подтверждение основных положений, выносимых на защиту. Это позволит дополнительно обосновать эффективность и потенциал предлагаемой системы в решении задач распознавания и разметки голосовых записей с применением нейронных сетей.

Таким образом, разработка предлагаемой платформы представляет собой важный шаг в направлении улучшения систем распознавания и разметки голосовых данных, что позволит компаниям более эффективно анализировать и использовать аудиоинформацию в своей деятельности.

## 2 ПРЕДЛАГАЕМОЕ РЕШЕНИЕ

					Д 09.04.01 2022 22286159 ВКРМ	Лист
		Осинцев А.В.		06.06.24		
Студ.	Фамилия	Подп.	Дата			21

## 2 ПРЕДЛАГАЕМОЕ РЕШЕНИЕ

### 2.1 Системный анализ проблемной ситуации/области

#### 2.1.1 Математическая модель объекта управления

Математическая модель, представленная ниже, описывает процесс оценки звонков операторов с использованием платформы, основанной на искусственном интеллекте. Эта модель является ключевым инструментом для анализа и улучшения качества обслуживания клиентов в различных сферах бизнеса. Математическая модель включает в себя несколько компонентов, начиная от преобразования речи в текст до анализа содержания звонков и вывода оценки на основе чеклистов[22].

#### Транскрибация звонка:

Преобразование звонка в текст можно представить в данном виде (1):

$$f: X \rightarrow T \quad (1)$$

где  $f$  - функция, принимающая  $X$  аудио запись звонка и выводящая  $T$  его текстовое представление.

#### Анализ текста по чеклистам:

После получения текста  $T$  (1) его необходимо проанализировать по чеклистам, чтобы выявить наличие или отсутствие ключевых слов в разговоре оператора. Чеклисты представлены набором ключевых слов или фраз, которые свидетельствуют о выполнении или невыполнении определенных критериев. Пусть  $C = \{c_1, c_2, \dots, c_n\}$  - набор чеклистов, тогда  $A(T)$  проверяет выполнение условий каждого чеклиста (2):

$$A(T) = \sum_{i=1}^n \delta(c_i(T)) \quad (2)$$

Где  $A(T)$  функция для анализа текста по заданным чеклистам.

$\delta(c_i(T))$  индикатор выполнения условия  $c_i$  в тексте  $T$

#### Вывод итоговой оценки:

После анализа текста по чеклистам мы можем вычислить итоговую оценку  $E$  на основе результатов анализа. Это может быть сделано с использованием функции  $f$  (3):

$$E = f(A(T)) \quad (3)$$

где  $f$  - функция, преобразующая результат анализа в итоговую оценку, например, нормализованную от 0 до 100.

### **Общее описание процесса**

**Преобразование речи в текст:** звуковой сигнал звонка преобразуется в текстовый формат с использованием алгоритмов распознавания речи. (1)

**Анализ текста:** текст звонка проходит через алгоритмы анализа, которые ищут ключевые слова или фразы, связанные с качеством обслуживания. (2)

**Вывод оценки:** на основе результатов анализа (3) формируется итоговая оценка, отражающая качество обслуживания оператором.

Таким образом, математическая модель платформы для оценки звонков операторов представляет собой комплексный процесс, начиная от преобразования речи в текст, через анализ текста по чеклистам, и заканчивая выводом итоговой оценки, что обеспечивает систематизированный и формализованный подход к оценке качества обслуживания клиентов.

### **2.1.2 Концептуальное моделирование**

Объектом автоматизации является система распознавания и разметки голосовых записей с применением нейронных сетей. В рамках диссертации создана концептуальная ER-диаграмма Питера Чена, содержащая в себе сущности: Пользователь, Телефон, Слово, Очередь назначения тега, Отвечающий, Цель, Сегмент, Тип отвечающего, Тег, Трек и слабые сущности: Звонок, Чек-лист, Отчет, Секция. Сущности по отдельности представлены на рисунках 3-16 и отношения между ними на рисунке 17.

На рисунке 3 представлена сильная сущность – Пользователь. Этот элемент системы играет фундаментальную роль, поскольку включает в себя основные данные о пользователе. Атрибуты этой сущности позволяют хранить различные аспекты профиля пользователя и обеспечивать его идентификацию в системе. Включает в себя:

id (primary key), Статус, а также составной атрибут – Данные пользователя, который в свою очередь включает в себя атрибуты Имя и Фамилия.

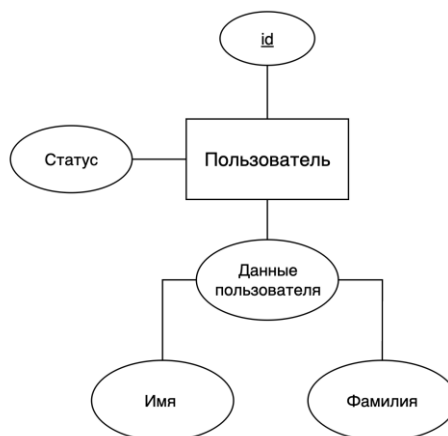


Рисунок 3 - сущность “Пользователь”

Сущность "Звонок" представлена на рисунке 4. Звонок является важным компонентом системы, поскольку содержит информацию о каждом звонке и его характеристиках. Включает в себя: id (Primary key), Длительность и Распознаемость звонка.

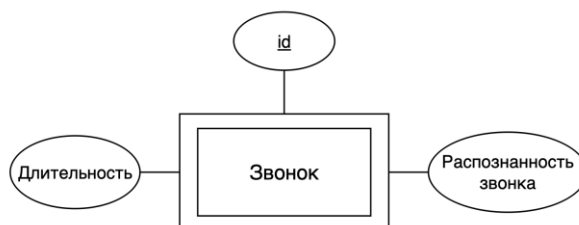


Рисунок 4 - сущность “Звонок”

Сущность "Чек-лист" – представлена на рисунке 5. Эта сущность представляет собой перечень критериев, используемых для оценки и проверки качества совершаемых пользователями звонков. Она играет важную роль в процессе обработки данных, включая в себя выбор и настройку алгоритмов, а также проверку качества разметки записей. Включает в себя id (Primary Key), Название, Описание, признак Архивированности и Частоту.

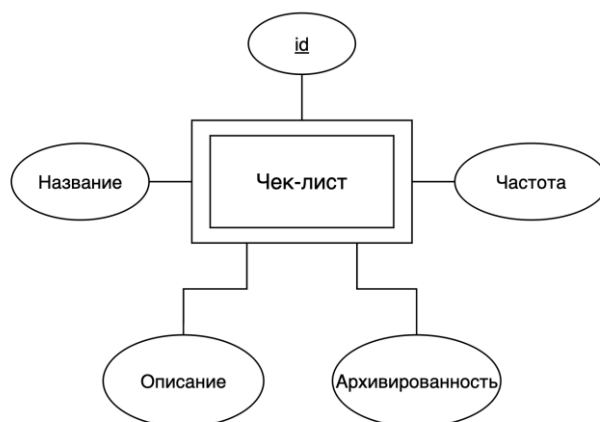


Рисунок 5 - сущность “Чек-лист”

Рисунок 6 – сущность “Телефон”. Данная сущность хранит в себе внутренние номера телефонов. включает в себя телефон (Primary key) и Дату создания.



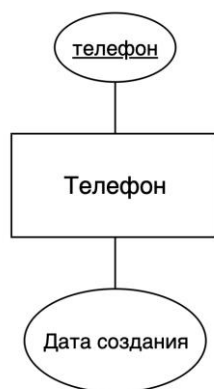


Рисунок 6 - сущность “Телефон”

Сущность "Слово" – представлена на рисунке 7. В контексте данной темы, сущность "слово" относится к основной единице информации в голосовых записях, которую необходимо распознать и разметить. Она играет важную роль распознавании и разметки слов, что в свою очередь, является одной из ключевых задач в системах обработки речи. Включает в себя id (Primary Key), Начало, Конец и Текст.

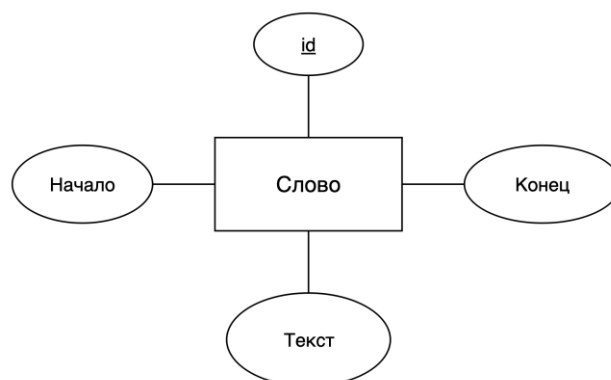


Рисунок 7 - сущность “Слово”

На рисунке 8 представлена сущность “Очередь назначения тегов”. Эта сущность относится к процессу или структуре, используемой для назначения и присвоения тегов или меток голосовым записям после их распознавания. Эта очередь представляет собой последовательность действий или задач, которые необходимо выполнить для определения и присвоения соответствующих тегов каждому элементу или фрагменту записи. В контексте систем распознавания и разметки голосовых данных с применением нейронных сетей, эта очередь может включать в себя такие этапы, как классификация слов, определение контекста и применение соответствующих тегов в зависимости от задачи или цели обработки. Включает в себя атрибуты id (Primary Key) Массив тегов и Дату создания.

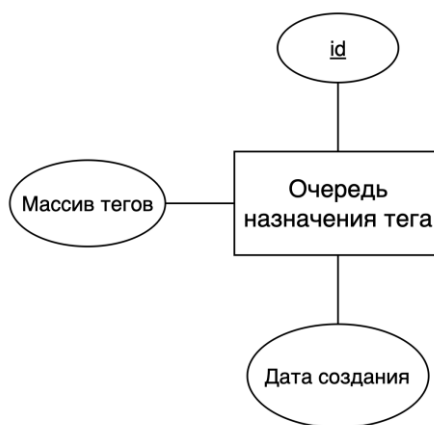


Рисунок 8 - сущность “Очередь назначения тега”

Сущность "Отчет" (см. Рисунок 9) является важной частью системы распознавания и разметки голосовых записей, поскольку представляет собой документ или набор данных, содержащих информацию о результатах работы системы. Этот отчет может включать в себя различные метрики и показатели, оценивающие эффективность и качество работы. Он играет важную роль в оценке производительности системы, так как представляет аналитические выводы. Id (Primary key), Период от, Период до, Результат и Количество звонков – атрибуты данной сущности.

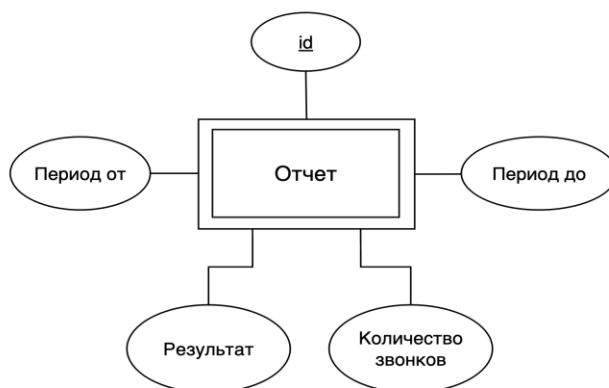


Рисунок 9 - сущность “Отчет”

Рисунок 10 – сущность “Отвечающий” описывает данные о контрагенте. Содержит такие атрибуты, как Id (Primary key), а также составной атрибут – Данные отвечающего, который в свою очередь включает в себя атрибуты Полное имя и Краткой имя.

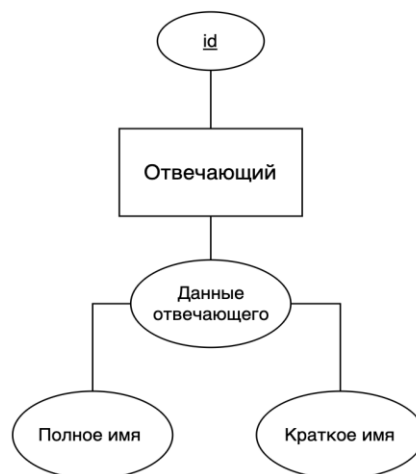


Рисунок 10 - сущность “Отвечающий”

Сущность "Цель", изображенная на рисунке 11, описывает цель звонка. Включает в себя атрибут – id (Primary key) и Имя.

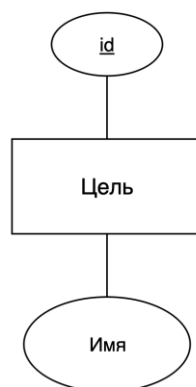


Рисунок 11 - сущность “Цель”

На рисунке 12 представлена сильная сущность – Сегмент. Эта сущность играет ключевую роль в системе, поскольку содержит информацию, которая классифицирует контрагента, например клиент может быть: потенциальным, лояльным, неизвестным или постоянным. Включает в себя: id (primary key) и Имя.

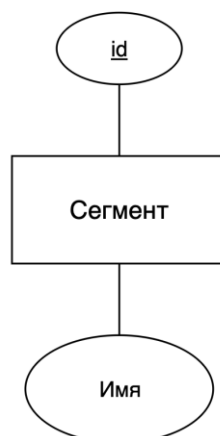


Рисунок 12 - сущность “Сегмент”

Сущность "Тип отвечающего", которая содержит информацию о клиенте, как о субъекте права, то есть контрагент может быть физ. лицом, юр. лицом или неизвестным, представлена на рисунке 13. Включает в себя: id (Primary key) и Имя.



Рисунок 13 - сущность "Тип отвечающего"

Рисунок 14 – сущность "Секция". Эта сущность позволяет разбить теги по категориями, которые, в свою очередь, называются секциями и существуют в рамках одного чек-листа, то есть являются уникальными. Id (Primary key), Имя и Название связаны с сущностью Секция.

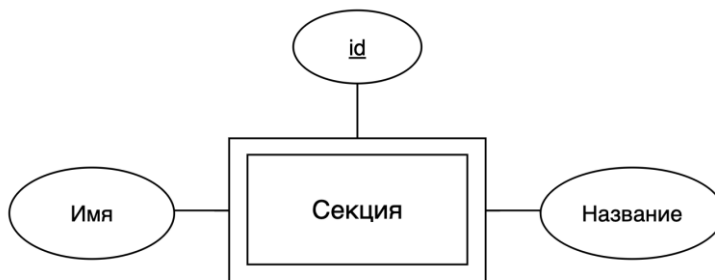


Рисунок 14 - сущность "Секция"

Сущность "Тег" – представлена на рисунке 15. Эта сущность является важным компонентом системы распознавания и разметки голосовых записей так как представляет собой метку или ярлык, который присваивается голосовым записям или их фрагментам после их распознавания и разметки. Эти теги обычно используются для классификации или идентификации содержания записей, облегчая последующий поиск, фильтрацию или анализ данных. В системах распознавания и разметки голосовых данных с применением нейронных сетей, теги играют важную роль в организации и обработке информации, что позволяет эффективно управлять и использовать голосовые данные в различных приложениях и задачах. Включает в себя id (Primary Key), Имя, Композиция и Цвет.

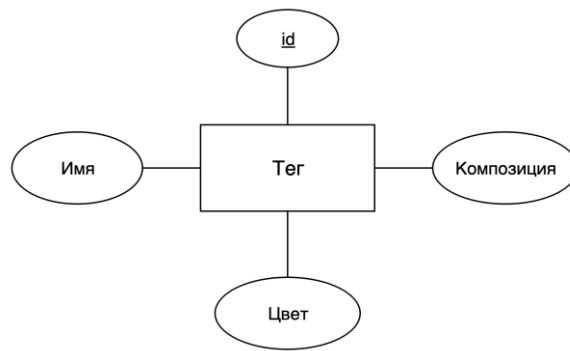


Рисунок 15 - сущность “Тег”

На рисунке 16 представлена сущность “Трек”. Эта сущность представляет собой конкретную дорожку из звонка. Помимо этого, внутри данной сущности находится информация о пути звонка, слова, а также транскрибированное(текстовое) представление звонка. В процессе работы системы распознавания и разметки голосовых записей с применением нейронных сетей, треки используются для организации и структурирования данных, обеспечивая точное и эффективное распознавание речи и других аудиофрагментов. Включает в себя атрибуты id (Primary Key), Сторона звонка, Длительность и Дата создания.

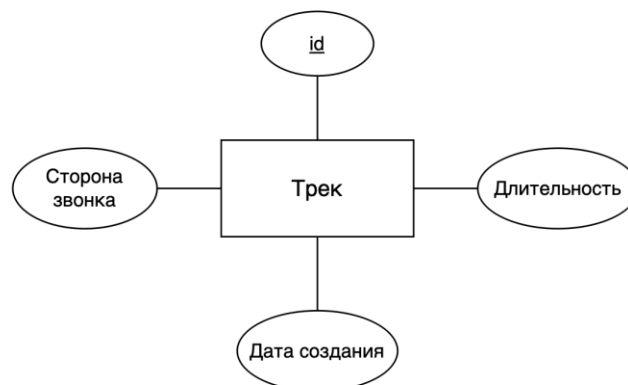


Рисунок 16 - сущность “Трек”

На рисунке 17 представлена концептуальная ER-диаграмма Питера Чена, содержащая в себе все сущности, их первичные ключи, а также данная диаграмма показывает отношения между описанными сущностями. Можно выделить такие компоненты системы, как Чек-лист, Звонок и Тег в связи с тем, что данные сущности имеют больше всего связей, то есть взаимодействуют со многими аспектами системы. Важно отметить, что многие элементы системы имеют связь “многие ко многим”, для упрощения отношений между ними, созданы искусственные сущности, которые приводят связь к виду “один ко многим”.

