



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н. Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ «Информатика, искусственный интеллект и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

К КУРСОВОЙ РАБОТЕ

НА ТЕМУ:

«Мобильное приложение для изучения японского языка»

Студент ИУ7-61Б
(Группа)

(Подпись, дата)

Корниенко К. Ю.
(И. О. Фамилия)

Руководитель курсовой работы

(Подпись, дата)

Шибанова Д. А.
(И. О. Фамилия)

2023 г.

СОДЕРЖАНИЕ

ОПРЕДЕЛЕНИЯ	3
ВВЕДЕНИЕ	4
1 Аналитический раздел	5
1.1 Использование цифровых технологий в изучении языков	5
1.2 Базы данных, системы управления базами данных	5
1.3 Выбор системы управления базами данных	6
1.4 Обзор существующего программного обеспечения для упроще- ния изучения японского языка	7
1.5 Проектирование базы данных	9
1.5.1 Диаграмма базы данных в нотации Чена	9
1.5.2 Пользовательские роли проектируемого приложения . . .	9
2 Конструкторский раздел	12
2.1 Концептуальная модель системы	12
2.2 Диаграмма прецедентов	13
2.3 Схема базы данных	15
3 Технологический раздел	16
4 Исследовательский раздел	17
ЗАКЛЮЧЕНИЕ	18
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	20
ПРИЛОЖЕНИЕ А	21

ОПРЕДЕЛЕНИЯ

В настоящей расчетно-пояснительной записке применяют следующие термины с соответствующими определениями.

Сетевое обучение — парадигма учебной деятельности, основой которой является открытость образовательных ресурсов, а также повсеместное сотрудничество участников образовательного процесса.

Кана — японские слоговые алфавиты: хирагана и катакана.

Хирагана — японский слоговой алфавит, который используется для записи слов, не имеющих кандзи.

Катакана — японский слоговой алфавит, который используется для записи иностранных слов и выделения важных слов в тексте.

Кандзи — китайские иероглифы, которые используются в японской письменности. Кандзи были заимствованы из китайского языка и адаптированы для использования в японском языке

ВВЕДЕНИЕ

В последние годы распространена идея использования цифровых технологий для помощи студентам в преодолении трудностей, связанных с изучением японского языка, таких как сложная система письма, грамматика и синтаксис, а также необходимость запоминания большого количества новых слов и символов [1]. Также рассматривается применение сетевого обучения, как парадигмы в дистанционном обучении для улучшения взаимодействия между учащимися путем создания виртуальных общностей учащихся [2].

Одной из наиболее важных проблем, с которым сталкивается человек при изучении японского языка — сложная система письма которая включает в себя два алфавита и иероглифику. Оптическое распознавание символов может облегчить изучение японского языка, преобразуя изображения, содержащие японский язык, в текстовые документы, для более легкой обработки, перевода и изучения иероглифов [3].

Цель работы — разработать мобильное приложение для упрощения изучения японского языка.

Для достижения поставленной цели, необходимо выполнить следующие задачи:

- провести анализ предметной области;
- описать взаимодействие компонентов системы и спроектировать базу данных;
- описать интерфейс доступа к базе данных и произвести тестирование функционала;
- провести сравнительный анализ работы со строками в хранимых процедурах и на клиенте.

1 Аналитический раздел

В данном разделе представлен обзор существующего ПО и подходов к упрощению изучения японского языка с помощью систем оптического распознавания текстов. Также введены основные сведения о построении баз данных для текстов на японском языке и словарей иероглифики.

1.1 Использование цифровых технологий в изучении языков

Использование цифровых технологий в изучении японского языка является актуальным и эффективным подходом для изучения японского языка и повышения мотивации в процессе изучения [4]. Проблемы, с которыми сталкиваются изучающие японский язык, включают необходимость запоминания большого количества новых слов и иероглифов, а также сложности в понимании контекста и культурных отличий.

Во время круглого стола по технологиям для устойчивого развития, организованного ЮНЕСКО в Париже в 2013 г., проведенные исследования и изученные отчеты, выделяющие значительное преимущество применения информационных технологий в изучении языков и в преподавании, позволили сделать вывод о главенствующей роли цифровизации в формировании коллективных знаний [5].

1.2 Базы данных, системы управления базами данных

База данных (БД) — это компьютеризированная система, основное назначение которой — хранить информацию, предоставляя пользователям средства ее извлечения и модификации [6, с. 46]. Базы данных используются для различных целей, таких как управление бизнесом, научные исследования и медицинские записи.

Работа напрямую с базой данных может привести к нарушению целостности данных. Код для работы с базой данных сложен и неудобен, что может привести к ошибкам и проблемам с безопасностью. Избежать проблемы при работе с данными позволяет система управления базами данных.