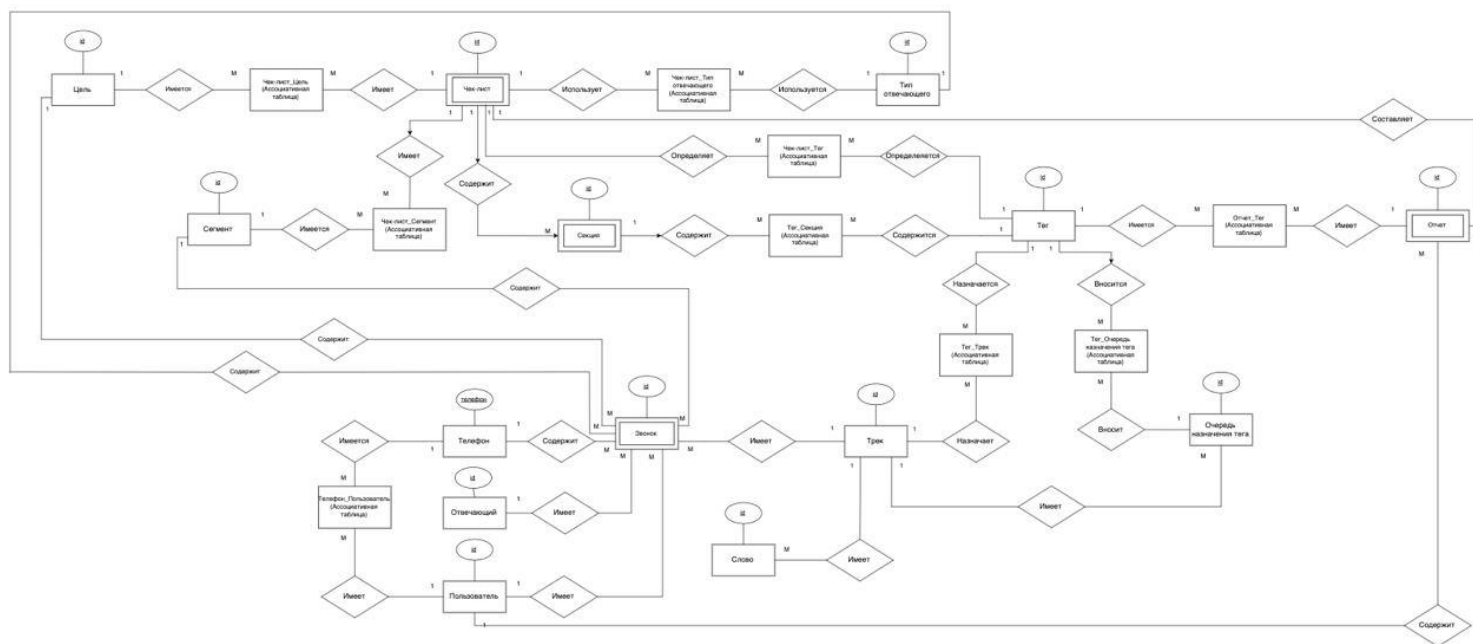


Рисунок 17 - Концептуальная ER диаграмма Питера Чена

## 2.2 Функциональная модель

В функциональной модели IDEF0 «как есть» (см. рисунок 18) вводными данными являются такие параметры как: хранилище звонков и базовая информация о звонке. Регламентирующие механизмы этой модели включают в себя такие нормативные документы как: регламент обработки текста. В качестве ресурсов, достижения необходимого результата можно выделить следующие параметры: speech-to-text и менеджеры. Выходным результатом совокупности данных параметров будет являться “оценка звонков”.

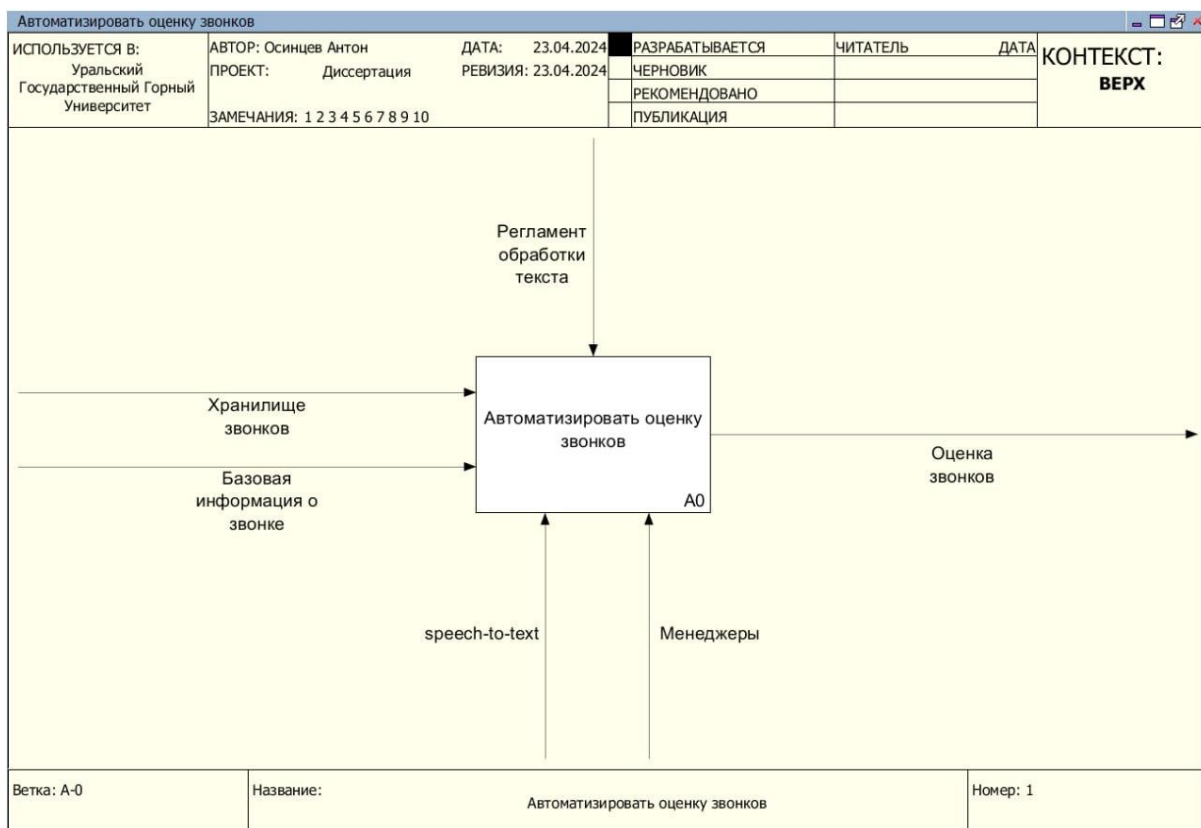


Рисунок 18 - IDEF0 модель уровня A0 “как есть”

Процесс автоматизации оценки звонков можно разделить на 3 основных, взаимосвязанных между собой, блока (рисунок 19):

- 1) Получить звонки из хранилища
- 2) Преобразовать записи звонков в текст
- 3) Составить отчет

Для большинства блоков можно выделить различные типы входных данных. Так для блока «Получить звонки из хранилища» входные данные – это данные из хранилища звонков. Для блока «Преобразовать записи звонков в текст» входными параметрами будет базовая информация о звонке. Регламентирующими механизмами, влияющими на каждый из блоков, являются нормативные документы, а именно регламент обработки текста. Также для достижения необходимого результата необходимы ресурсы. Ресурс “менеджеры” касается двух этапов, а именно “Получить звонки из хранилища” и “Составить отчет”. Ресурс “speech-to-text” необходим только для второго блока. Конечными выходными данными модели будет “запущенный бот”.

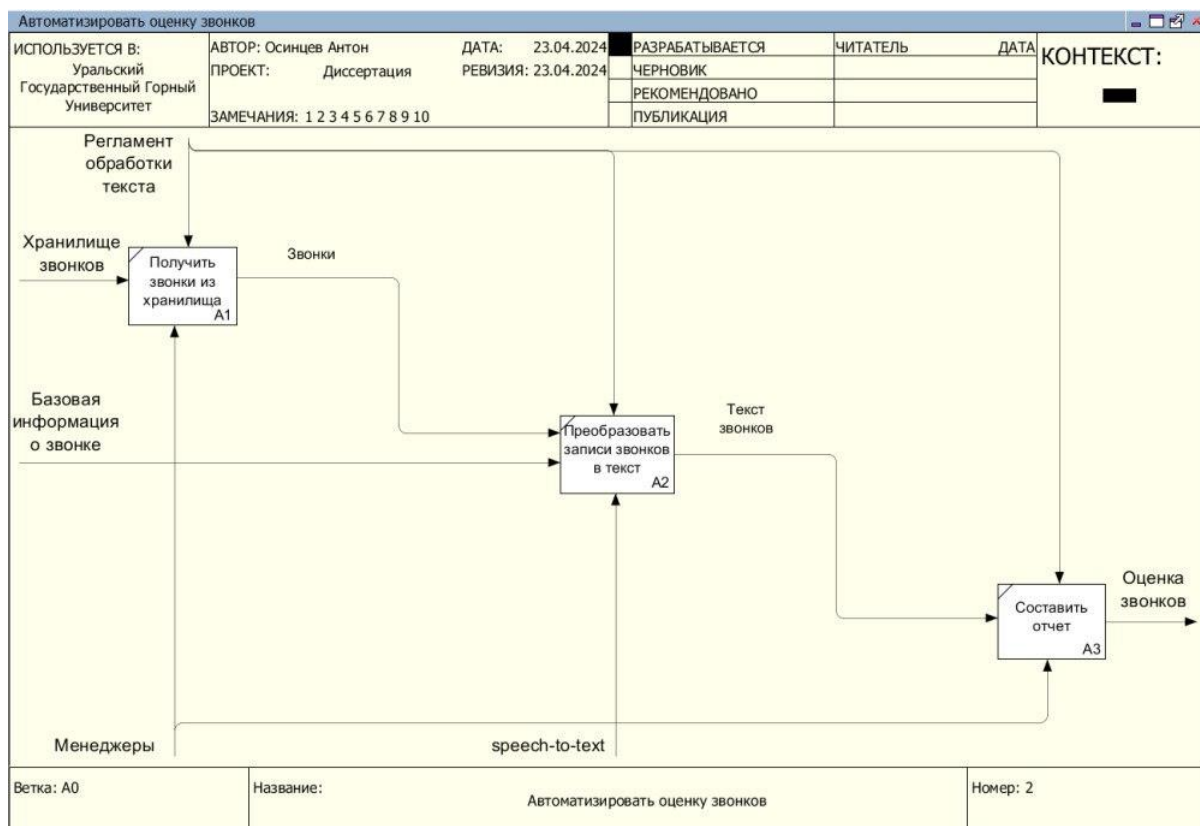


Рисунок 19 – IDEF0 модель уровня A1 “как есть”

Но для должного выполнения процесса “автоматизировать оценку звонков” и получения удовлетворительного результата входных данных, получаемых существующими системами недостаточно. Существующие решения не реализуют и не предоставляют часть факторов, необходимых для полноценного функционирования процесса. Так в функциональной модели IDEF0 «как будет» (рисунок 20) представлены

недостающие компоненты входных данных, такие как данные сотрудников, данные контрагентов и метаданные о звонке. Благодаря которым на выходе реализуются “Чек-лист с оценкой звонков”, “Текст звонков”, “Теги звонков” и “Отчеты”, благодаря которым процесс становится целостным и завершенным. Учитывая эти входные данные при создании системы, можно добиться нужного результата на выходе.



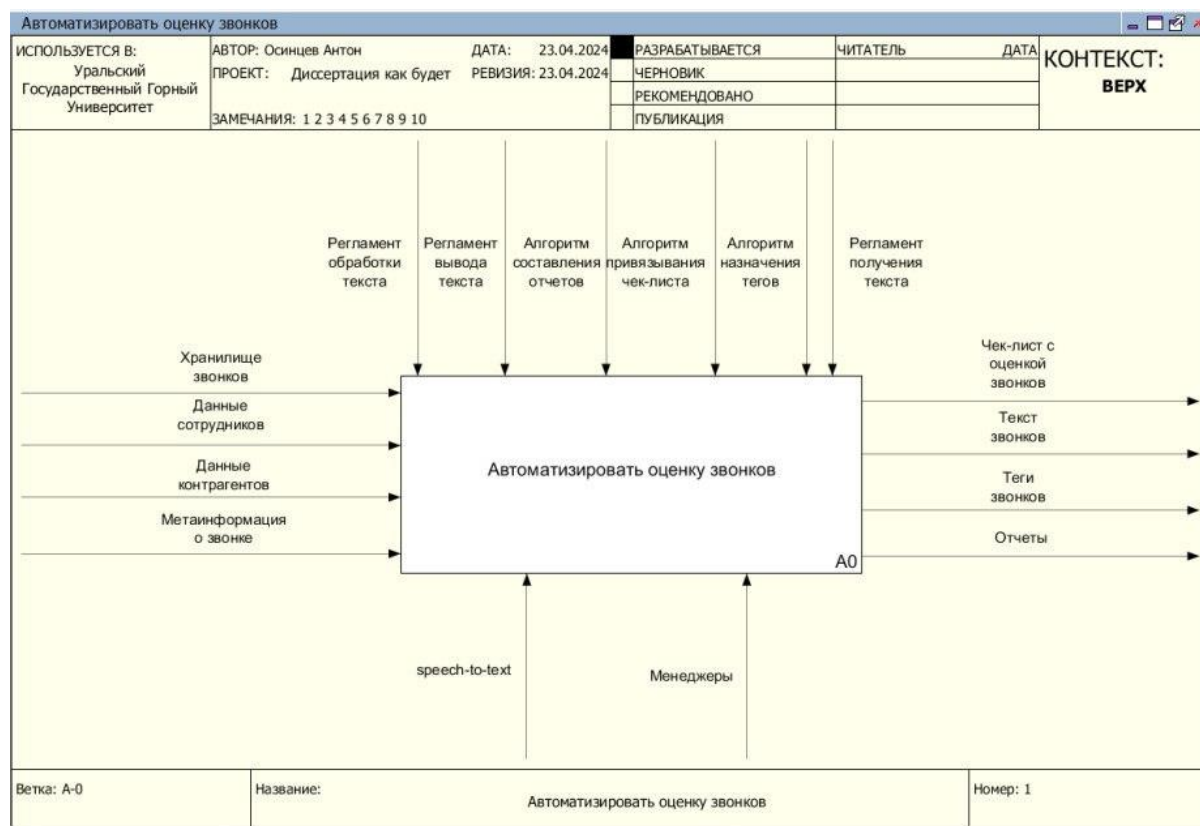


Рисунок 20 – IDEF0 модель уровня A0 “как будет”

Декомпозиция диаграммы «Как будет» (см. рисунок 21), в свою очередь, значительно отличается от той, которая показана на рисунке 19. На данной диаграмме видно, что появились пункты, раскрывающие более полно данный аспект, а именно: назначить теги по распознанному тексту и коррелировать чек-лист со звонками. Таким образом, данные вышеизложенные пункты позволяют сделать процесс более отлаженным, последовательным и детализированным, что в свою очередь дает возможность на выходе получить качественный продукт, а именно структурированную, готовую к дальнейшему анализу информацию. Выходными данными всего процесса являются “Чек-лист с оценкой звонков”, “Текст звонков”, “Теги звонков” и “Отчеты”. Также блок “Коррелировать чек-лист со звонками” обзавелся индивидуальными выходными данными, а конкретно “данные контрагентов”, “данные сотрудников” и “метаинформация о звонке”.

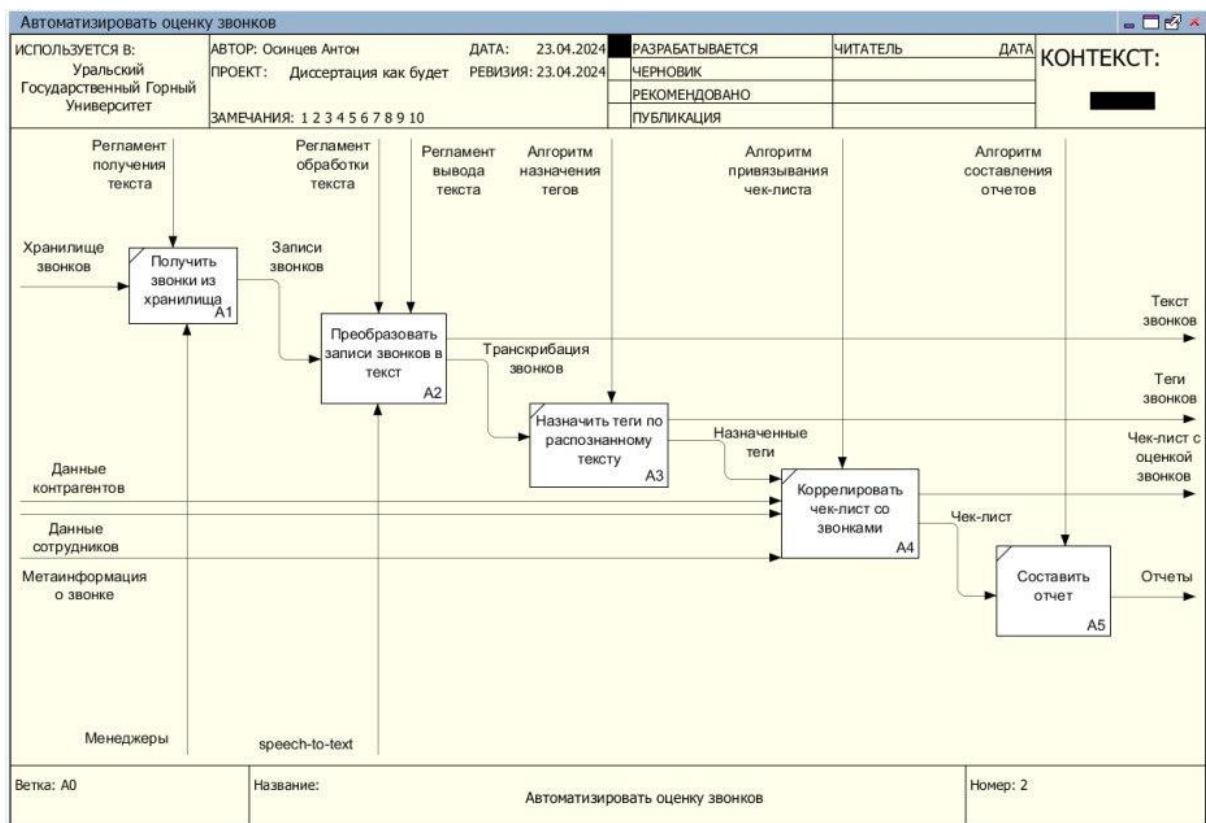


Рисунок 21 – IDEF0 модель уровня A1 “как будет”

Декомпозиция блока “Преобразовать записи звонков в текст” (см. Рисунок 22) включает в себя следующие процессы:

1. Определить активные участки записи звонка;
2. Объединить активные участки;
3. Транскрибировать аудио в текст;
4. Валидировать временные метки слов.

Входными данными служат: записи звонков. Нормативными документами для данных процессов служат “Регламент обработки текста” и “Регламент вывода текста”. Ресурсы управления, в свою очередь, наследовались из диаграммы IDEF0 модели уровня A1 “как будет”. Выходными данными всего процесса являются “Транскрибация звонков” и “Текст звонков”.

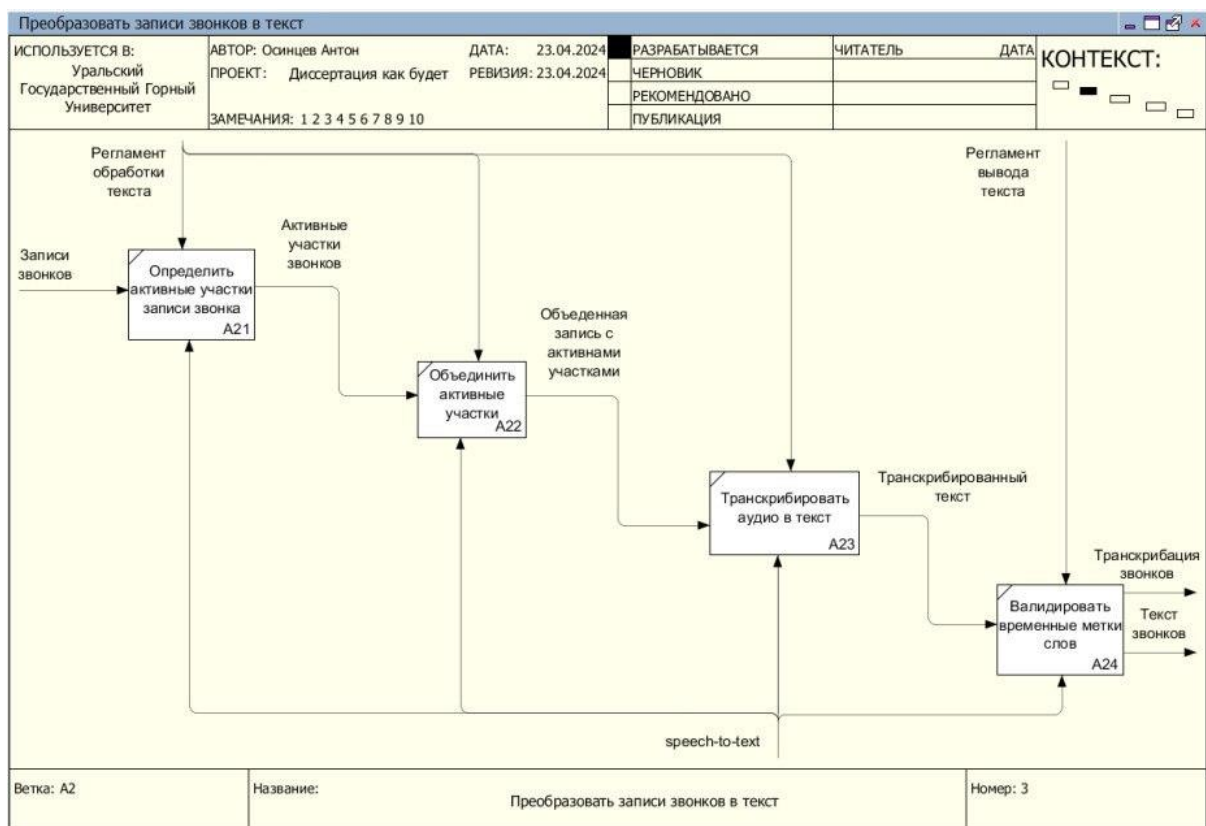


Рисунок 22 – IDEF0 модель уровня A2 “как будет”

Функциональное моделирование наглядно демонстрирует процессы, применяемые в настоящее время для решения поставленных задач и их потенциальную оптимизацию в будущем. Диаграммы "как есть" подробно отображают текущий процесс автоматизации оценки звонков, предоставляя информацию о текущем состоянии решаемой проблемы. А диаграммы "как будет" описывают финальный процесс после внедрения задуманного сервиса, позволяя представить его будущее воплощение.

## 2.3 Имитационная модель

Имитационная модель для исследования и разработки системы распознавания и разметки голосовых записей с применением нейронных сетей представляет собой комплексный инструмент, предназначенный для автоматического анализа звонков, представлен на рисунке 23. В данной модели акцент сделан на расшифровку разговоров, определение их содержания и оценку качества взаимодействия менеджера с клиентом.

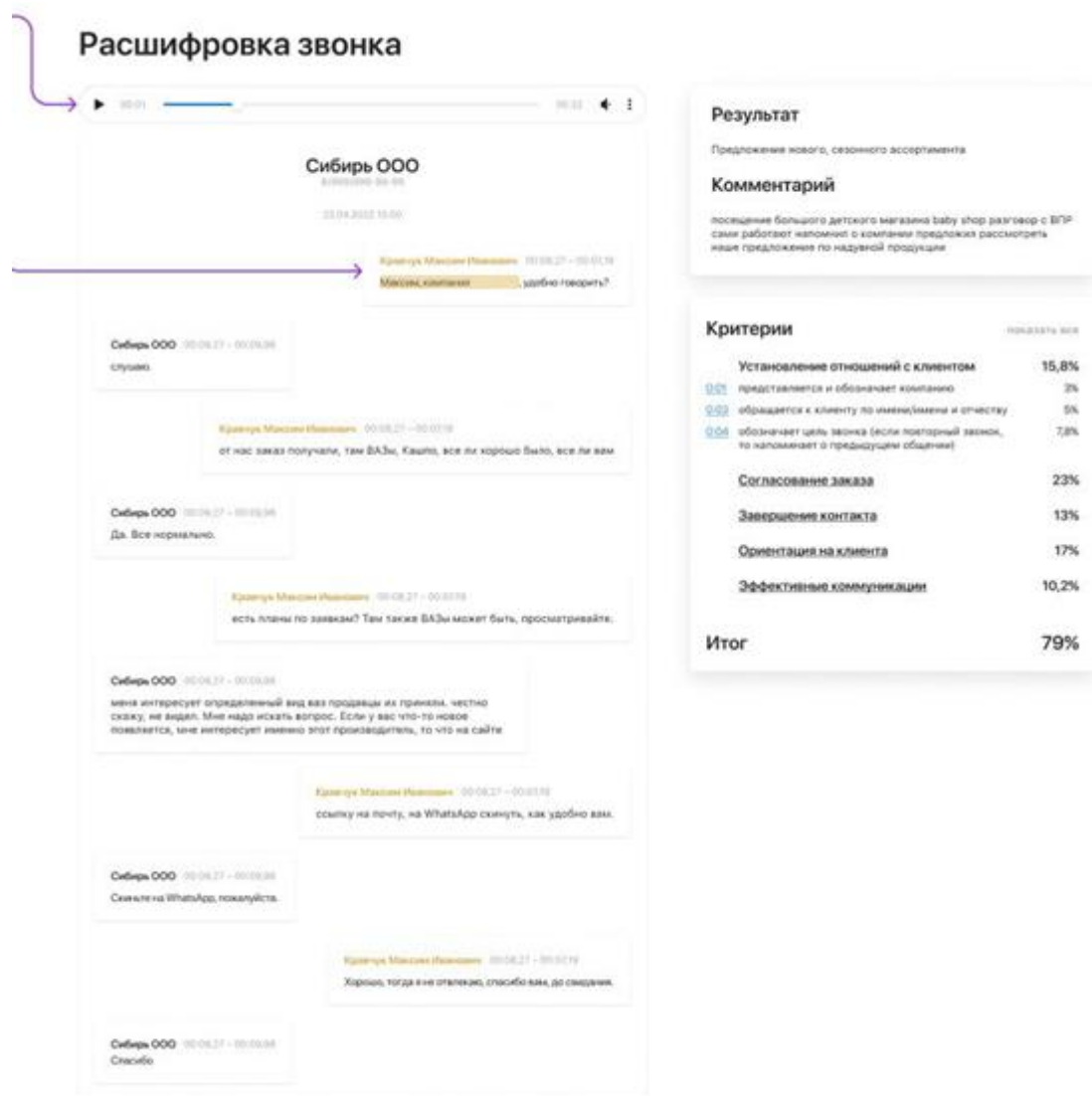


Рисунок 23 - имитационная модель “Расшифровка звонка”

#### Общая структура модели

1. Сбор данных: Запись звонков, включающих диалоги между менеджерами и клиентами.
2. Предобработка данных: Очистка аудиозаписей от шумов, нормализация громкости, разбиение на сегменты.
3. Распознавание речи: Использование нейронных сетей (например, Recurrent Neural Networks (RNN) или Transformer-based моделей, таких как BERT) для преобразования аудиозаписей в текст.
4. Анализ текста: Классификация и разметка текста на основе содержимого разговора (например, предложение товаров, обсуждение сезонного ассортимента).
5. Оценка качества: Оценка эффективности взаимодействия на основе заданных критериев.

#### Расшифровка звонка

#### Технологии и методы

Для расшифровки звонка используются современные алгоритмы распознавания речи, основанные на нейронных сетях:

Модели типа RNN: Модели с долгосрочной кратковременной памятью (LSTM) и gated recurrent unit (GRU).

Transformer-based модели: Использование архитектур, подобных BERT или GPT, для точного распознавания и интерпретации речи.

Эти модели обучаются на больших наборах данных с различными акцентами и шумовыми фонами, что позволяет достигать высокой точности в расшифровке.

Выводимый системой результат звонка

После расшифровки система должна выводить структурированный текст, который содержит полное содержание разговора. Например:

Текст расшифровки:

Менеджер: "Максим, компания X, удобно говорить?"

Клиент: "Слушаю"

На данном этапе видно, как аудио разговор преобразуется в текст. Таким образом, после преобразования, система начнет анализировать качество работы сотрудника, контекст диалога и его результат.

В поле “Результат” отображается цель звонка, который система определяет исходя из проанализированного диалога.

Также система автоматически генерирует краткий комментарий о содержании разговора, который может включать ключевые моменты и темы, обсуждавшиеся в ходе звонка.

Критерии оценки качества разговора

Для оценки качества взаимодействия менеджера с клиентом, система должна учитывать несколько ключевых критериев. Промежуточные итоги и итоговая оценка выражаются в процентном виде.

Критерии:

- Установление отношений с клиентом (20%)
- Согласование заказа (20%)
- Завершение контакта (20%)
- Ориентация на клиента (20%)
- Эффективные коммуникации (20%)

Система определяет таймлайн и подзадача, которую выполнял менеджер в данный момент. По итогам, выставляется оценка в % соотношении насколько корректно сотрудник выполнил работу. Следующим этапом, подсчитывается итоговый результат, выраженный в %, определяющий качество проведенного диалога.

Таким образом, модель представляет собой инструмент для распознавания и разметки голосовых записей, анализа обработанной информации и вывода итогового результата.

## **2.4 Формирование требований к проектируемой системе**

### **2.4.1 Требования к системе**

#### **2.4.1.1. Требования к структуре и функционированию системы**

Перечень подсистем для системы распознавания и разметки голосовых записей с применением нейронных сетей можно выделить следующие подсистемы, их назначение, основные характеристики, а также требования к уровням иерархии и степени централизации системы:

Подсистема "calls-fetcher" играет ключевую роль в системе, обеспечивая получение и загрузку аудиофайлов с звонками для последующей обработки и анализа. Ее целью является автоматический сбор информации о звонках и загрузка аудиофайлов, содержащих голосовые записи разговоров. Вот более подробное описание этой подсистемы:

##### **1. Назначение:**

- Автоматический сбор информации о звонках из различных источников, таких как телефонные системы, VoIP-серверы, CRM-системы и другие.
- Загрузка аудиофайлов, содержащих голосовые записи звонков, для последующей обработки и анализа.
- Подготовка и структурирование данных для дальнейшего использования в системе распознавания и разметки голосовых записей.

##### **2. Основные характеристики:**

- Поддержка различных источников данных, таких как телефонные системы, VoIP-серверы, CRM-системы и другие.
- Механизмы автоматического сбора информации о звонках, включая получение метаданных (например, даты, времени, участников звонка) и загрузку аудиофайлов.
- Обработка больших объемов данных и эффективная передача их для дальнейшей обработки в системе распознавания речи.
- Обеспечение безопасности и конфиденциальности данных во время их сбора и передачи.

##### **3. Требования к уровням иерархии и степени централизации системы:**

- Уровень иерархии: Подсистема "calls-fetcher" чаще всего является частью

централизованной системы, так как данные о звонках могут быть централизованно храниться и управляться в организации.

- Степень централизации: В зависимости от требований к системе и архитектуры приложения, "calls-fetcher" может быть реализована как централизованная, децентрализованная или дистрибутивная система. Например, если все звонки обрабатываются на едином сервере, подсистема будет централизованной. В то время как, если организация имеет несколько офисов с собственными телефонными системами, "calls-fetcher" может быть реализована децентрализованно с точками сбора данных в каждом офисе.

Подсистема "employee-info-fetcher" выполняет функцию получения дополнительной информации о инициаторе или отвечающем в рамках телефонного звонка. Ее целью является обогащение звонка дополнительными данными о сотруднике, такими как его должность, отдел, контактная информация и другие характеристики, которые могут быть необходимы для проведения дальнейшего анализа или для повышения эффективности обработки звонков. Вот более подробное описание этой подсистемы:

#### 1. Назначение:

- Получение дополнительной информации о сотрудниках из источников, таких как базы данных компании, системы управления персоналом или внешние сервисы.
- Обогащение звонков дополнительными данными о сотрудниках, участвующих в звонке.
- Предоставление информации о сотрудниках для дальнейшего анализа или использования в других приложениях.

#### 2. Основные характеристики:

- Интеграция с базами данных компании или другими системами управления персоналом для получения актуальной информации о сотрудниках.
- Механизмы обработки и хранения полученных данных с учетом требований к безопасности и конфиденциальности.
- Возможность работы в реальном времени для быстрого получения информации о сотрудниках во время звонка.
- Поддержка различных форматов и протоколов для интеграции с разными источниками данных.

### 3. Требования к уровням иерархии и степени централизации системы:

- Уровень иерархии: Подсистема "employee-info-fetcher" чаще всего является частью централизованной системы, так как информация о сотрудниках может быть централизованно храниться и управляться в компании.
- Степень централизации: В зависимости от особенностей компании и архитектуры приложения, "employee-info-fetcher" может быть реализована как централизованная, децентрализованная или дистрибутивная система. Например, если информация о сотрудниках хранится в единой централизованной базе данных, подсистема также может быть централизованной. Однако, если каждый отдел имеет свою собственную базу данных с информацией о сотрудниках, подсистема может быть реализована децентрализованно с возможностью доступа к данным из разных источников.

Подсистема "track-processor" выполняет важные функции по обработке аудиозаписей, которые не были определены или распознаны ранее в системе. Ее целью является отправка этих аудиозаписей на определение речи (speech-to-text), чтобы получить текстовое представление содержимого звонков. Вот более подробное описание этой подсистемы:

#### 1. Назначение:

- Обработка аудиозаписей, которые не были предварительно определены или распознаны в системе, например, из-за ошибок в процессе сбора или недостаточной информации о звонке.
- Отправка этих аудиозаписей на определение речи с целью получения текстового представления содержимого разговоров.
- Подготовка текстовых данных для дальнейшего анализа и обработки в системе, например, для выявления ключевых слов или паттернов.

#### 2. Основные характеристики:

- Механизмы автоматической обработки и отправки аудиозаписей на определение речи.
- Использование сервисов распознавания речи или моделей нейронных сетей для преобразования аудио в текст.
- Обработка различных форматов аудиофайлов и поддержка различных языков.
- Обеспечение точности и качества распознавания речи.



### 3. Требования к уровням иерархии и степени централизации системы:

- Уровень иерархии: Подсистема "track-processor" может быть частью централизованной системы, так как обработка и распознавание речи может осуществляться на едином сервере или наборе серверов.
- Степень централизации: В зависимости от объема и сложности аудиозаписей, а также требований к скорости обработки, "track-processor" может быть реализован как централизованная или децентрализованная система. Например, если обработка аудиозаписей происходит на высокопроизводительном сервере с доступом из различных источников, система будет централизованной. В то время как, если обработка происходит на мобильных устройствах или распределенных серверах, "track-processor" может быть реализована децентрализованно.

Подсистема "track-tag-assigner" в системе распознавания и разметки голосовых записей с применением нейронных сетей отвечает за назначение тегов аудиозаписям. Она создает связь между аудиозаписями (треками) и тегами, определяя, какие теги должны быть присвоены каждой записи. Вот более подробное описание этой подсистемы:

#### 1. Назначение:

- Назначение тегов аудиозаписям с целью классификации, категоризации или идентификации их содержимого.
- Создание связи между аудиозаписями и соответствующими тегами для обеспечения последующего анализа и поиска по ним.
- Управление процессом назначения тегов, включая создание, обновление и удаление связей между треками и тегами.

#### 2. Основные характеристики:

- Использование таблицы с очередью для управления операциями назначения тегов, где создаются записи для каждого трека в момент создания нового тега.
- Механизмы для определения соответствия между аудиозаписями и тегами на основе содержания записей или дополнительной метаданных информации.
- Обеспечение целостности и безопасности данных при операциях назначения и удаления тегов.
- Отслеживание статуса операций назначения тегов для контроля и мониторинга

процесса.

### 3. Требования к уровням иерархии и степени централизации системы:

- Уровень иерархии: Подсистема "track-tag-assigner" может быть частью централизованной системы, так как операции назначения тегов могут осуществляться на едином сервере или наборе серверов.
- Степень централизации: В зависимости от требований к системе и архитектуры приложения, "track-tag-assigner" может быть реализована как централизованная, децентрализованная или дистрибутивная система. Например, если все операции назначения тегов происходят на высокопроизводительном сервере с доступом из различных источников, система будет централизованной. В то время как, если операции выполняются на различных серверах или устройствах, "track-tag-assigner" может быть реализована децентрализованно.

Подсистема "report" в системе отвечает за создание отчетов на основе обработанных данных. Компонент "report" создает отчеты, которые предоставляют пользователю информацию о различных аспектах обработки голосовых записей, результатов распознавания и разметки. Вот более подробное описание этой подсистемы:

#### 1. Назначение:

- Создание отчетов, предоставляющих пользователю информацию о различных аспектах обработки голосовых записей, результатов распознавания и разметки.
- Предоставление пользователю аналитических данных о производительности системы, точности распознавания, распределении ключевых фраз и других метриках.
- Обеспечение возможности мониторинга и анализа работы системы, а также принятия решений на основе полученных данных.

#### 2. Основные характеристики:

- Генерация отчетов в различных форматах, таких как текстовые документы, таблицы, графики, диаграммы и дашборды.
- Возможность настройки содержания отчетов в соответствии с потребностями пользователей и особенностями проекта.
- Использование данных из различных источников, включая результаты обработки голосовых записей, статистические данные, метаданные и т.д.

- Обеспечение безопасного доступа к отчетам для соответствующих пользователей с учетом прав доступа и конфиденциальности данных.

### 3. Требования к уровням иерархии и степени централизации системы:

- Уровень иерархии: Подсистема "report" часто является частью централизованной системы, так как отчеты обычно создаются на основе данных, централизованно хранимых и обрабатываемых в системе.
- Степень централизации: Поскольку отчеты основываются на общих данных и метриках, необходимо обеспечить централизацию данных для создания единых и надежных отчетов. Пользователи из разных подразделений или уровней иерархии могут иметь доступ к различным отчетам в зависимости от их ролей и функций.

Подсистема "speech-to-text" (STT) играет важную роль в системе распознавания и разметки голосовых записей с применением нейронных сетей. Ее целью является преобразование аудиозаписей с голосовым содержанием в текстовый формат. Вот более подробное описание этой подсистемы:

#### 1. Назначение:

- Преобразование речевых сигналов, содержащих голосовую речь, в текстовый формат с помощью алгоритмов и моделей машинного обучения.
- Поддержка процесса анализа и обработки голосовых данных, делая их доступными для дальнейшей обработки и анализа.
- Обеспечение возможности поиска, анализа и хранения текстовой информации, полученной из аудиозаписей.