Система управления базами данных (СУБД) - это программное обеспе-

чение, которое позволяет пользователям создавать, управлять и обрабатывать данные в БД [7, с. 10—12]. СУБД предоставляет интерфейс для работы с данными, а также обеспечивает безопасность и целостность данных.

СУБД и БД тесно связаны, поскольку системы управления базами данных обеспечивают доступ к данным, хранящимся в БД, и позволяют пользователям выполнять операции с данными.

1.3 Выбор системы управления базами данных

Для разработки мобильного приложения, которое будет преобразовывать фотографии текстов на японском языке в текстовые документы, может быть использована как реляционная система управления базами данных, так и нереляционная. При разработке приложения будет использована реляционная система управления базами данных, так как она позволяет оформить ролевую модель и хранить структурированные данные [8].

Реляционные системы управления базами данных характеризуются использованием реляционной модели управления, отличающуюся табличной формой представления данных, а также применением формальной математики и реляционных вычислений для обрабатываемых данных [9].

Данные, в реляционных моделях, представляют собой двумерный массив и характеризуются следующими особенностями:

- любая составляющая таблицы является одной составляющей данных;
- любой столбец имеет свое уникальное имя;
- отсутствие одинаковых строк в таблице;
- все составляющие в столбцах имеют однородный тип;
- строки и столбцы имеют произвольный порядок [10].

Основные современные СУБД основаны на реляционной модели данных, в таблице 1.1 представлен сравнительный анализ некоторых из них.

Таблица 1.1 – Сравнительный анализ реляционных СУБД

СУБД	Лицензия	Масштабируемость	Скорость выполнения запросов	Управление транзакциями
MSSQL [11]	проприетарная	вертикальная, комплексная горизонтальная	средняя	пессимистический
MariaDB [12]	GNU GPL	вертикальная горизонтальная	средняя	частично-оптимистический
PostgreSQL [13]	открытый исходный код	вертикальная горизонтальная	высокая	частично-оптимистический
Oracle [14]	проприетарная	вертикальная горизонтальная	высокая	оптимистическое
IBM DB2 [15]	проприетарная	вертикальная горизонтальная	средняя	пессимистический
SQLite [16]	MIT	вертикальная	низкая	частично-пессимистический

PostgreSQL обладает высокой производительностью и безопасностью, а также является бесплатным программным продуктом с открытым исходным кодом [10], что делает его доступным для использования при разработке мобильного приложения. Кроме того, PostgreSQL имеет хорошую поддержку хранимых процедур и триггеров, имеет хорошую поддержку для языков SQL и Unicode, что важно для работы с японским языком. Например, PostgreSQL поддерживает полнотекстовый поиск на японском языке и имеет встроенную поддержку для японских иероглифов [13], что делает его хорошим выбором для разрабатываемого приложения.

1.4 Обзор существующего программного обеспечения для упрощения изучения японского языка

При изучении иностранных языков часто используются мобильные приложения, изучение языков с поддержкой мобильных устройств позволяет обучающемуся получить доступ к знаниям о грамматике и лексике иностран-

ного языка, не накладывая ограничений на место и время изучения [17]. Особенно информационные технологии важны при изучении японского языка, который отличается двумя азбуками и иероглификой, ведь система письменности является важным аспектом при обучении любому иностранному языку.

В таблице 1.2 представлен сравнительный анализ четырех приложений для изучения японского языка.

Таблица 1.2 – Обзор приложений для упрощения изучения японского языка

ПО	Duolingo [18]	Memrise [19]	Lingodeer [20]	WaniKani [21]
Цена	Бесплатно с платными функциями	Бесплатно с платными функциями	Бесплатно с платными функциями	платно
Типы упражнений	перевод, аудирование, грамматика	перевод, аудирование, грамматика	перевод, аудирование, грамматика	кана, кандзи, слова, грамматика
Платформа	веб, мобильное приложение	веб, мобильное приложение	веб, мобильное приложение	веб, мобильное приложение

Исходя из приведенных в таблице данных, можно сделать вывод о том, что приложения не уделяют достаточно внимания чтению текстов на японском языке. В то же время именно чтение текстов на японском языке представляет наибольшую сложность в процессе обучения [22]. Для упрощения изучения японского языка в разрабатываемом мобильном приложении будет сделан акцент на снижение сложности чтения и анализа иероглифических текстов путем синтеза текстовых документов из изображений отрывков из книг, журналов и других источников текстов на японском языке.