#### 2. Основные характеристики:

- Использование передовых алгоритмов и моделей машинного обучения для распознавания речи с высокой точностью.
- Поддержка различных языков и диалектов, а также способность адаптироваться к различным голосам и акцентам.
- Возможность работы в режиме реального времени для немедленного преобразования речи в текст во время воспроизведения аудио.
- Интеграция с другими компонентами системы для дальнейшей обработки и

анализа текстовой информации.

### 3. Требования к уровням иерархии и степени централизации системы:

- Уровень иерархии: подсистема "speech-to-text" обычно является частью централизованной системы, так как обработка и распознавание речи обычно требует высоких вычислительных ресурсов и доступа к данным.
- Степень централизации: для обеспечения единой и надежной работы системы обработки речи рекомендуется централизованное размещение моделей и алгоритмов на высокопроизводительных серверах. Однако для специфических случаев, например, при работе с мобильными устройствами или в режиме реального времени, может потребоваться децентрализованный или распределенный подход.

Возможности, которые пользователь может использовать и которые система должна автоматически выполнять:

Функционал для пользователя:

- Возможность просмотра совершенных звонков сотрудника.
- Сортировка звонков по чек-листам, тегам, отчетам, по интервалу, по сотруднику, по типу сегмента, по типу отвечающего и по цели.
- Отображение звонков в структурированном виде.
- Просмотр эффективности сотрудника исходя из отчетов.
- Прослушивание аудиозаписи разговора с отображением распознанного текста и выделением фраз, произносимых в текущий момент диалога.

Функционал для системы:

- Получение и обработка записей разговоров.
- Выявление совпадений в распознанных записях по тегам и чек-листам.
- Составление отчетов.
- Мониторинг и анализ качества работы системы.
- Создание записей о событиях и аудита с целью последующего исследования и составления отчетов.

Исходя из вышесказанного, система распознавания и разметки голосовых записей должна обеспечивать пользователям удобные инструменты для эффективного анализа аудио

дорожек в текстовом виде, а также спектр возможностей для работы с полученной после транскрибации информацией.

Система распознавания и разметки голосовых записей должна поддерживать следующие режимы функционирования:

Нормальный режим:

В нормальном режиме система должна обеспечивать непрерывный процесс распознавания и разметки голосовых записей с использованием нейронных сетей. Она должна непрерывно собирать аудиоданные и обрабатывать их для извлечения текстовой информации. Помимо этого, система должна корректно анализировать распознанный текст, предоставляя пользователю информацию о содержании записей и разметке. Также, она должна гарантировать выполнение основных функций, таких как: настройка нейронных сетей, сбор и обработка данных, мониторинг процесса распознавания, управление данными и предоставление отчетности.

Аварийный режим:

В аварийном режиме система должна немедленно остановить процесс распознавания и разметки голосовых записей. Это означает приостановку всех текущих операций и предотвращение создания новых задач обработки. Однако, система не должна удалять текущие задачи. Дополнительно, система должна оперативно информировать администраторов и других уполномоченных лиц о возникшем аварийном случае, предоставляя подробную информацию о его причинах и принимаемых мерах для устранения проблемы. Для обеспечения непрерывности работы в условиях аварии система должна автоматически активировать резервные ресурсы с целью сокращения времени простоя и обеспечения доступности сервисов. Это может включать переключение на альтернативные серверы или другие аспекты инфраструктуры. Также необходимо, чтобы система сохраняла подробную информацию о произошедшей аварии, включая записи событий, принятые действия и результаты их выполнения. Это позволит провести анализ ситуации и принять меры для предотвращения подобных инцидентов в будущем. После восстановления нормальной работы система должна предоставить возможности для восстановления данных и настроек, а также осуществить проверку работоспособности всех компонентов перед возобновлением обработки голосовых записей.

Требования к диагностированию системы:

- Мониторинг и диагностика:
- Использование стандартных инструментов для мониторинга и диагностики,

включая:

- Средства администрирования баз данных для отслеживания работы системы управления базами данных.
- Инструменты ETL для контроля процессов извлечения, трансформации и загрузки данных.
- Визуализационные инструменты для мониторинга работы системы и отображения данных.

Ведение журналов:

- Запись инцидентов в электронном формате.
- Создание графиков и журналов проведения профилактических работ.

Техническое обслуживание:

- Постоянное и организованное обследование состояния и обслуживание всех элементов системы.

## **2.4.2 ТРЕБОВАНИЯ К НАДЕЖНОСТИ**

### **2.4.2.1. СОСТАВ ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ ДЛЯ СИСТЕМЫ В ЦЕЛОМ**

Для системы распознавания и разметки голосовых записей с применением нейронных сетей требования к надежности также играют ключевую роль, учитывая значимость операций и автоматизацию процессов. Вот основные требования к надежности такой системы:

#### **1. Высокая доступность:**

- Система должна быть доступна для использования полные рабочие часы, учитывая потенциальную нагрузку поступление аудио информации для обработки.
- Необходимы механизмы обеспечения непрерывной доступности, такие как резервирование серверов и сетевых соединений.

#### **2. Стабильная работа:**

- Система должна гарантировать низкую вероятность возникновения проблем при функционировании основных компонентов.
- Проведение регулярных тестов на стойкость к нагрузкам и неожиданным ситуациям помогает обеспечить стабильную работу.

### 3. Защита от атак:

- Система должна быть защищена от вредоносных атак, включая DDoS-атаки, взломы и вирусы.
- Применение современных методов защиты, таких как брандмауэры и антивирусное ПО, обеспечивает безопасность системы.

### 4. Безопасность данных:

- Все данные, обрабатываемые и хранимые системой, должны быть защищены от несанкционированного доступа и утечек.
- Использование методов шифрования и строгих политик доступа помогает обеспечить безопасность данных.

### 5. Резервное копирование и восстановление:

- Система должна регулярно создавать резервные копии данных для возможности быстрого восстановления в случае сбоев или аварий.
- Резервные копии данных должны храниться в защищенных местах с ограниченным доступом.

### 6. Мониторинг и управление:

- Система должна быть оборудована средствами мониторинга ее состояния и автоматического реагирования на возникающие проблемы.
- Настройка систем оповещения и автоматических мер по восстановлению помогает поддерживать надежную работу системы.

### 7. Регулярные обновления и аудит безопасности:

- Система должна регулярно обновляться для исправления уязвимостей и обеспечения безопасности.
- Проведение регулярных аудитов безопасности помогает выявить потенциальные угрозы и уязвимости.

## 2.4.3 ТРЕБОВАНИЯ К КВАЛИФИКАЦИИ ПЕРСОНАЛА

#### Пользователь сервиса:

- Пользователь должен обладать доступом к сервису, то есть быть в белом листе для работы с системой.
- Использование платформы должно соответствовать ее целям, не нарушая

законодательство и правила пользования системой.

- Пользователь должен бережно обращаться с конфиденциальной информацией и не передавать ее третьим лицам и/или лицам, не находящимся в белом листе сервиса.
- Пользователю необходимо обладать достаточной компетенцией для анализа информации, предоставляемой после работы системы
- Пользователь имеет возможность получить оперативную и качественную поддержку от администрации системы при возникновении вопросов или проблем с использованием платформы.

Администратор платформы:

- Администратор системы должен обладать достаточным уровнем компетенции в области обработки голосовых данных и нейронных сетей.
- Администратор системы следит за работоспособностью всех модулей и компонентов платформы, проводя периодические проверки и контроль их функциональности.
- Администратор обязан следить за обновлениями и актуализацией информации о методах обработки голоса и современных технологиях в этой области.
- Он должен оперативно реагировать на запросы и обращения пользователей, предоставляя качественную и своевременную поддержку.
- Администратор регулярно обновляет и улучшает функционал системы в соответствии с потребностями пользователей и последними достижениями в области распознавания и разметки голосовых записей.
- Контроль за действиями пользователей на платформе также лежит на администраторе, включая выявление нарушений правил использования системы.
- Он обеспечивает адекватную поддержку пользователям, консультируя и обучая их по всем вопросам, связанным с использованием системы.
- Администратор содействует эффективному взаимодействию с разработчиками и технической поддержкой для решения любых проблем, связанных с функционированием платформы.
- Он соблюдает принципы этики и профессиональной ответственности в



общении с пользователями и обработке их данных.

#### **2.4.4 ПОКАЗАТЕЛИ НАЗНАЧЕНИЯ**

Система распознавания и разметки голосовых записей с применением нейронных сетей должна обладать высокой приспособляемостью к изменениям в методах анализа и обработки аудиоданных. Требуется обеспечить возможность интеграции новых методов обучения нейронных сетей без кардинальных изменений в программном обеспечении системы. Важно предусмотреть возможность масштабирования системы по производительности и объему обрабатываемой информации путем модернизации используемых технических средств и алгоритмов.

Система должна обеспечивать возможность хранения исторических данных с глубиной не менее 1-2 лет для последующего анализа и использования в обучении нейронных сетей. При этом, для подсистемы операционной деятельности система должна поддерживать одновременную работу более 5 пользователей, а для остальных подсистем не менее 2 пользователей. Система должна обеспечивать формирование отчета за время не более 10 минут. Система должна обеспечивать преобразование аудиозаписей в текст за время, не превышающее 12 часов. Система должна предоставлять отчетность в готовом виде в течении 24 часов.

Для системы распознавания и разметки голосовых записей с применением нейронных сетей можно определить следующие параметры, характеризующие ее соответствие целям использования:

1. Гибкость и адаптивность к изменениям в процессах и методах обработки данных, а также к изменениям параметров записей:

- Система должна обладать высокой приспособляемостью к изменениям в структуре и характере голосовых данных, таким как изменения в интонации, акценте и скорости речи.
- Она должна предоставлять возможности гибкой настройки алгоритмов распознавания и разметки, а также способности адаптироваться к различным стилям и особенностям речи.

2. Возможности модернизации и развития системы:

- Система должна быть разработана с учетом потенциала для будущих модернизаций и расширений функционала.
- Ее архитектура должна быть гибкой и масштабируемой, позволяя легко внедрять изменения и добавлять новые функции в процессе ее развития.

### 3. Временные показатели и доступность:

- Система должна обеспечивать высокую доступность и надежность, чтобы сохранить свою функциональность в различных ситуациях.
- Важно устанавливать целевые показатели для времени обработки записей, доступности системы и скорости реакции на изменения в аудиоданных.

Для представления условно-функционального описания системы распознавания и разметки голосовых записей с применением нейронных сетей в цифрах, можно выделить следующие ключевые показатели:

Точность распознавания голоса:

Этот показатель отражает процент правильно распознанных голосовых команд или фраз системой относительно общего числа вводимых данных за определенный период времени. Чем выше этот показатель, тем более эффективной является система в распознавании голоса.

Скорость обработки голосовых данных:

Данный показатель представляет собой среднее время, за которое система обрабатывает одну голосовую запись или команду. Чем быстрее система обрабатывает данные, тем более эффективной она является в работе.

Точность разметки голосовых записей:

Это процент правильно размеченных данных в голосовых записях относительно общего числа размеченных данных. Чем выше этот показатель, тем более точной и надежной является система в разметке голосовых записей.

Время реакции на запросы:

Этот параметр показывает, как быстро система реагирует на голосовые команды и запросы пользователей, генерируя соответствующие ответы или действия. Быстрая реакция на запросы может повысить удовлетворенность пользователей и эффективность системы.

Количество обрабатываемых записей в единицу времени:

Этот показатель указывает на производительность системы в обработке голосовых записей. Высокое количество обрабатываемых записей может свидетельствовать о высокой производительности и эффективности системы.

Надежность и стабильность работы:

Этот параметр отражает процент времени, в течение которого система работает без

сбоев или ошибок. Высокая надежность и стабильность работы важны для обеспечения бесперебойной работы системы и удовлетворения потребностей пользователей.

Таким образом, система должна быть способна обрабатывать голосовые данные с высокой производительностью, обеспечивая эффективное взаимодействие с пользователями и сохранение данных для долгосрочного анализа и обучения нейронных сетей.

#### **2.4.5.1 ТРЕБОВАНИЯ К ПРИСПОСОБЛЯЕМОСТИ СИСТЕМЫ К ИЗМЕНЕНИЯМ**

Для обеспечения приспособляемости системы распознавания и разметки голосовых записей с применением нейронных сетей важно учитывать следующие аспекты:

Регулярное администрирование:

Предусмотреть систематическое администрирование системы для поддержания её работоспособности и обновления компонентов.

Оптимизация работы алгоритмов:

Необходимо постоянно улучшать алгоритмы распознавания и разметки голосовых записей для оптимизации скорости обработки данных и повышения точности результатов.

Модернизация процессов обработки данных:

Обеспечить возможность модернизации процессов обработки голосовых данных в соответствии с развитием технологий и изменениями в требованиях к системе.

Гибкость в доступе и представлении данных:

Предусмотреть гибкость в доступе и представлении данных для пользователей, позволяя им настраивать интерфейс и получать необходимую информацию в соответствии с их потребностями.

Настройка и конфигурирование ПО:

Обеспечить наличие настроечных и конфигурационных файлов для программного обеспечения системы, чтобы пользователи могли гибко настраивать и конфигурировать систему под свои нужды и требования.

##### **2.4.5.1.2. ТРЕБОВАНИЯ К ПОЛЬЗОВАТЕЛЬСКОМУ ИНТЕРФЕЙСУ И УДОБСТВУ ИСПОЛЬЗОВАНИЯ**

Внешний вид:

1. Структурированные интерфейсы подсистем.
2. Обеспечение локализации интерфейса на русском языке.

3. Использование кроссплатформенного решения с адаптивной версткой для удобства пользователей.

4. Эффективная навигация по интерфейсу.

5. Выполнение операций с высокой степенью успешности (95–100%).

6. Минимизация количества кликов до любого модуля или функции (не более 2–3).

7. Быстрый отклик на изменения в материале (не более 300 мс).

Взаимодействие с пользователем:

1. Повторное использование предыдущих вводимых данных.

2. Предусмотрены "горячие" клавиши для частых операций.

3. Возможность восстановления после ошибок пользователя и контроля системных ресурсов для минимизации последствий ошибок.

4. Отображение сообщений об ошибках на экране с рекомендациями по их устранению на русском языке.

Процедуры ввода-вывода данных:

1. Поддержка многомерного анализа данных в табличном и графическом форматах.

#### **2.4.5.1.3. ТРЕБОВАНИЯ К ЗАЩИТЕ ИНФОРМАЦИИ ОТ НЕСАНКЦИОНИРОВАННОГО ДОСТУПА**

##### **2.4.5.1.3.1. ТРЕБОВАНИЯ К ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ**

Требования к безопасности информации в системе распознавания и разметки голосовых записей с применением нейронных сетей являются критически важными, учитывая чувствительность данных и потенциальные угрозы безопасности. Вот основные требования к защите информации:

Аутентификация:

Гарантировать строгую аутентификацию всех пользователей с использованием сильных паролей, механизмов двухфакторной аутентификации и других методов подтверждения личности.

Авторизация:

Обеспечить каждому пользователю строго определенные права доступа к различным функциям и данным системы. Администраторам предоставить возможность гибкого управления правами доступа.

Шифрование данных:

Зашифровать всю передаваемую и хранимую информацию для защиты от несанкционированного доступа. Использовать современные алгоритмы шифрования с длинными ключами для обеспечения безопасности данных.

Защита сетевого периметра:

Применить меры защиты сетевого периметра, включая брандмауэры, VPN-соединения и другие средства. Реализовать политики безопасности для ограничения доступа к сетевым ресурсам.

Мониторинг и обнаружение инцидентов:

Обеспечить мониторинг сетевой активности, входящих и исходящих запросов, а также поведения пользователей. Реализовать системы обнаружения нештатных ситуаций для быстрого реагирования на потенциальные угрозы.

Регулярное обновление и обучение персонала:

Регулярно обновлять систему для обеспечения актуальной защиты и обучать персонал по правилам безопасности информации и соблюдению политик защиты данных.

Эти требования гарантируют надежную защиту информации от несанкционированного доступа в системе распознавания и разметки голосовых записей с применением нейронных сетей.

Для обеспечения сохранности информации в системе распознавания и разметки голосовых записей с применением нейронных сетей в аварийных ситуациях необходимо установить определенные требования. Вот основные из них:

1. Регулярное резервное копирование данных:

Система должна периодически создавать резервные копии всех важных данных, включая голосовые записи, результаты распознавания, метаданные и технические настройки. Резервные копии должны храниться в отдельном безопасном хранилище, обеспечивающем защиту от потери данных в случае аварии.

2. Автоматизированное восстановление данных:

При возникновении аварийных ситуаций система должна автоматически восстанавливать данные из резервных копий. Это позволит минимизировать простой системы и предотвратить потерю важной информации.

3. Защита от повреждения данных:

Система должна обладать методами защиты данных от случайного или умышленного повреждения, такие как механизмы, которые предотвращают возможность записи или удаления файлов без соответствующих разрешений доступа. Также, важно внедрить инструменты для мониторинга целостности данных с целью обнаружения повреждений и их автоматического восстановления.

#### 4. Аварийное отключение и перенос работы:

Если восстановление работы системы на текущем месте невозможно, она должна быть способна к аварийному отключению и переносу работы на резервные серверы или инфраструктуру. Это позволит сократить время простоя и обеспечить сохранность данных.

#### 5. Тестирование восстановления после аварий:

Необходимо регулярно тестировать процедуры восстановления после аварийных ситуаций, чтобы проверить их эффективность и сократить время восстановления. Это включает проверку восстановления из резервных копий и переход на резервные ресурсы для обеспечения непрерывности работы системы.

Требования к обеспечению информационной безопасности системы распознавания и разметки голосовых записей с применением нейронных сетей:

- Для обеспечения защиты системы необходимо использовать комплекс программно-технических средств и соответствующих организационных мер.
- Система должна быть защищена на всех этапах обработки информации и во всех режимах работы, даже во время проведения технического обслуживания и регламентных работ.
- Программно-технические средства безопасности не должны значительно снижать основные функциональные характеристики системы, такие как надежность, скорость работы и возможность изменения конфигурации.
- Разграничение прав доступа пользователей и администраторов системы должно осуществляться в соответствии с принципом "запрещено все, что не разрешено".

### **ВЫВОД ПО 2 ГЛАВЕ:**

Таким образом, математическое моделирование является важным этапом работы, поскольку оно формулирует процесс анализа текста по чеклистам с математической точки зрения. Концептуальное моделирование, в свою очередь, декомпозирует систему, разбивая

её на мелкие сущности и их связи. ER-диаграмма Питера Чена объединяет эти сущности в одно целое, представляя работу сервиса как единый механизм. Функциональное моделирование позволяет проанализировать функционирование существующих платформ и разрабатываемой в рамках данного проекта системы, что подчеркивает актуальность выбранной темы. Имитационное моделирование позволяет провести апробацию модели и, на основе результатов тестирования, определить качество работы инструмента. Требования к системе, определяют условия для разрабатываемого сервиса.

Функционал будущей системы будет включать в себя:

Распознавание речи: Одной из ключевых задач является точное и эффективное распознавание речи из аудиозаписей различного качества и стилей. Это включает в себя различение слов, выражений, интонаций, акцентов и даже эмоциональных состояний говорящего.

Обработка больших объемов данных: С увеличением объемов аудиоданных становится важным обеспечить эффективную обработку и анализ голосовых записей в реальном времени или в пакетном режиме.

Обеспечение высокой точности и скорости: необходимо достичь высокой точности распознавания и разметки речи при сохранении быстродействия системы, что особенно важно для приложений, требующих реального времени.

Обработка информации о собеседнике: необходимо определить тип контрагента, целевой смысл диалога с клиентом, а также цели, преследуемые сотрудником в данном диалоге.

Итоговая информация о разговоре: необходимо проанализировать работу сотрудника при контакте с клиентами, выявить проблемные стороны разговора и недочеты, и исходя из данной информации поставить оценку сотруднику за проведенный диалог.

Таким образом у менеджера отдела и вышестоящих лиц, появиться возможность в короткие сроки определить качество работы сотрудника и его компетентность для принятия решений о его дальнейшей судьбе в составе компании.

### 3 ПРОВЕДЕНИЕ ЭКСПЕРИМЕНТОВ И АНАЛИЗ РЕЗУЛЬТАТОВ

					Д 09.04.01 2022 22286159 ВКРМ	Лист
		Осинцев А.В.		06.06.24		
Студ.		Фамилия	Подп.	Дата		56



## **3 ПРОВЕДЕНИЕ ЭКСПЕРИМЕНТОВ И АНАЛИЗ РЕЗУЛЬТАТОВ**

### **3.1 Проектирование информационной системы**

#### **3.1.1 Выбор архитектуры информационной системы**

Микросервисная архитектура является популярным подходом к разработке сложных и масштабируемых приложений. Она подразумевает разделение системы на независимые сервисы, каждый из которых отвечает за выполнение определённых задач. В рамках диссертации на тему “Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей”, использование микросервисной архитектуры позволяет использовать множество преимуществ.

Основные принципы микросервисной архитектуры:

- Разделение на независимые сервисы: Каждый микросервис отвечает за определённую функцию.
- Межсервисное взаимодействие: Сервисы взаимодействуют друг с другом через чётко определённые API.
- Автономность и независимость: Каждый сервис может разрабатываться, деплоиться и масштабироваться независимо от других.
- Использование различных технологий: Каждый микросервис может быть написан на наиболее подходящем для конкретной задачи языке программирования и использовать нужные инструменты и библиотеки.

Преимущества микросервисной архитектуры:

#### **1. Гибкость и масштабируемость:**

Легко масштабировать отдельные сервисы в зависимости от их нагрузки. Например, если сервис распознавания речи требует больше ресурсов, его можно масштабировать независимо от других сервисов.

#### **2. Ускорение разработки:**

Команды могут работать над разными сервисами параллельно, что ускоряет процесс разработки и внедрения новых функций.

#### **3. Повышенная отказоустойчивость:**

Отказ одного сервиса не приводит к полному падению системы. Например, если сервис разметки временно недоступен, остальные сервисы продолжают работать.

#### **4. Легкость в тестировании и деплое:**

Микросервисы можно тестировать и развертывать независимо друг от друга, что упрощает процессы CI/CD (Continuous Integration/Continuous Deployment).

#### 5. Разделение ответственности:

Каждый сервис имеет свою чёткую зону ответственности, что упрощает управление кодовой базой и способствует лучшему пониманию системы.

В связи с нюансами и особенностями разрабатываемой платформы, используется Микросервисная архитектура. Основные критерии, повлиявшие на выбор, представлены ниже:

##### Сложность и разнообразие задач:

Распознавание и разметка голосовых записей с применением нейронных сетей включает в себя различные этапы обработки данных (предварительная обработка, распознавание, разметка, анализ и т.д.). Микросервисная архитектура позволяет разрабатывать и оптимизировать каждый этап отдельно.

##### Требования к производительности:

Возможность масштабировать наиболее ресурсоемкие сервисы (например, нейронные сети для распознавания речи) позволяет эффективно распределять ресурсы.

##### Модульность и возможность расширения:

Легко добавлять новые функции и модули без необходимости вносить изменения в уже существующие сервисы. Это важно в исследовательских проектах, где часто появляются новые гипотезы и эксперименты.

##### Поддержка современных методов DevOps:

Использование оркестрации (Kubernetes) хорошо сочетается с микросервисной архитектурой, что облегчает управление инфраструктурой и ускоряет процессы деплоя.

Выбор микросервисной архитектуры для системы распознавания и разметки голосовых записей с применением нейронных сетей является обоснованным решением, учитывая преимущества в гибкости, масштабируемости и автономности разработки. Такой подход позволяет более эффективно управлять сложной системой, улучшать производительность и поддерживать модульность, что особенно важно в контексте научных исследований и разработок, где требуются постоянные изменения и адаптации.

### 3.2 Инфологическое проектирование

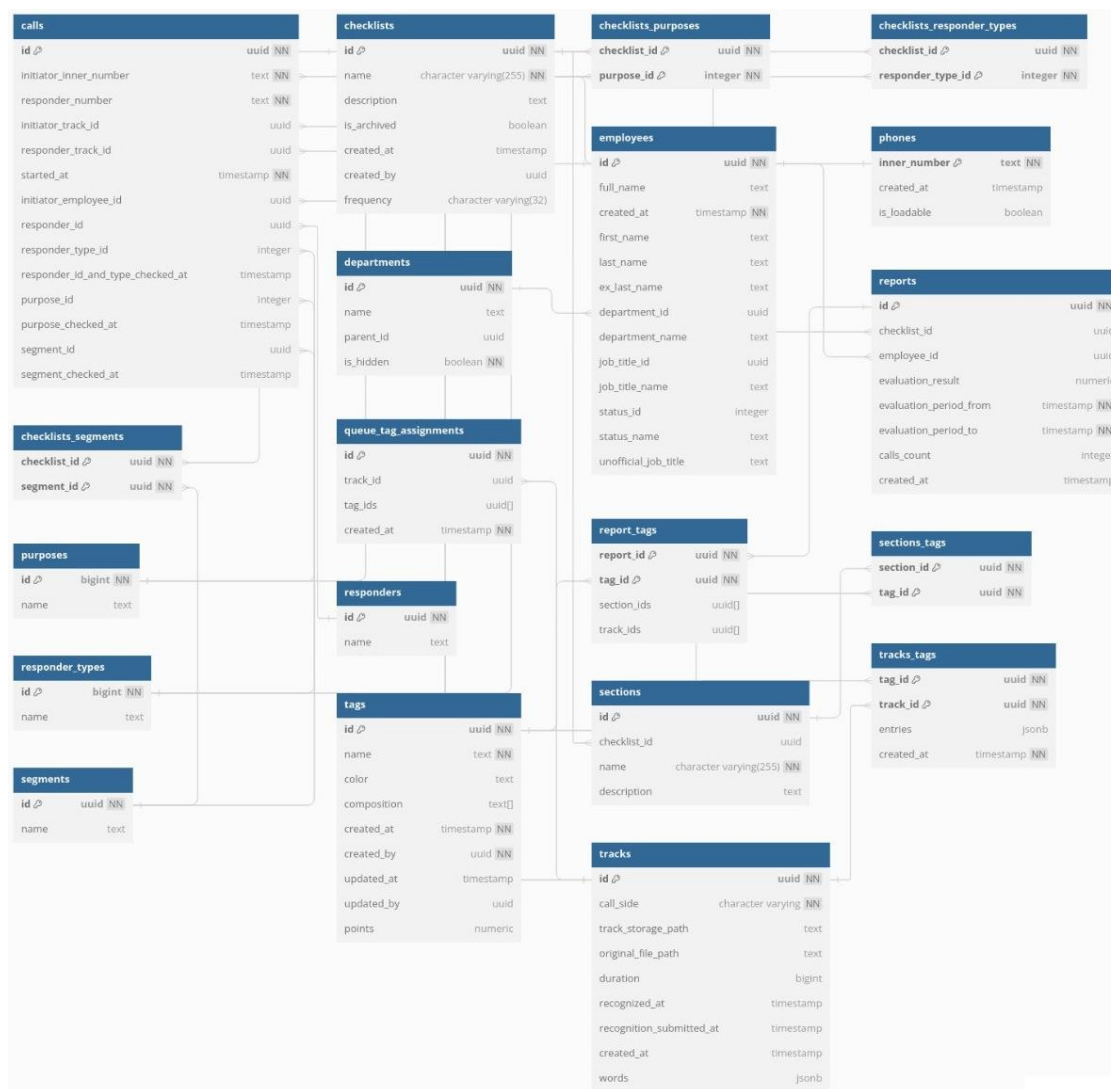


Рисунок 24 - инфологическая модель

Данная инфологическая модель (см. Рисунок 24) представляет собой диаграмму сущность-связь, которая отображает объекты и их взаимосвязи в системе, связанной с отчетами, звонками, чеклистами и сотрудниками.

Звонки (calls) в системе характеризуются уникальным идентификатором (UUID), внутренним номером инициатора и номером абонента, который отвечает на звонок. Также фиксируются идентификаторы дорожек (track) для обеих сторон, время начала звонка и идентификаторы сотрудников и абонентов. В системе учитываются типы абонентов, цели звонков и сегменты, а также время их проверки.

Чеклисты (checklists) включают уникальный идентификатор, название, описание и статус архивации. Фиксируются время создания чеклиста, идентификатор создателя и частота выполнения.

Цели чеклистов (checklists\_purposes): Каждый чеклист может быть связан с одной или несколькими целями, для чего используются идентификаторы чеклиста и цели.

Типы отвечающих чеклиста (checklists\_responder\_types): чеклисты также могут иметь связь с различными типами отвечающих, что позволяет создать разделение чеклистов между

различными типами клиентов.

Сотрудники (employees) в системе имеют уникальный идентификатор, полное имя, время создания записи и прочую информацию, такую как имя, фамилия, дополнительная фамилия, отдел, должность и статус

Телефонные номера (phones) сотрудников включают внутренний номер, время создания записи и статус загрузки телефона.

Отчеты (reports) связаны с чеклистами и сотрудниками через уникальные идентификаторы. Они содержат результаты оценки, период оценки, количество звонков и время создания.

Отделы (departments) характеризуются уникальным идентификатором, названием и статусом видимости.

Назначения тегов в очереди (queue\_tag\_assignments): Для назначения тегов в системе используются уникальные идентификаторы, идентификаторы дорожек и массив идентификаторов тегов, а также время создания назначения.

Отвечающие (responders): содержит список отвечающих с информацией о них

Теги (tags) характеризуются уникальным идентификатором, названием, цветом, составом, временем создания и обновления, а также количеством присвоенных очков.

Теги отчетов (report\_tags): Отчеты могут содержать теги, что фиксируется через идентификаторы отчетов и тегов, а также массив идентификаторов разделов и дорожек.

Разделы/Секции (sections) включают уникальный идентификатор, идентификатор чеклиста, название и описание.

Теги разделов (sections\_tags) могут быть связаны с тегами через идентификаторы разделов и тегов.

Сегменты (segments) в системе характеризуются уникальным идентификатором и названием.

Дорожки (tracks) включают уникальный идентификатор, сторону звонка (например, входящий, исходящий), путь к файлу дорожки, исходный путь к файлу, длительность, время распознавания и подачи на распознавание, время создания и массив слов.

Сегменты чеклистов (checklists\_segments) могут быть связаны с сегментами через идентификаторы чеклиста и сегмента.

Типы отвечающих (responder\_types) лиц включают уникальный идентификатор и название типа отвечающего

Эти поля и их связи позволяют эффективно организовать данные о звонках, чеклистах, сотрудниках и отчетах, обеспечивая структуру для хранения и обработки информации в системе.системе.

### 3.3 Проектирование интерфейса

На рисунке 25 изображен список отделов таким образом, каким его видит сотрудник компании в своем аккаунте. По клику можно перейти в любой из доступных разделов. Помимо этого, доступен поиск по контрагенту, выделенный синим светом.

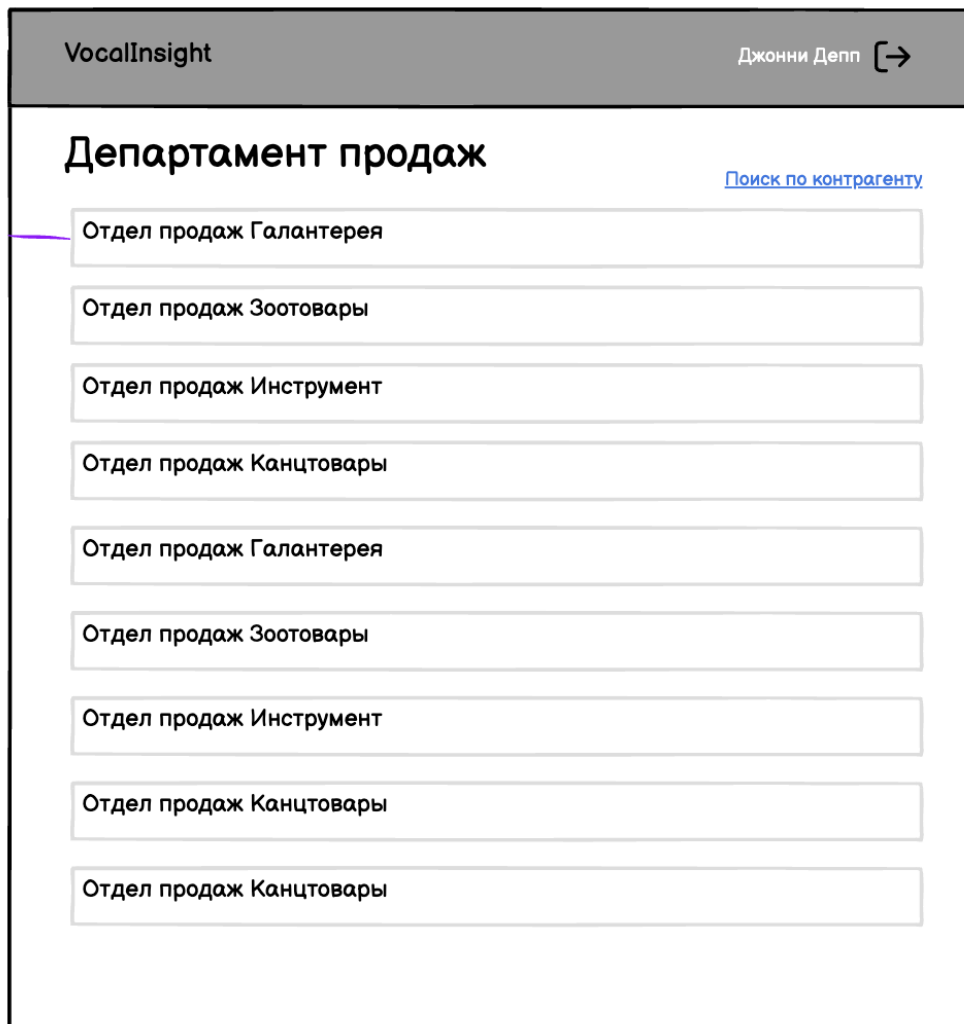


Рисунок 25 - Департамент продаж

При переходе в “отдел продаж Галантерея” сотрудник увидит страницу, изображенную на рисунке 26.

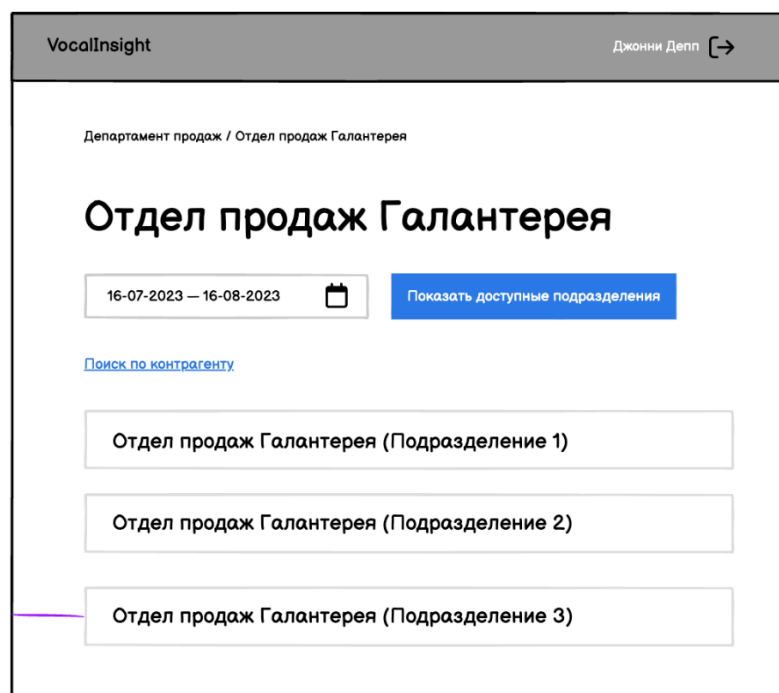


Рисунок 26 - Отдел продаж Галантерея

После перехода в раздел “Отдел продаж Галантерея (Подразделение 3)” отобразится страница, которая показана на рисунке 27. Данный раздел сайта показывает сотрудников, работающих по данному направлению. Помимо этого, можно увидеть поиск по контрагенту, выбрать временной промежуток, исходя из которого будет показан результат, а также доступна сортировка по всем отображаемым пунктам.

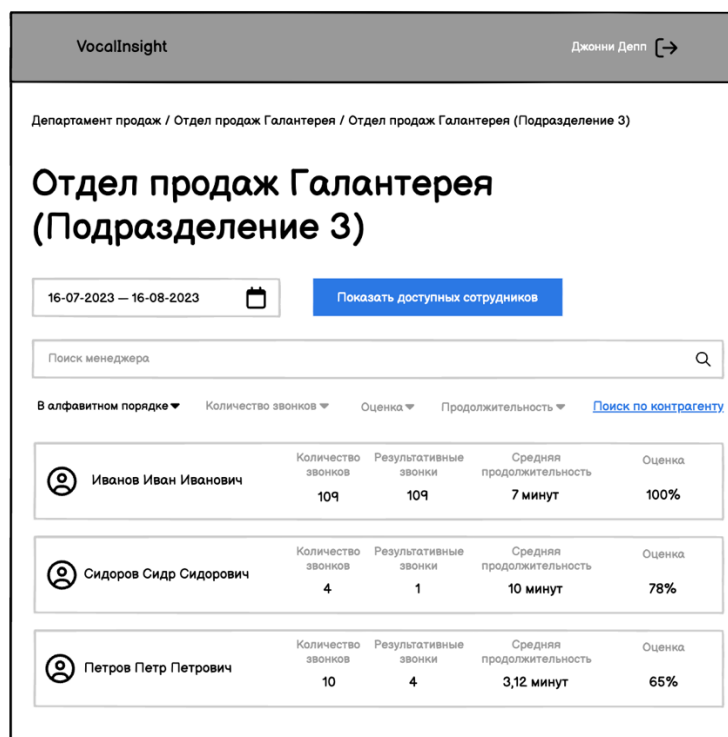


Рисунок 27 - Отдел продаж Галантерея (Подразделение 3)

Перейдем в поиск по контрагентам, где можно посмотреть всю хронологию звонков

определенному клиенту, подходящую под заданные параметры, можно увидеть на рисунке 28. Также, можно осуществить поиск по параметрам или организации клиента.