1.5 Проектирование базы данных

1.5.1 Диаграмма базы данных в нотации Чена

Для проектируемой базы данных создана ER-диаграмма в нотации Чена, представленная на рисунке 1.1.

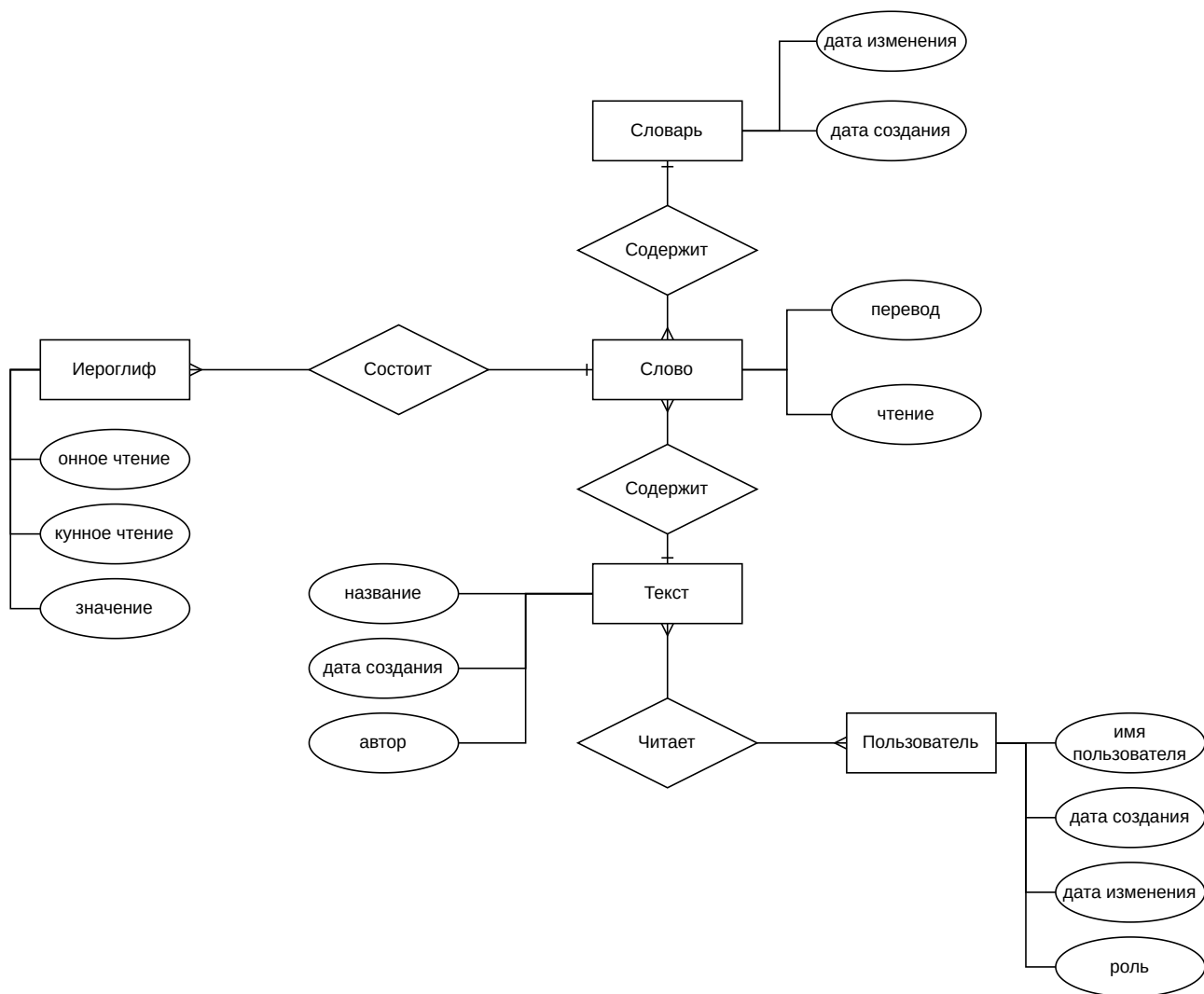


Рисунок 1.1 – Диаграмма сущность-связь в нотации Чена

1.5.2 Пользовательские роли проектируемого приложения

В разрабатываемом приложении выделяются три роли: администратор, модератор и пользователь.

1. Администратор — имеет полный доступ к приложению и базе данных. Он может создавать и удалять учетные записи пользователей и модераторов,

а также имеет доступ к полной информации о пользователях и их активности в приложении.