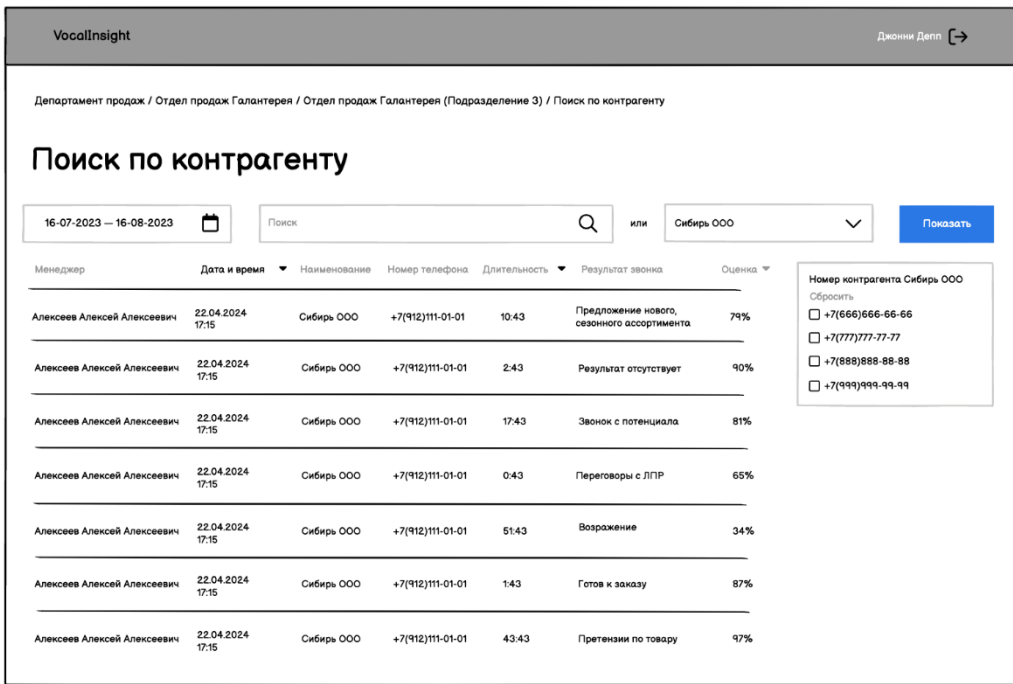


Рисунок 28 – Поиск по контрагенту

Погрузившись в сотрудника, отобразится страница с расшифровкой звонка (рисунок 29). В данном разделе сайта представлена информация о звонке, а именно: голосовая запись, расшифрованный звонок, результат и комментарий, а также анализ качества работы сотрудника и итоговая оценка, выраженная в %.

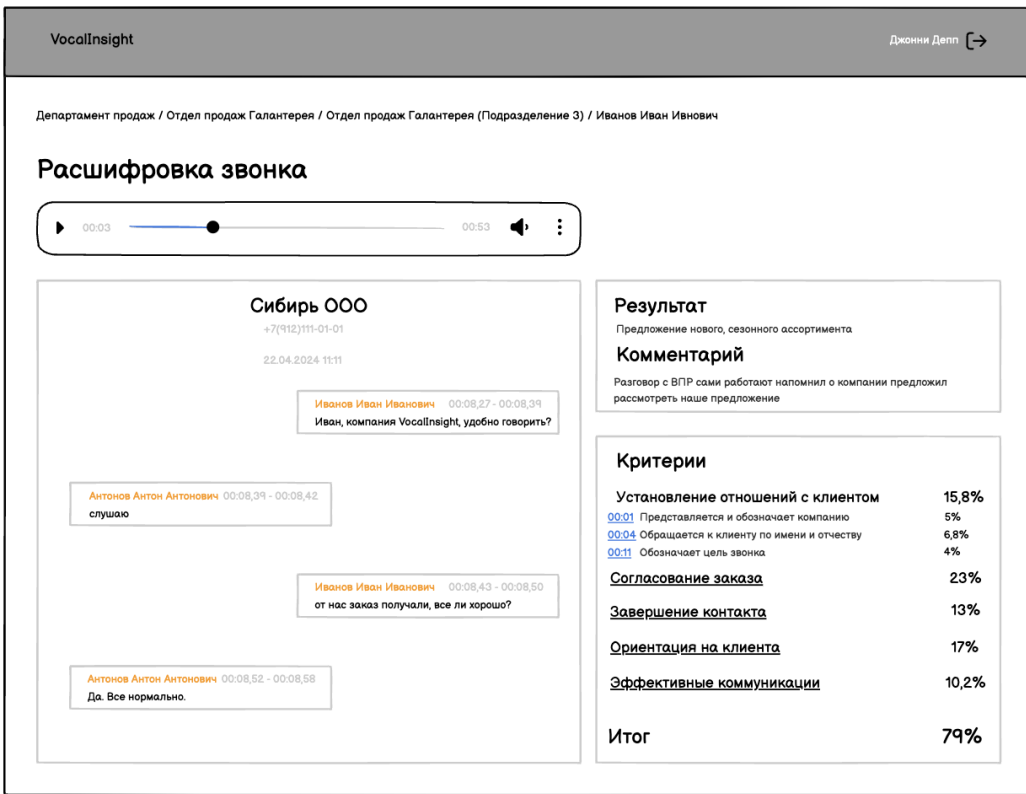


Рисунок 29 - Расшифровка звонка

Помимо этого, есть возможность посмотреть все звонки выбранного менеджера. На рисунке 30 представлена страница с информацией о работе сотрудника Иванов Иван Иванович. Помимо фильтрации поиска, отображается средняя оценка, благодаря которой можно определить эффективность проведенных диалогов.

**Vocallnsight** Джонни Депп

Департамент продаж / Отдел продаж Галантерей / Отдел продаж Галантерей (Подразделение 3) / Иванов Иван Иванович

Средняя оценка **65%**

**Иванов Иван Иванович**

16-07-2023 — 16-08-2023

Поиск

**Чеклист**  
 АКБ ПКС Сервис  
☒ Содержит ☐ Не содержит  
☐ +7(666)666-66-66  
☐ +7(777)777-77-77  
☐ +7(888)888-88-88  
☐ +7(999)999-99-99

**Результат звонка**  
 АКБ ПКС Сервис  
☐ Звонок с потенциала  
☐ Переговоры с ЛПР  
☐ Предложил сотрудничество  
☐ Предложение новых товаров  
☐ Передал в работу

[Показать доступные звонки](#)

Менеджер	Дата и время	Наименование	Номер телефона	Длительность	Результат звонка	Оценка
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	10:43	Предложение нового, сезонного ассортимента	79%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	2:43	Результат отсутствует	90%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	17:43	Звонок с потенциала	81%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	0:43	Переговоры с ЛПР	65%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	51:43	Возражение	34%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	1:43	Готов к заказу	87%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	43:43	Претензии по товару	97%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	10:43	Предложение нового, сезонного ассортимента	79%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	2:43	Результат отсутствует	90%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	17:43	Звонок с потенциала	81%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	0:43	Переговоры с ЛПР	65%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	51:43	Возражение	34%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	1:43	Готов к заказу	87%
Алексеев Алексей Алексеевич	22.04.2024 17:15	Сибирь ООО	+7(912)111-01-01	43:43	Претензии по товару	97%

Рисунок 30 - Список звонков Иванова Ивана

## 3.4 Программная реализация проекта

### 3.4.1 Выбор средств реализации

Для реализации системы распознавания и разметки голосовых записей с применением нейронных сетей был выбран набор технологий, включающий PostgreSQL, Golang, Kubernetes и gRPC для реализации основной части системы, а также whisper.cpp в качестве модели для транскрипции и C++. Каждая из этих технологий имеет свои особенности и преимущества, которые делают их подходящими для решения поставленной задачи.

PostgreSQL — это мощная, объектно-реляционная система управления базами данных с открытым исходным кодом, известная своей надежностью, масштабируемостью и



обширными возможностями[20].

Преимущества:

Надежность и отказоустойчивость: PostgreSQL обеспечивает высокую степень целостности данных и поддерживает ACID-транзакции.

Масштабируемость: поддерживает масштабирование как вертикально (увеличение ресурсов сервера), так и горизонтально (шардинг).

Богатый набор функций: включает поддержку сложных запросов, индексов, полнотекстового поиска, хранимых процедур и расширений.

Активное сообщество: Широкое сообщество разработчиков и пользователей, что обеспечивает постоянное обновление и поддержку.

Система распознавания и разметки голосовых записей требует надежного хранения и управления данными, такими как аудиофайлы, результаты распознавания и метаданные. PostgreSQL подходит для этих целей благодаря своим возможностям и надежности.

Golang, также известный как Go, — это язык программирования с открытым исходным кодом, разработанный в Google. Он известен своей простотой, высокой производительностью и поддержкой параллелизма[18].

Преимущества:

Высокая производительность: компилируется в машинный код, что обеспечивает высокую скорость выполнения программ.

Параллелизм: Встроенная поддержка горутин и каналов упрощает разработку многопоточных приложений.

Простота и эффективность: Язык прост для изучения и использования, что сокращает время разработки.

Богатая стандартная библиотека: включает обширный набор инструментов для работы с сетью, файлами, базами данных и т.д.

Golang подходит для создания микросервисов благодаря своей производительности и поддержке параллелизма. Для системы распознавания и разметки голосовых записей, где требуется быстрая обработка данных и взаимодействие между сервисами, Golang является оптимальным выбором.

Kubernetes — это система оркестрации контейнеров с открытым исходным кодом, разработанная для автоматизации развертывания, масштабирования и управления

контейнеризированными приложениями[21].

Преимущества:

Автоматизация управления: обеспечивает автоматическое развертывание, масштабирование и управление контейнерами.

Масштабируемость: поддерживает горизонтальное масштабирование приложений.

Устойчивость: автоматически восстанавливает упавшие контейнеры и заменяет их новыми.

Управление конфигурацией: Поддержка декларативных конфигураций для управления инфраструктурой и приложениями.

Причины выбора:

В системе с микросервисной архитектурой необходимо эффективно управлять множеством контейнеров, их масштабированием и отказоустойчивостью. Kubernetes предоставляет все необходимые инструменты для этого, обеспечивая стабильную и масштабируемую среду.

gRPC — это современный фреймворк для удаленного вызова процедур (RPC) с открытым исходным кодом, разработанный Google. Он использует Protocol Buffers (protobuf) для сериализации данных и поддерживает работу с различными языками программирования[19].

Преимущества:

Высокая производительность: gRPC обеспечивает быструю и эффективную передачу данных.

Простота определения интерфейсов: Использование protobuf для определения сервисов и сообщений упрощает разработку и поддержание интерфейсов.

Поддержка нескольких языков: gRPC имеет поддержку множества языков программирования, что упрощает интеграцию в разнородных средах.

Стриминг данных: Поддержка потоковой передачи данных (bidirectional streaming), что особенно полезно для реального времени и больших объемов данных.

Для микросервисной архитектуры важна быстрая и эффективная коммуникация между сервисами. gRPC обеспечивает низкую задержку и высокую производительность, что критично для задач распознавания и разметки голосовых записей.

whisper.cpp — это библиотека для высокоэффективного распознавания речи,

написанная на C++, основанная на модели Whisper от OpenAI. Проект `whisper.cpp` позволяет выполнять преобразование аудио в текст, используя преимущества производительности и низкого уровня доступа, предоставляемые языком C++[23].

Выбор PostgreSQL, Golang, Kubernetes и gRPC для реализации системы распознавания и разметки голосовых записей с применением нейронных сетей обусловлен их преимуществами и соответствием требованиям проекта. PostgreSQL обеспечивает надежное и масштабируемое хранилище данных. Golang предоставляет высокую производительность и поддержку параллелизма, что критично для обработки голосовых записей. Kubernetes обеспечивает автоматизированное управление и масштабирование контейнеров, что упрощает развертывание и управление микросервисами. gRPC предоставляет эффективные средства коммуникации между микросервисами, обеспечивая быструю и надежную передачу данных. C++ — это компилируемый язык программирования общего назначения, разработанный как расширение языка C, включающее объектно-ориентированные, обобщенные и функциональные возможности. Все эти технологии вместе создают мощную и гибкую инфраструктуру, способную справиться с задачами, поставленными в рамках диссертации.

### **3.5 Проведение экспериментов и анализ результатов**

В рамках данной работы был проведен эксперимент для подтверждения эффективности внедрения автоматизированной системы для разметки и распознавания голосовых данных, улучшения точности обработки, выявления недостатков в коммуникации и повышения удовлетворенности клиентов. Для эксперимента были использованы данные разговоров из контакт-центра за один месяц, общее количество записей составило 2000. Эти записи были распределены между двумя группами: контрольной и экспериментальной.

Контрольная группа включала 10 менеджеров, которые работали вручную без использования автоматизированной системы. Эти менеджеры вручную размечали и анализировали записи разговоров. Процесс включал прослушивание записей, выделение ключевых фраз и оценку взаимодействий. Среднее время на один разговор составляло около 9 минут. Менеджеры работали по графику 5 дней в неделю, по 8 часов в день. Время, затраченное на разметку и анализ данных в контрольной группе, составило 150 часов

В экспериментальной группе также находилось 10 менеджеров, но они использовали

автоматизированную систему для разметки и анализа записей. Система автоматически выделяла ключевые фразы и предоставляла аналитические данные, что значительно сократило время обработки каждой записи до 3 минут. Общее время, затраченное на разметку и анализ данных в экспериментальной группе, составило 50 часов (см. таблицу 3).

Разница во времени между двумя группами показывает значительное сокращение затрат времени при использовании автоматизированной системы. Это подтверждает целесообразность внедрения таких систем для повышения эффективности анализа записей разговоров. (см. таблицу 3)

Точность распознавания речи также была оценена в обеих группах. В контрольной группе точность определялась путем ручной проверки и корректировки распознанных фраз. Процент ошибок в распознавании составил 15%, что соответствует точности 85%. В экспериментальной группе использовалась модель транскрибации речи, что позволило достичь точности 95%, с процентом ошибок всего 5% (см. таблицу 3).

Таблица 3 – Результаты проведения эксперимента

Группа	Контрольная	Экспериментальная
1 Точность распознавания, %	85	95
2 Выявленные ошибки, шт	50	30
3 Удовлетворенность клиента, балл	3.8	4.5
4 Затраченное время, ч	150	50

Количество выявленных ошибок в работе сотрудников также различалось между группами. В контрольной группе ошибки фиксировались вручную на основе прослушивания записей и анализа разговоров, и было выявлено 50 ошибок за месяц. В экспериментальной группе автоматизированные системы автоматически выявляли и классифицировали ошибки, что позволило снизить их количество до 30. (см. таблицу 3). Удовлетворенность клиентов оценивалась по двум параметрам: непосредственная оценка после звонка оператору и дополнительный опрос, проводимый менеджером через неделю после звонка. Клиенты оценивали взаимодействие по шкале от 1 до 5. В контрольной группе средний балл удовлетворенности составил 3.8, в то время как в экспериментальной группе он был значительно выше — 4.5. (см. таблицу 3).

Помимо приведенного выше, для подтверждения того, что внедрение

автоматизированной системы помогает своевременно выявлять и устранять недостатки в коммуникации, улучшая общую эффективность взаимодействия с клиентами и повышая удовлетворенность клиентов, было проведено дополнительное тестирование после завершения основного тестирования, в котором участвовали обе группы, принимавшие участие в основном эксперименте.

В рамках дополнительного эксперимента были использованы записи разговоров за полгода и эти данные были загружены в автоматизированную систему, которая анализировала коммуникации на предмет выявления различных типов ошибок и проблем, таких как:

- Пропущенные ключевые слова и фразы
- Неправильное использование профессиональной лексики
- Длительные паузы или перебивания
- Негативный тон общения

После анализа система автоматически генерировала отчеты с выявленными недостатками и рекомендациями для их устранения, которые можно выявить исходя из результата анализа системы (см. рисунок 31). Эти отчеты были предоставлены менеджерам для последующего обсуждения с сотрудниками

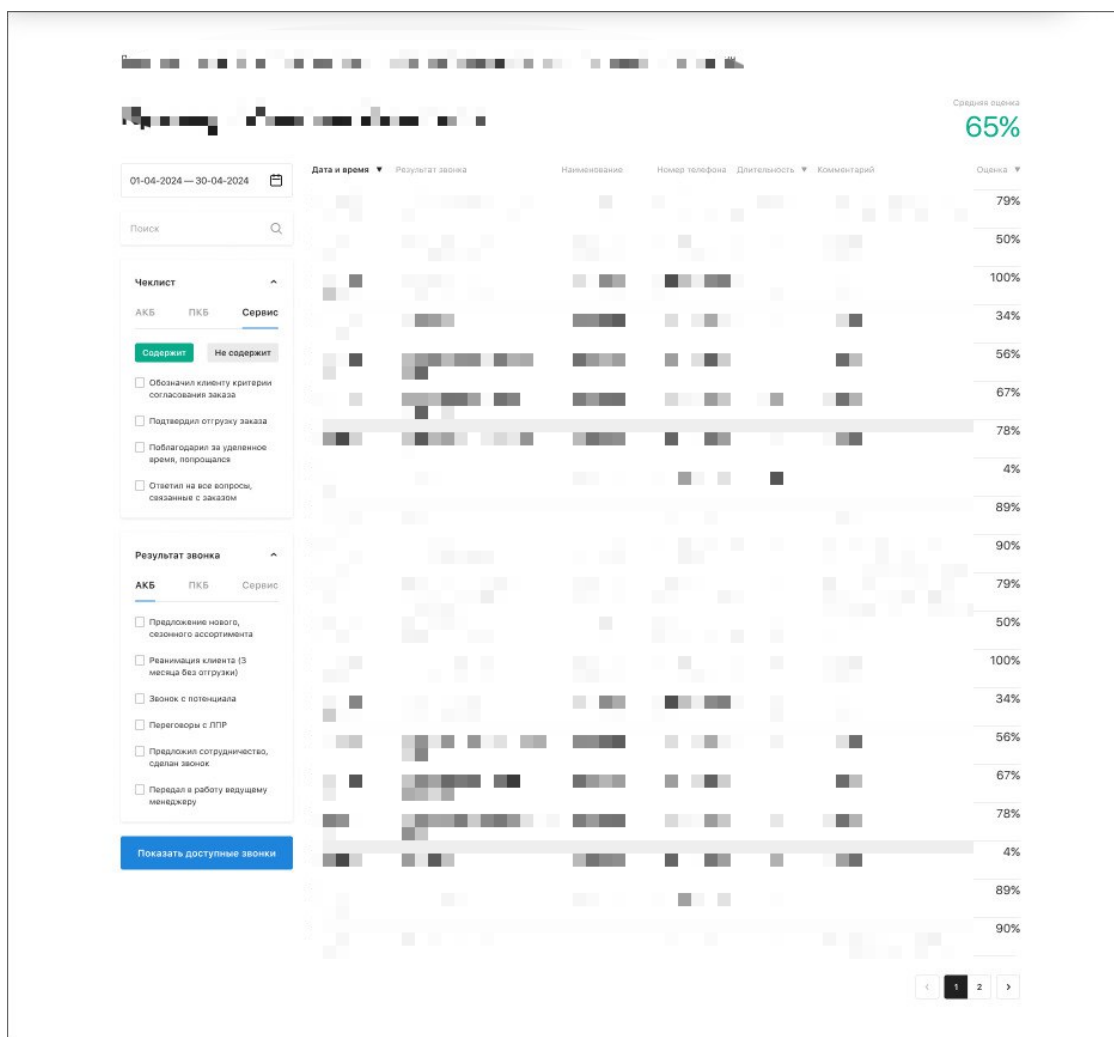


Рисунок 31 – Отчет, построенный системой

Контрольная группа отметила, что данная платформа помогла улучшить объективность оценки работы сотрудников. Также при проведении эксперимента стало известно, что менеджеры недооценивают или переоценивают работу своих подчиненных. Однако, благодаря автоматизированной системе, которая обеспечивает более точную и систематизированную обратную связь, оценка стала более объективной. В экспериментальной группе менеджеры также выделили ряд преимуществ использования платформы. Они отметили простоту поиска звонков от определенных клиентов (см. рисунок 32), возможность отслеживания истории разговоров с клиентом в случае конфликтных ситуаций, а также понятность интерфейса и его интуитивную понятность. Эти факторы значительно повысили эффективность работы менеджеров и улучшили качество обслуживания клиентов.

## Поиск по контрагенту



Рисунок 32 – Просмотр хронологии звонков

Таким образом, результаты эксперимента показали, что использование автоматизированных систем для разметки и распознавания голосовых данных значительно улучшает эффективность анализа и точность обработки, сокращает количество ошибок в работе сотрудников и повышает удовлетворенность клиентов. Эти данные подтверждают целесообразность и пользу дальнейшего внедрения и поддержки разрабатываемой системы.

### ВЫВОДЫ ПО 3 ГЛАВЕ

Выбор микросервисной архитектуры информационной системы оказался удачным решением для выполнения задач, поставленных при разработке данного сервиса. Инфологическое проектирование позволило выделить ключевые объекты системы, которые обеспечивают организацию данных, их хранение и обработку. Проектирование интерфейса упростило дальнейшую разработку дизайна системы при непосредственном создании сервиса. Выбор средств реализации подробно описал используемые в системе платформы, сервисы и языки программирования. Проведение эксперимента завершило разработку сервиса и наглядно продемонстрировало его работоспособность. Результаты исследования демонстрируют, что использование автоматизированных систем для разметки и распознавания голосовых данных значительно повышает эффективность анализа, точность обработки, снижает количество ошибок и повышает удовлетворенность клиентов. Эти данные подтверждают целесообразность дальнейшего внедрения и поддержки разрабатываемой системы, что способствует улучшению качества работы контакт-центров и повышению удовлетворенности клиентов.

## ЗАКЛЮЧЕНИЕ

					Д 09.04.01 2022 22286159 ВКРМ	Лист
		Осинцев А.В.		06.06.24		72
Студ.	Фамилия	Подп.	Дата			



## ЗАКЛЮЧЕНИЕ

В ходе работы был проведен анализ проблемной области, который выявил, что, колл-центры являются важной частью большого количества компаний, но большой объем данных и высокая время затратность необходимая для ручной проверки, требует автоматизации данного процесса. Были рассмотрены и проанализированы существующие решения, которых оказалось крайне мало, что позволило однозначно подтвердить актуальность данной проблемной области.

Была поставлена цель создания системы и определен основной функционал, который должен быть реализован в конечном продукте. Кроме того, были определены основные входные и выходные данные, представленные на диаграмме IDEF0 «Как будет», которые выявляет основные конкурентные преимущества платформы. На основе этой информации были сформулированы требования к проектируемой системе.

В процессе разработки использовались следующие технологии: база данных PostgreSQL, язык программирования Golang, система Kubernetes, а также фреймворк gRPC и библиотека whisper.cpp.

Были проведены эксперименты с группой сотрудников колл-центра, которые показали, что сервис является эффективным инструментом для повышения точности, объективности и скорости оценки работы операторов. Это на практике подтвердило актуальность проделанной работы.

Положения, выносимые на защиту, были доказаны: внедрение системы позволило своевременно выявлять и устранять недостатки в коммуникации, улучшая общую эффективность взаимодействия с клиентами и повышая удовлетворенность клиентов; использование автоматизированных систем обеспечило непрерывный процесс разметки и распознавания голосовых данных, что значительно повысило эффективность анализа записей разговоров; постоянное обновление и оптимизация моделей распознавания речи уменьшили ошибки и улучшили точность обработки данных.

Проведенная работа включала в себя глубокий анализ проблемной области, разработку концептуальной и инфологической моделей, а также реализацию и тестирование системы с использованием современных технологий. Результаты показали, что созданный сервис эффективно решает поставленные задачи, обеспечивая автоматизацию процесса оценки качества работы сотрудников колл-центров. Таким образом, данное исследование подтвердило актуальность и значимость разработки, а также продемонстрировало её практическую ценность.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

					Д 09.04.01 2022 22286159 ВКРМ	Лист
		Осинцев А.В.		06.06.24		74
Студ.	Фамилия	Подп.	Дата			

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Morteza Saberi, Omar Khadeer Hussain, Elizabeth Chang «Past, present and future of contact centers: a literature review». [Электронный ресурс]: <https://www.emerald.com/insight/content/doi/10.1108/BPMJ-02-2015-0018/full/html> [1]
2. Ko de Ruyter, Martin Wetzels, Richard Feinberg «Role stress in call centers » [Электронный ресурс]: <https://www.sciencedirect.com/science/article/abs/pii/S1094996801701788> [2]
3. Zeynep Aksin, Mor Armony, Vijay Mehrotra «The Modern Call Center: A Multi-Disciplinary Perspective on Operations Management Research». [Электронный ресурс]: <https://journals.sagepub.com/doi/10.1111/j.1937-5956.2007.tb00288.x> [3]
4. Richard A. Feinberg, Leigh Hokama, Rajesh Kadam, IkSuk Kim «Operational determinants of caller satisfaction in the banking/financial services call center». [Электронный ресурс]: <https://www.emerald.com/insight/content/doi/10.1108/02652320210432954/full/html> [4]
5. Amy Neustein «Advances in Speech Recognition». [Электронный ресурс]: <https://link.springer.com/book/10.1007/978-1-4419-5951-5> [5]
6. Haipeng Shen «Statistical Analysis of Call-Center Operational Data: Forecasting Call Arrivals, and Analyzing Customer Patience and Agent Service». [Электронный ресурс]: <https://onlinelibrary.wiley.com/doi/10.1002/9780470400531.eorms0827> [6]
7. G. Zweig, O. Siohan, G. Saon, B. Ramabhadran, D. Povey, L. Mangu, B. Kingsbury «AUTOMATED QUALITY MONITORING FOR CALL CENTERS USING SPEECH AND NLP TECHNOLOGIES». [Электронный ресурс]: <https://aclanthology.org/N06-4011.pdf> [7]
8. L.R. Rabiner «Applications of speech recognition in the area of telecommunications». [Электронный ресурс]: <https://ieeexplore.ieee.org/abstract/document/659129> [8]
9. Oleg Borisov, Mohammad Aliannejadi, Fabio Crestani «Keyword Extraction for Improved Document Retrieval in Conversational Search». [Электронный ресурс]: <https://arxiv.org/pdf/2109.05979> [9]
10. Yu Zhang, Mingxiang Tuo, Qingyu Yin, Le Qi, Xuxiang Wang, Ting Liu «Keywords extraction with deep neural network model». [Электронный ресурс]: <https://www.sciencedirect.com/science/article/abs/pii/S092523121931687X> [10]
11. Nazanin Firoozeh, Adeline Nazarenko, Fabrice Alizon, Béatrice Daille «Keyword extraction: Issues and methods». [Электронный ресурс]: <https://www.cambridge.org/core/journals/natural-language-engineering/article/abs/keyword-extraction-issues-and-methods/84BFD5221E2CA86326E5430D03299711> [11]
12. «Twin Speech Analytics». [Электронный ресурс]: <https://twin24.ai/> [12]

13. «Natural Language AI». [Электронный ресурс]: <https://cloud.google.com/natural-language/> [13]
14. «IBM Watson Speech to Text». [Электронный ресурс]: <https://www.ibm.com/products/speech-to-text> [14]
15. «Amazon Transcribe». [Электронный ресурс]: <https://aws.amazon.com/transcribe/> [15]
16. «Amazon Comprehend». [Электронный ресурс]: <https://aws.amazon.com/comprehend/> [16]
17. «Azure Microsoft Speech to text». [Электронный ресурс]: <https://azure.microsoft.com/en-us/products/ai-services/speech-to-text> [17]
18. Go (Golang). Официальный сайт. Режим доступа: <https://go.dev/> [18]
19. gRPC. Официальный сайт. Режим доступа: <https://grpc.io/s://go-app.dev/> [19]
20. PostgreSQL. Официальный сайт. Режим доступа: <https://www.postgresql.org/> [20]
21. Kubernetes. Официальный сайт. Режим доступа: <https://kubernetes.io/> [21]
22. Осинцев А.В., Волкова Е.А. ИССЛЕДОВАНИЕ ТЕХНИЧЕСКИХ АСПЕКТОВ ПОИСКА КЛЮЧЕВЫХ СЛОВ В ТЕКСТОВОМ ПРЕДСТАВЛЕНИИ // Уральская горная школа регионам: сборник докладов международной научно-практической конференции – Екатеринбург: Изд-во УГГУ, 2024. [22]
23. Осинцев А.В., Волкова Е.А. ИССЛЕДОВАНИЕ ВОПРОСА МОДЕЛЕЙ РАСПОЗНАВАНИЯ РЕЧИ // Уральская горная школа регионам: сборник докладов международной научно-практической конференции – Екатеринбург: Изд-во УГГУ, 2024. [23]

## ОТЗЫВ

### на выпускную квалификационную работу магистра магистранта группы ИНФ.м-22

Осинцева Антона Владимировича

«Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей»

Выпускная квалификационная работа магистранта группы ИНФ.м-22 А.В. Осинцева выполнена в соответствии с заданием и содержит все разделы, предусмотренные пунктами задания.

Выпускная квалификационная работа магистра направлена на повышение эффективности анализа записей разговоров для улучшения качества обслуживания клиентов. Цель исследования: повышение эффективности анализа записей разговоров, выявление процента соответствий заданным критериям в совершенных звонках. Научная новизна работы заключается в реализации комплексной системы, которая автоматизирует разметку и распознавание голосовых данных, интегрируется с другими корпоративными системами, также вводит новые методологии оценки работы, которые ранее не применялись. Работа содержит 76 страниц, на которых размещены 32 рисунка, 3 таблицы, 23 литературных источника.

В работе выполнено исследование методологии автоматизации разметки и распознавания голосовых данных, спроектирована база данных, интерфейс, прототип системы анализа голосовых записей, дано техническое обоснование использованию моделей распознавания речи, сформулированы требования к системе.

Спроектировано, разработано и представлено программное обеспечение, представляющее собой прототип автоматизированной информационной системы для анализа голосовых данных. Внедрение предлагаемого решения позволит обеспечить повышение эффективности анализа записей разговоров, улучшение точности обработки голосовых данных, снижение количества ошибок в работе сотрудников и повышение удовлетворенности клиентов. Имеется акт внедрения разработанного решения на предприятии ООО «Софтмаркет».

Выпускная квалификационная работа состоит из аннотации, введения, трех глав, заключения.

#### **Существенных замечаний по работе нет.**

За время работы над выпускной квалификационной работой А.В. Осинцев продемонстрировал умение самостоятельно ставить и решать научные и практические задачи, проводить теоретические исследования и внедрять полученные результаты на практике.

Выполненная работа полностью соответствует требованиям, предъявляемым к выпускной квалификационной работе, магистрант А.В. Осинцев достоин звания дипломированного магистра. Рекомендуемая оценка **«отлично»**.

Научный руководитель,  
к.т.н., доцент кафедры ИНФ



Е.А. Волкова

05 июня 2024 г.



# Отчет о проверке

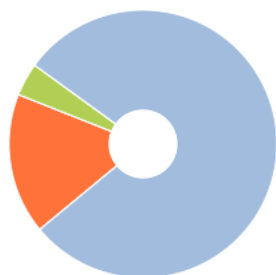
Автор: Осинцев А В

Проверяющий: Oh A

Название документа: Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей

## РЕЗУЛЬТАТЫ ПРОВЕРКИ

Тариф: FULL



Совпадения:  
17,2%



Оригинальность:  
78,58%



Цитирования:  
4,22%



Самоцитирования:  
0%



1

«Совпадения», «Цитирования», «Самоцитирования», «Оригинальность» являются отдельными показателями, отображаются в процентах и в сумме дают 100%, что соответствует проверенному тексту документа.



Есть подозрения на следующие группы маскировки заимствований: Сгенерированный текст на страницах: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20... еще на 29 стр.

- Совпадения** — фрагменты проверяемого текста, полностью или частично сходные с найденными источниками, за исключением фрагментов, которые система отнесла к цитированию или самоцитированию. Показатель «Совпадения» — это доля фрагментов проверяемого текста, отнесенных к совпадениям, в общем объеме текста.
- Самоцитирования** — фрагменты проверяемого текста, совпадающие или почти совпадающие с фрагментом текста источника, автором или соавтором которого является автор проверяемого документа. Показатель «Самоцитирования» — это доля фрагментов текста, отнесенных к самоцитированию, в общем объеме текста.
- Цитирования** — фрагменты проверяемого текста, которые не являются авторскими, но которые система отнесла к корректно оформленным. К цитированиям относятся также шаблонные фразы; библиография; фрагменты текста, найденные модулем поиска «СПС Гарант: нормативно-правовая документация». Показатель «Цитирования» — это доля фрагментов проверяемого текста, отнесенных к цитированию, в общем объеме текста.
- Текстовое пересечение** — фрагмент текста проверяемого документа, совпадающий или почти совпадающий с фрагментом текста источника.
- Источник** — документ, проиндексированный в системе и содержащийся в модуле поиска, по которому проводится проверка.
- Оригинальный текст** — фрагменты проверяемого текста, не обнаруженные ни в одном источнике и не отмеченные ни одним из модулей поиска. Показатель «Оригинальность» — это доля фрагментов проверяемого текста, отнесенных к оригинальному тексту, в общем объеме текста.

Обращаем Ваше внимание, что система находит текстовые совпадения проверяемого документа с проиндексированными в системе источниками. При этом система является вспомогательным инструментом, определение корректности и правомерности совпадений или цитирований, а также авторства текстовых фрагментов проверяемого документа остается в компетенции проверяющего.

## ИНФОРМАЦИЯ О ДОКУМЕНТЕ

Номер документа: 6

Тип документа: Магистерская диссертация

Дата проверки: 08.06.2024 22:40:44

Дата корректировки: 08.06.2024 22:49:27

Количество страниц: 76

Символов в тексте: 112574

Слов в тексте: 13209

Число предложений: 4454

Комментарий: не указано

ПАРАМЕТРЫ ПРОВЕРКИ

Выполнена проверка с учетом редактирования: Да

Выполнено распознавание текста (OCR): Нет

Выполнена проверка с учетом структуры: Нет

**Модули поиска:** Переводные заимствования (KyEn), Переводные заимствования по Интернету (KyRu), Диссертации НББ, Переводные заимствования по Интернету (KkRu), Перефразирования по Интернету (EN), ИПС Адилет, Шаблонные фразы, СПС ГАРАНТ: аналитика, Сводная коллекция ЭБС, СМИ России и СНГ, Медицина, Публикации eLIBRARY (переводы и перефразирования), Переводные заимствования (KkEn), Перефразирования по коллекции IEEE, Публикации eLIBRARY, Переводные заимствования\*, Цитирование, Коллекция НБУ, Патенты СССР, РФ, СНГ, Библиография, Публикации РГБ, СПС ГАРАНТ: нормативно-правовая документация, Переводные заимствования по Интернету (EnRu), Перефразирования по Интернету, Перефразированные заимствования по коллекции Интернет в английском сегменте, Переводные заимствования (RuEn), IEEE, Кольцо вузов, Переводные заимствования по коллекции Интернет в английском сегменте, Перефразированные заимствования по коллекции Интернет в русском сегменте, Переводные заимствования IEEE, Переводные заимствования по коллекции Интернет в русском сегменте, Кольцо вузов (переводы и перефразирования), Переводные заимствования по коллекции Гарант: аналитика, Публикации РГБ (переводы и перефразирования), Перефразирования по СПС ГАРАНТ: аналитика, Интернет Плюс\*

➔ **Модули, недоступные в рамках тарифа:** Интернет Free

ИСТОЧНИКИ

№	Доля в тексте	Источник	Актуален на	Модуль поиска	Комментарий
[01]	8,1%	ИССЛЕДОВАНИЕ И РАЗРАБОТКА АВТОМАТИЗИ...	27 Июн 2018	Кольцо вузов (переводы и перефразирования)	
[02]	5,91%	ПРОЕКТИРОВАНИЕ АВТОМАТИЗИРОВАННОЙ ...	18 Июн 2018	Кольцо вузов (переводы и перефразирования)	
[03]	4,77%	ИССЛЕДОВАНИЕ И РАЗРАБОТКА АВТОМАТИЗИ...	27 Июн 2018	Кольцо вузов	
[04]	3,35%	не указано	13 Янв 2022	Библиография	
[05]	2,68%	Проектирование и разработка автоматизиров...	28 Июн 2018	Кольцо вузов (переводы и перефразирования)	
[06]	2,5%	Исследовательская работа Соловьева.docx	18 Фев 2023	Кольцо вузов (переводы и перефразирования)	
[07]	2,45%	ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ГРАФИЧЕСК...	18 Июн 2018	Кольцо вузов (переводы и перефразирования)	
[08]	2,22%	<a href="https://www.ursmu.ru/userfiles/media/default/s...">https://www.ursmu.ru/userfiles/media/default/s...</a> <a href="https://ursmu.ru">https://ursmu.ru</a>	22 Дек 2023	Перефразированные заимствования по коллекции Интернет в русском сегменте	
[09]	1,85%	ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ИНФОРМА...	15 Июн 2018	Кольцо вузов	
[10]	1,38%	ПРОЕКТИРОВАНИЕ АВТОМАТИЗИРОВАННОЙ ...	18 Июн 2018	Кольцо вузов	
[11]	1,31%	Проектирование и разработка автоматизиров...	18 Июн 2018	Кольцо вузов	
[12]	1,02%	Сайфулин Сергей Олегович SSayfulin_Diploma...	05 Июн 2018	Кольцо вузов	
[13]	0,97%	Разработка информационной подсистемы по ...	16 Июн 2023	Кольцо вузов	
[14]	0,95%	АНАЛИЗ ИСПОЛЬЗОВАНИЯ МИКРОСЕРВИСНО...	22 Сен 2023	Публикации eLIBRARY (переводы и перефразирования)	
[15]	0,87%	не указано	13 Янв 2022	Шаблонные фразы	
[16]	0,81%	Росляков, Александр Станиславович Оптимиз... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	01 Янв 2012	Публикации РГБ	
[17]	0,81%	Механизмы обеспечения эффективности и ре... <a href="http://elibrary.ru">http://elibrary.ru</a>	09 Июл 2020	Публикации eLIBRARY (переводы и перефразирования)	
[18]	0,81%	<a href="https://www.ursmu.ru/userfiles/media/default/s...">https://www.ursmu.ru/userfiles/media/default/s...</a> <a href="https://ursmu.ru">https://ursmu.ru</a>	22 Дек 2023	Интернет Плюс*	
[19]	0,72%	не указано	13 Янв 2022	Цитирование	Источник исключен. Причина: Маленький процент пересечения.