2. Модератор — имеет ограниченный доступ к приложению и базе данных. Он может создавать тексты и выдавать доступ к ним для пользователей. Модератор не имеет доступа к полной информации о пользователях и их активности в приложении.
3. Пользователь — имеет доступ к основным функциям приложения, таким как изучение японского языка, создание своего профиля. Пользователь не имеет полного доступа к базе данных.

Описание пользовательских сценариев представлено на рисунке 1.2.

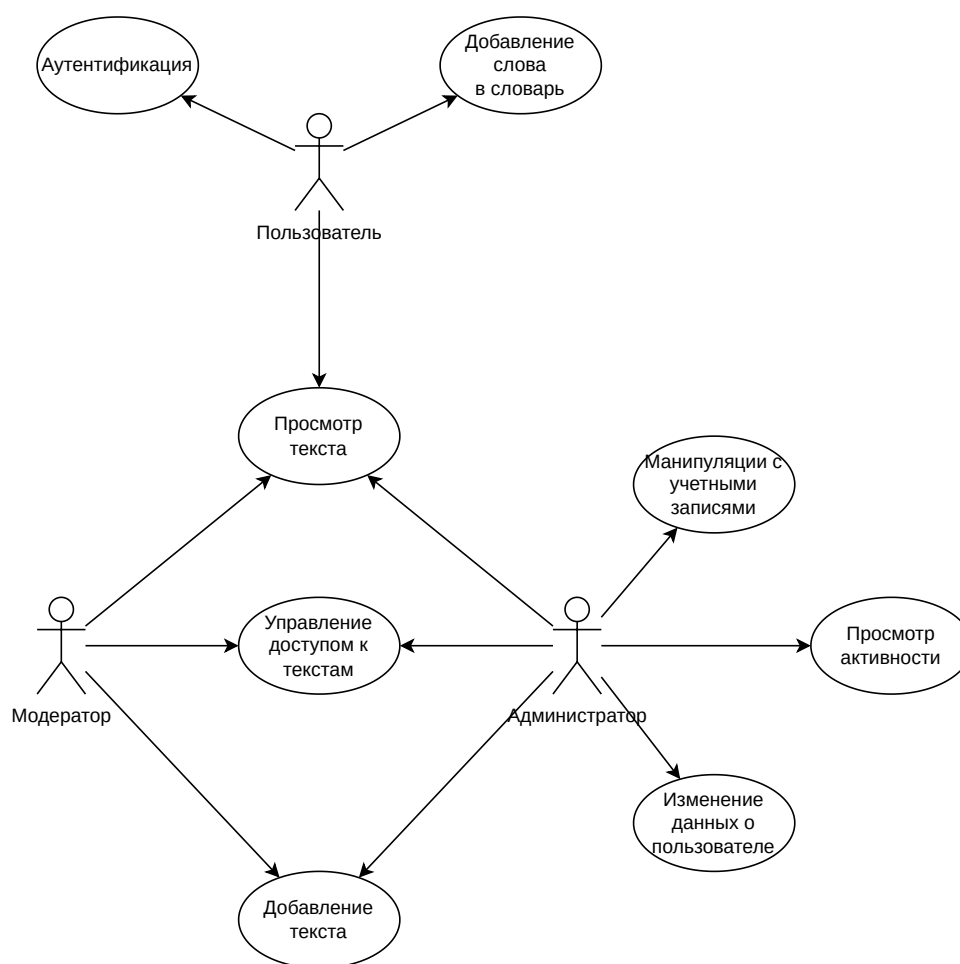


Рисунок 1.2 – Диаграмма вариантов использования приложения

Вывод из аналитического раздела

В данном разделе была рассмотрена предметная область и подходы к изучению японского языка. Также были представлены основные сведе-

ния о базах данных и системах управления базами данных для хранения японских иероглифов и проведен анализ существующих приложений для упрощения изучения японского языка. В ходе работы будет использоваться СУБД PostgreSQL, так как она имеет открытый исходный код, высокую производительность и безопасность, а также поддерживает хранение в базе данных иероглифов, процедур и триггеров.

При разработке приложения будет учитываться отсутствие у крупных приложений для изучения японского языка механизмов для чтения печатных текстов. Приложение будет доступно как в веб версии, так и на мобильных платформах и персональных компьютерах.

2 Конструкторский раздел

В данном разделе представлена схема базы данных и диаграмма прецедентов. Также описано взаимодействие компонентов системы.

2.1 Концептуальная модель системы

Фундаментальной функцией разрабатываемого приложения является преобразование изображения, содержащего текст, написанный на японском языке, в текстовый документ. На рисунках 2.1–2.2 представлена модель системы в нотации IDEF0.