[20]	0,67%	%D0%A7%D0%B0%D1%81%D1%82%D1%8C%20... <a href="https://kosygin-rgu.ru">https://kosygin-rgu.ru</a>	05 Дек 2023	Перефразированные заимствования по коллекции Интернет в русском сегменте	
[21]	0,59%	Диссертация на тему «Формирование критиче... <a href="https://dissercat.com">https://dissercat.com</a>	05 Фев 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[22]	0,57%	Информационно-управляющая система для п... <a href="http://elibrary.ru">http://elibrary.ru</a>	01 Янв 2023	Публикации eLIBRARY (переводы и перефразирования)	
[23]	0,53%	Разработка приложения для распознавания р...	18 Окт 2023	Публикации eLIBRARY (переводы и перефразирования)	Источник исключен. Причина: Маленький процент пересечения.
[24]	0,51%	Беспалько, Виталий Алексеевич Стратегическ... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	12 Янв 2021	Публикации РГБ	Источник исключен. Причина: Маленький процент пересечения.
[25]	0,51%	Беспалько, Виталий Алексеевич Стратегическ... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	25 Окт 2019	Публикации РГБ	Источник исключен. Причина: Маленький процент пересечения.
[26]	0,45%	РАЗРАБОТКА ТЕХНИЧЕСКОГО ЗАДАНИЯ НА ВН... <a href="https://scienceforum.ru">https://scienceforum.ru</a>	08 Мар 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[27]	0,42%	<a href="https://landrailpip.ru/wp-content/uploads/2024/...">https://landrailpip.ru/wp-content/uploads/2024/...</a> <a href="https://landrailpip.ru">https://landrailpip.ru</a>	07 Апр 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[28]	0,41%	АвдеевИИ_ББИ-19-3	13 Июн 2023	Кольцо вузов	Источник исключен. Причина: Маленький процент пересечения.
[29]	0,38%	Диплом Кривелевич С.Б..docx	24 Мая 2023	Кольцо вузов	Источник исключен. Причина: Маленький процент пересечения.
[30]	0,33%	Алексанов А.К., Доронин А.М., Демчев И.А. и д... <a href="http://ivo.garant.ru">http://ivo.garant.ru</a>	19 Окт 2013	Перефразирования по СПС ГАРАНТ: аналитика	
[31]	0,3%	zaharkov_a_a_tehniko-ekonomicheskoe-obosnov...	17 Мая 2022	Кольцо вузов	Источник исключен. Причина: Маленький процент пересечения.
[32]	0,3%	Микросервисная архитектура как инновацион...	17 Янв 2024	Публикации eLIBRARY (переводы и перефразирования)	
[33]	0,27%	ВНЕДРЕНИЕ МОДУЛЯ ВЗАИМОДЕЙСТВИЯ С БА... <a href="https://scienceforum.ru">https://scienceforum.ru</a>	27 Мая 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[34]	0,27%	Характеристика объектов автоматизации - Тех... <a href="https://topuch.ru">https://topuch.ru</a>	04 Мар 2022	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[35]	0,25%	Демидов, Роман Алексеевич Выявление угроз ... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	01 Янв 2018	Публикации РГБ (переводы и перефразирования)	
[36]	0,23%	Наугольнова, Ирина Александровна Процессн... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	01 Янв 2023	Публикации РГБ	Источник исключен. Причина: Маленький процент пересечения.
[37]	0,23%	<a href="http://mhs-journal.ru/wp-content/uploads/2019/...">http://mhs-journal.ru/wp-content/uploads/2019/...</a> <a href="http://mhs-journal.ru">http://mhs-journal.ru</a>	14 Фев 2022	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[38]	0,23%	Микросервисная архитектура: основные конц... <a href="https://it-vacancies.ru">https://it-vacancies.ru</a>	03 Янв 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[39]	0,22%	Информационно-технологическое обеспечен... <a href="http://ivo.garant.ru">http://ivo.garant.ru</a>	25 Июн 2022	Перефразирования по СПС ГАРАНТ: аналитика	
[40]	0,22%	Искусственный интеллект и обработка времен... <a href="https://nauchniestati.ru">https://nauchniestati.ru</a>	17 Апр 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[41]	0,21%	<a href="https://ieomsociety.org/proceedings/2022istanb...">https://ieomsociety.org/proceedings/2022istanb...</a> <a href="https://ieomsociety.org">https://ieomsociety.org</a>	08 Июн 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[42]	0,21%	Predicting Individual-Level Call Arrival from Onlin... <a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>	24 Янв 2019	IEEE	Источник исключен. Причина: Маленький процент пересечения.
[43]	0,21%	<a href="http://se.sportedu.ru/sites/se.sportedu.ru/files/s...">http://se.sportedu.ru/sites/se.sportedu.ru/files/s...</a> <a href="http://se.sportedu.ru">http://se.sportedu.ru</a>	10 Апр 2023	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.
[44]	0,2%	On the modeling and forecasting of call center ar... <a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>	21 Фев 2013	IEEE	Источник исключен. Причина: Маленький процент пересечения.
[45]	0,19%	Операционные системы. Концепции построен... <a href="http://studentlibrary.ru">http://studentlibrary.ru</a>	20 Дек 2016	Медицина	Источник исключен. Причина: Маленький процент пересечения.
[46]	0,18%	С. Л. Шнырев Базы данных учебное пособие д... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	31 Мар 2014	Публикации РГБ (переводы и перефразирования)	Источник исключен. Причина: Маленький процент пересечения.
[47]	0,18%	vk_r_final_vers_1.docx	24 Июн 2023	Кольцо вузов	Источник исключен. Причина: Маленький процент пересечения.
[48]	0,18%	Карпов, Алексей Анатольевич диссертация ... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	раньше 2011	Публикации РГБ (переводы и перефразирования)	Источник исключен. Причина: Маленький процент пересечения.
[49]	0,18%	Customised Data Dashboard for Contact Centres... <a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>	30 Дек 2018	IEEE	Источник исключен. Причина: Маленький процент пересечения.
[50]	0,17%	сборник научных статей х международной на... <a href="https://bigdataminsk.bsuir.by">https://bigdataminsk.bsuir.by</a>	16 Апр 2024	Переводные заимствования по	Источник исключен. Причина: Маленький процент пересечения.



					коллекции Интернет в русском сегменте	
[51]	0,17%	СОВЕРШЕНСТВОВАНИЕ МЕТОДИКИ УЧЁТА И Н... <a href="http://elibrary.ru">http://elibrary.ru</a>	01 Янв 2018	Публикации eLIBRARY	Источник исключен. Причина: Маленький процент пересечения.	
[52]	0,17%	ISO_IEC_17799_2000_rus - Стр 9 <a href="http://studfiles.ru">http://studfiles.ru</a>	08 Янв 2017	Перефразирования по Интернету	Источник исключен. Причина: Маленький процент пересечения.	
[53]	0,16%	komashinskij v.i., smirnov d.a. nejronnye seti i ih ... <a href="http://inethub.olvi.net.ua">http://inethub.olvi.net.ua</a>	22 Апр 2014	Переводные заимствования по коллекции Интернет в русском сегменте	Источник исключен. Причина: Маленький процент пересечения.	
[54]	0,15%	Simulation Assessment of Data-Driven Channel A... <a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>	07 Апр 2023	IEEE	Источник исключен. Причина: Маленький процент пересечения.	
[55]	0,15%	Газуль, Станислав Михайлович Формировани... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	22 Фев 2019	Публикации РГБ (переводы и перефразирования)	Источник исключен. Причина: Маленький процент пересечения.	
[56]	0,14%	Об утверждении Правил приема, хранения св... <a href="http://adilet.zan.kz">http://adilet.zan.kz</a>	21 Янв 2016	ИПС Адилет	Источник исключен. Причина: Маленький процент пересечения.	
[57]	0,14%	Likability Estimation of Call-center Agents by Sup... <a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>	30 Июн 2023	Перефразирования по коллекции IEEE	Источник исключен. Причина: Маленький процент пересечения.	
[58]	0,14%	Speech Sentiment and Customer Satisfaction Esti... <a href="http://arxiv.org">http://arxiv.org</a>	21 Янв 2023	Перефразированные заимствования по коллекции Интернет в английском сегменте	Источник исключен. Причина: Маленький процент пересечения.	
[59]	0,14%	Likability Estimation of Call-center Agents by Sup... <a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>	30 Июн 2023	IEEE	Источник исключен. Причина: Маленький процент пересечения.	
[60]	0,14%	[Lecture Notes in Computer Science] Computatio... <a href="https://doi.org">https://doi.org</a>	16 Окт 2019	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[61]	0,14%	Mining conversational text for procedures with a... <a href="https://doi.org">https://doi.org</a>	08 Июн 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[62]	0,14%	Electronics   Free Full-Text   Voice-Controlled Int... <a href="https://mdpi.com">https://mdpi.com</a>	08 Июн 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[63]	0,13%	Speech Sentiment and Customer Satisfaction Esti... <a href="http://arxiv.org">http://arxiv.org</a>	21 Янв 2023	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[64]	0,12%	<a href="https://t4.kai.ru/images/digest_2021/5.pdf">https://t4.kai.ru/images/digest_2021/5.pdf</a> <a href="https://t4.kai.ru">https://t4.kai.ru</a>	08 Сен 2022	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[65]	0,12%	Саркисян, Анна Ашотовна Цифровизация суде... <a href="http://dlib.rsl.ru">http://dlib.rsl.ru</a>	01 Янв 2023	Публикации РГБ	Источник исключен. Причина: Маленький процент пересечения.	
[66]	0,09%	<a href="https://bigdataminsk.bsuir.by/files%202024/2024...">https://bigdataminsk.bsuir.by/files%202024/2024...</a> <a href="https://bigdataminsk.bsuir.by">https://bigdataminsk.bsuir.by</a>	20 Апр 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[67]	0,09%	ПОСТРОЕНИЕ И АНАЛИЗ МОДЕЛИ КОНТАКТ-Ц... <a href="https://vaael.ru">https://vaael.ru</a>	31 Мая 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[68]	0,09%	Salesperson Self-Regulated Learning and Online ... <a href="https://ncbi.nlm.nih.gov">https://ncbi.nlm.nih.gov</a>	08 Июн 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[69]	0,09%	Подходы к классификации текста с использов... <a href="https://viola37.ru">https://viola37.ru</a>	20 Дек 2023	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[70]	0,08%	Автоматическое распознавание речи и аудио... <a href="https://airobotic.ru">https://airobotic.ru</a>	08 Мая 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[71]	0,07%	Микросервисы глубинного изучения: эффекти... <a href="https://nauchniestati.ru">https://nauchniestati.ru</a>	05 Мая 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[72]	0,07%	<a href="https://www.ursmu.ru/userfiles/media/default/o...">https://www.ursmu.ru/userfiles/media/default/o...</a> <a href="https://ursmu.ru">https://ursmu.ru</a>	01 Апр 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[73]	0,07%	Об эффективности Call-центров, применяемых... <a href="https://cyberleninka.ru">https://cyberleninka.ru</a>	08 Июн 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[74]	0,06%	Потолок в продажах: причины и способы пре... <a href="https://amsales.ru">https://amsales.ru</a>	31 Мая 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[75]	0,05%	KPI логиста по транспорту: основные показате... <a href="https://logistics.by">https://logistics.by</a>	03 Июн 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[76]	0,05%	<a href="https://elar.urfu.ru/bitstream/10995/129220/1/m...">https://elar.urfu.ru/bitstream/10995/129220/1/m...</a> <a href="https://elar.urfu.ru">https://elar.urfu.ru</a>	29 Апр 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[77]	0,04%	<a href="https://xn--80aa3afkgvdf5he.xn--p1ai/RNSM-47_...">https://xn--80aa3afkgvdf5he.xn--p1ai/RNSM-47_...</a> <a href="https://xn--80aa3afkgvdf5he.xn--p1ai">https://xn--80aa3afkgvdf5he.xn--p1ai</a>	29 Мая 2024	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	
[78]	0,04%	<a href="https://books.ifmo.ru/file/pdf/3257.pdf">https://books.ifmo.ru/file/pdf/3257.pdf</a> <a href="https://books.ifmo.ru">https://books.ifmo.ru</a>	28 Окт 2023	Интернет Плюс*	Источник исключен. Причина: Маленький процент пересечения.	

# 1. Результаты проверки по базам данных ВКР-ВУЗ:

- Предложений после анализа документа: **870**
- Символов после анализа документа: **106525**
- Оригинальность текста документа: **99.57%**
- Корректных заимствований (цитирований): **0%**
- Фактических некорректных заимствований: **0.43%**
  - Из которых некорректных цитирований: **0%**

Источник	Ссылка на источник	Коллекция/ модуль поиска	Доля в тексте	Доля в отчете
Работа «Мотивация в системе управления предприятием: роль, состояние, анализ и направление совершенствования» студента «Ахмедзаде Эльхан Магомед Оглы» организации «Филиал Российского экономического университета им. Г.В.Плеханова в г. Пятигорск»	()	Интернет	24.19%	0.11%
Тема: Информационные технологии в телекоммуникациях	<a href="http://bibliofond.ru/view.aspx?id=722127">http://bibliofond.ru/view.aspx?id=722127</a> ( <a href="http://bibliofond.ru/view.aspx?id=722127">http://bibliofond.ru/view.aspx?id=722127</a> )	Интернет	12.74%	0.06%
Тема: Методика программирования в LabView	<a href="http://bibliofond.ru/view.aspx?id=607501">http://bibliofond.ru/view.aspx?id=607501</a> ( <a href="http://bibliofond.ru/view.aspx?id=607501">http://bibliofond.ru/view.aspx?id=607501</a> )	Интернет	11.45%	0.05%
Массаж как средство регуляции эмоционального состояния студенток, связанного со сдачей государственного экзамена	<a href="http://www.km.ru/referats/332497-massazh-kak-sredstvo-regulyatsii-emotsionalnogo-sostoyaniya-studentok-svyazannogo-so-sdachei-">http://www.km.ru/referats/332497-massazh-kak-sredstvo-regulyatsii-emotsionalnogo-sostoyaniya-studentok-svyazannogo-so-sdachei-</a> ( <a href="http://www.km.ru/referats/332497-massazh-kak-sredstvo-regulyatsii-emotsionalnogo-sostoyaniya-studentok-svyazannogo-so-sdachei-">http://www.km.ru/referats/332497-massazh-kak-sredstvo-regulyatsii-emotsionalnogo-sostoyaniya-studentok-svyazannogo-so-sdachei-</a> )	Интернет	10.15%	0.04%
Тема: Реинжиниринг в страховой организации	<a href="http://bibliofond.ru/view.aspx?id=562636">http://bibliofond.ru/view.aspx?id=562636</a> ( <a href="http://bibliofond.ru/view.aspx?id=562636">http://bibliofond.ru/view.aspx?id=562636</a> )	Интернет	9.72%	0.04%
Тема: Слуховой анализатор	<a href="http://bibliofond.ru/view.aspx?id=720973">http://bibliofond.ru/view.aspx?id=720973</a> ( <a href="http://bibliofond.ru/view.aspx?id=720973">http://bibliofond.ru/view.aspx?id=720973</a> )	Интернет	8.42%	0.04%
Информационная защита системы 1С:Бухгалтерия 8.0	<a href="http://www.ronl.ru/referaty/informatika-programmirovanie/100595/">http://www.ronl.ru/referaty/informatika-programmirovanie/100595/</a> ( <a href="http://www.ronl.ru/referaty/informatika-programmirovanie/100595/">http://www.ronl.ru/referaty/informatika-programmirovanie/100595/</a> )	Интернет	7.99%	0.03%

Источник	Ссылка на источник	Коллекция/ модуль поиска	Доля в тексте	Доля в отчете
Тема: Экспериментальные исследования облачных вычислений как основы для создания информационного пространства научных коммуникаций	<a href="http://bibliofond.ru/view.aspx?id=725442">http://bibliofond.ru/view.aspx?id=725442</a> ( <a href="http://bibliofond.ru/view.aspx?id=725442">http://bibliofond.ru/view.aspx?id=725442</a> )	Интернет	7.78%	0.03%
Тема: Совершенствование коммерческой работы по продажам цветов	<a href="http://bibliofond.ru/view.aspx?id=817486">http://bibliofond.ru/view.aspx?id=817486</a> ( <a href="http://bibliofond.ru/view.aspx?id=817486">http://bibliofond.ru/view.aspx?id=817486</a> )	Интернет	7.56%	0.03%

## РЕЦЕНЗИЯ

на магистерскую диссертацию студента

А. В. Осинцева

«Исследование и разработка системы распознавания и разметки голосовых записей с применением нейронных сетей»

Выпускная квалификационная работа студента Осинцева А. В. выполнена в соответствии с заданием и содержит все основные разделы, предусмотренные пунктами задания.

В работе выполнено исследование методологии автоматизации разметки и распознавания голосовых данных.

### Вопросы по работе:

1. В работе не указаны оценки времени при распознавании фрагментов звонка.
2. Отсутствуют диаграммы развертывания сервисов в Kubernetes.
3. При работе ИС возможны ситуации что один или несколько микросервисов становятся недоступны. Как решается проблема надежности передачи данных между сервисами? Какие есть альтернативные решения?
4. В разделе 2.1.2 Дано описание сущности «Слово» у которого указаны атрибуты начало и конец. Прошу объяснить почему одно и то же слово в различных треках и в различных позициях треков должны быть различными сущностями?

Выполненная работа соответствует требованиям, предъявляемым к магистерской диссертации, студент Осинцев А. В. достоин звания магистр. Рекомендована оценка хорошо.

Ведущий инженер



Д.С. Рыжков

14 июня 2024 г.