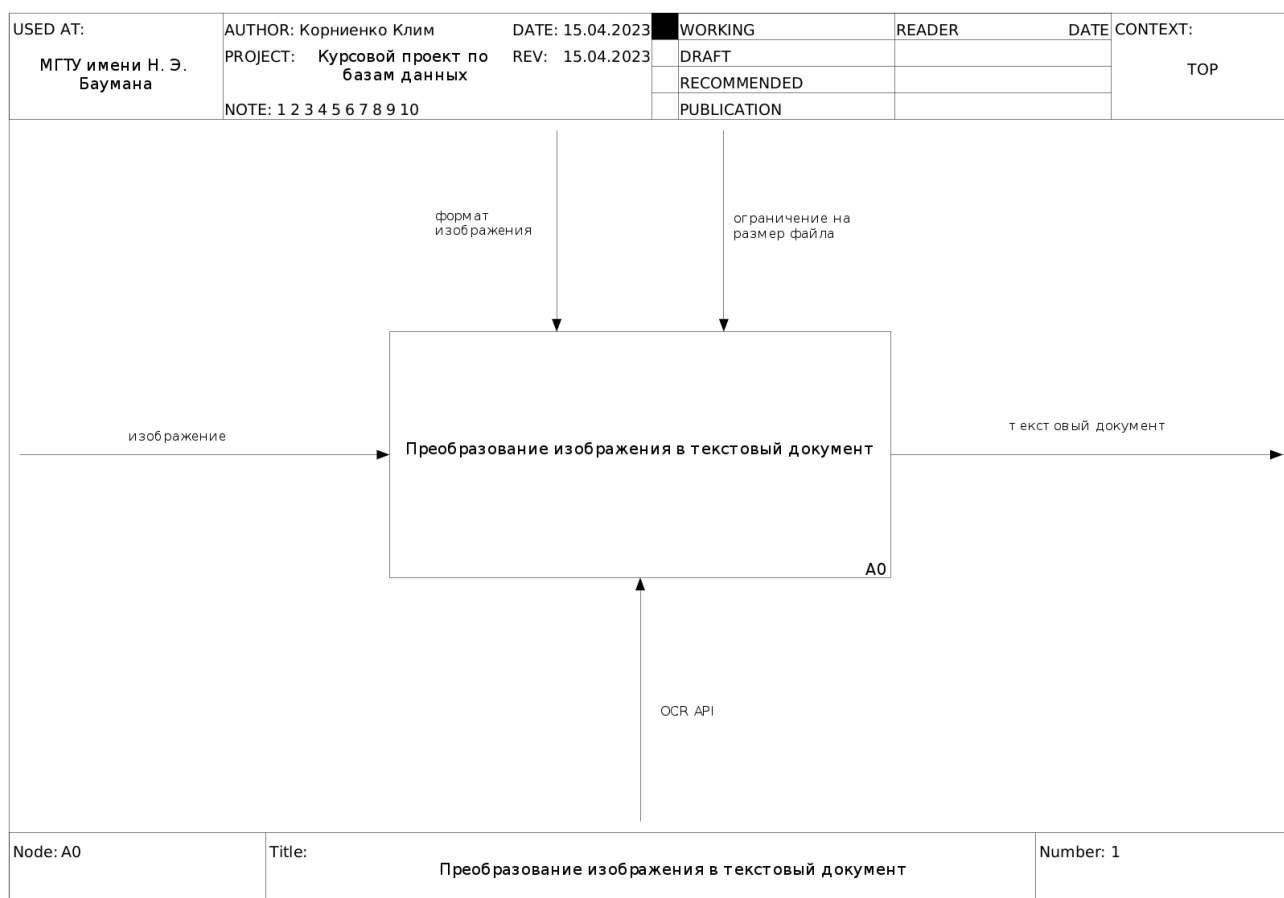


Рисунок 2.1 – Модель преобразования изображения в текст (верхний уровень)

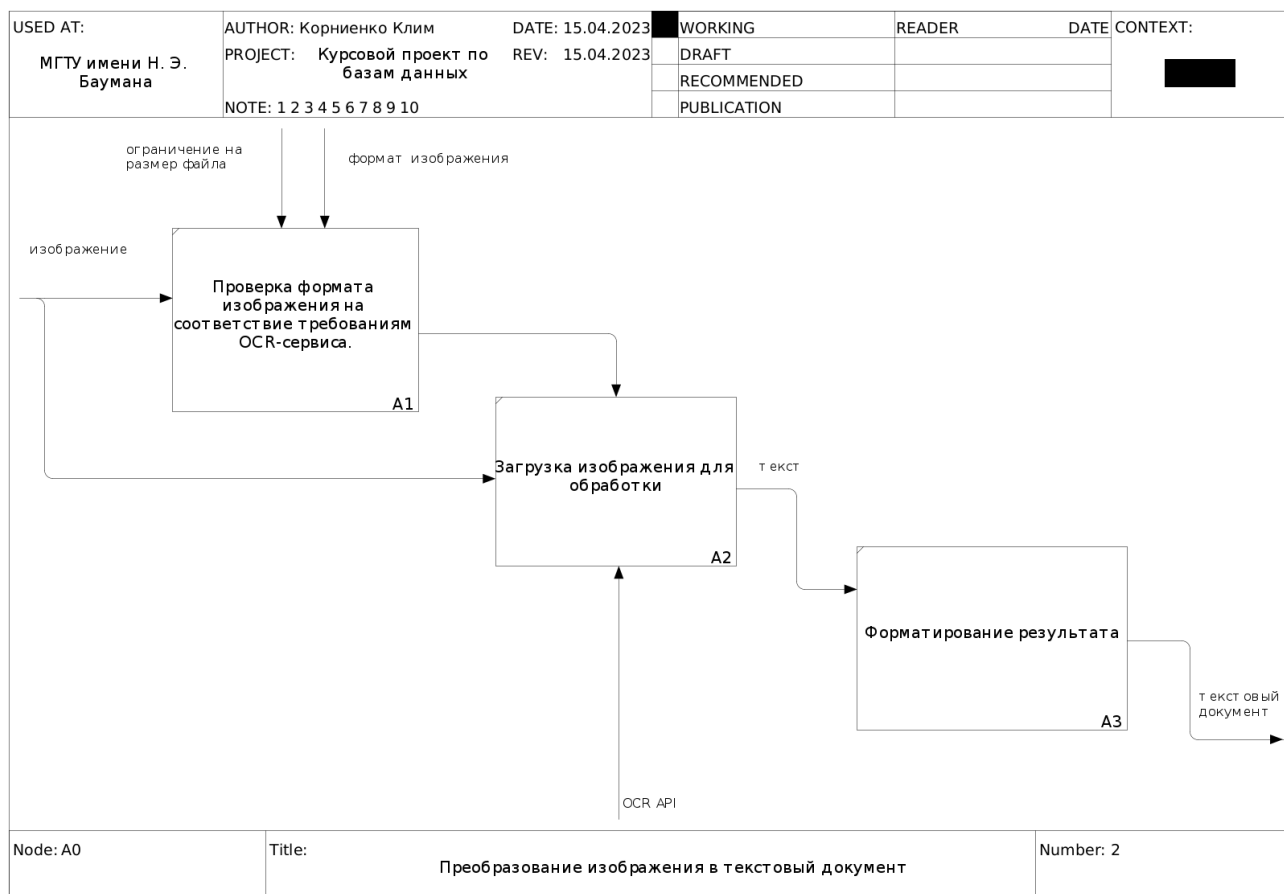


Рисунок 2.2 – Модель преобразования изображения в текст (первый уровень)

2.2 Диаграмма прецедентов

На рисунках 2.3–2.5 представлены диаграммы прецедентов.

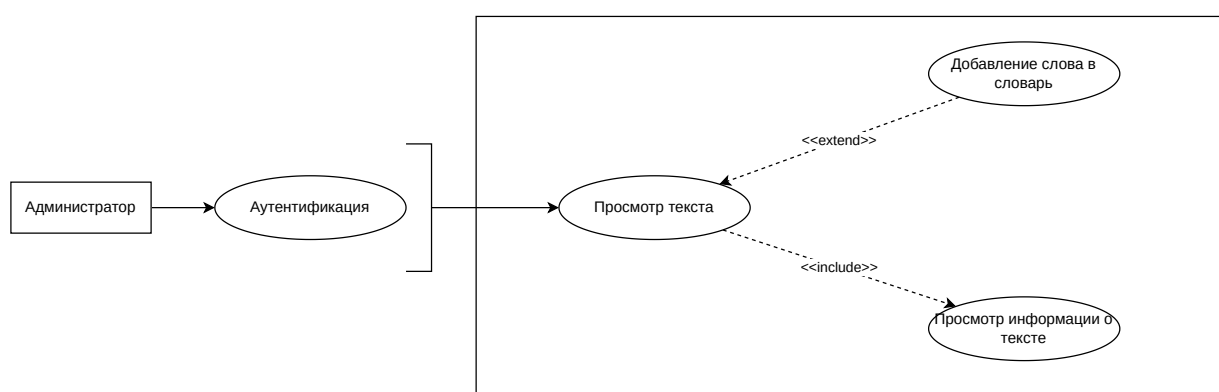


Рисунок 2.3 – Диаграмма прецедентов (пользователь)

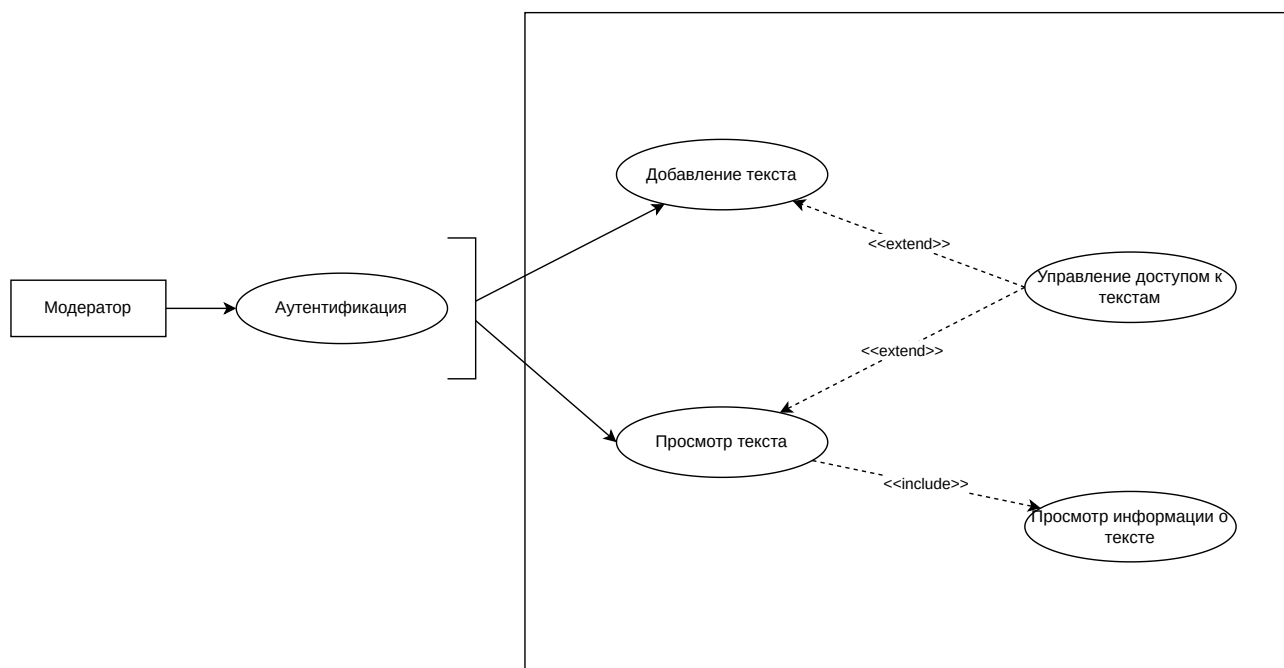


Рисунок 2.4 – Диаграмма прецедентов (модератор)

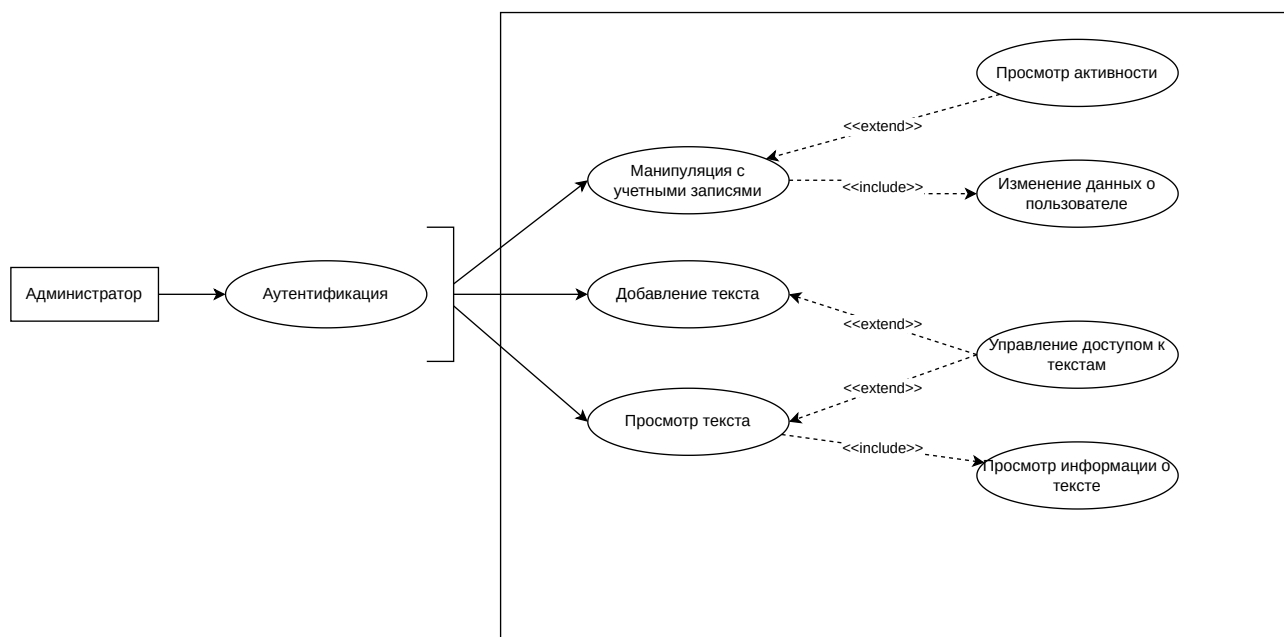


Рисунок 2.5 – Диаграмма прецедентов (администратор)

2.3 Схема базы данных

3 Технологический раздел

4 Исследовательский раздел

ЗАКЛЮЧЕНИЕ

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Мухтарова А. А.* Использование ИТ-технологий при изучении английского и японского языка // Умная цифровая экономика. — 2022. — Т. 2, № 3. — С. 34—41.
2. *Cheng T.* Applying networked learning to improve learner interactions: A new paradigm of teaching and learning in ODL // Asian Association of Open Universities Journal. — 2013. — Т. 8. — С. 67—85. — DOI: 10.1108/AAOUJ-08-02-2013-B006.
3. Recognizing modern Japanese magazines by combining Deep Learning with language models / N. T. Nguyen [и др.] // 2021 13th International Conference on Knowledge and Systems Engineering (KSE). — 2021. — С. 1—6. — DOI: 10.1109/KSE53942.2021.9648643.
4. *Sasanti N. S.* Japanese Language Learning Consistency in the Digital Era // Jurnal Ilmiah Lingua Idea. — 2022. — Т. 13. — С. 207—219.
5. *Febrianty F., Ricardo R.* Information Technology for Japanese Learning // IOP Conference Series: Materials Science and Engineering. — 2019. — Т. 662, № 2. — С. 022117. — DOI: 10.1088/1757-899X/662/2/022117. — URL: <https://dx.doi.org/10.1088/1757-899X/662/2/022117>.
6. *Дейт К. Д.* Введение в системы баз данных. — М. : Издательский дом "Вильямс", 2005.
7. *Elmasri R., Navathe S. B.* Fundamentals of Database Systems / под ред. M. Hirsch. — London : Pearson, 2010.
8. *A Malik A Burney F. A.* A Comparative Study of Unstructured Data with SQL and NO-SQL Database Management Systems // Journal of Computer and Communications. — 2020. — Т. 8. — С. 59—71. — DOI: 10.4236/jcc.2020.84005.
9. *Мейер М.* Теория реляционных баз данных. — М. : Мир, 1987.
10. *К Н Васильева Г. Я. Х.* Реляционные базы данных // Colloquium-Journal. — Warszawa, 2020. — С. 22—23.
11. Microsoft: SQL Server. [Электронный Ресурс]. — Режим Доступа: <https://www.microsoft.com/en-us/sql-server/> (дата обращения: 13.04.2023).

12. MariaDB Server: The open source relational database. [Электронный Ресурс]. — Режим Доступа: <https://mariadb.org/> (дата обращения: 13.04.2023).
13. PostgreSQL: The World's Most Advanced Open Source Relational Database. [Электронный Ресурс]. — Режим Доступа: <https://www.postgresql.org/> (дата обращения: 13.04.2023).
14. Database | Oracle. [Электронный Ресурс]. — Режим Доступа: <https://www.oracle.com/database/> (дата обращения: 16.04.2023).
15. IBM Db2 | IBM. [Электронный Ресурс]. — Режим Доступа: <https://www.ibm.com/products/db2> (дата обращения: 16.04.2023).
16. SQLite Home Page. [Электронный Ресурс]. — Режим Доступа: <https://sqlite.org/index.html> (дата обращения: 16.04.2023).
17. *Ramya Gangaianaran M. P.* Review on Use of Mobile Apps for Language Learning // International Journal of Applied Engineering Research. — 2017. — Т. 12, № 21. — С. 11242—11251. — ISSN 0973-4562.
18. Duolingo - The world's best way to learn a language. [Электронный Ресурс]. — Режим Доступа: <https://www.duolingo.com/> (дата обращения: 14.04.2023).
19. Learn a language. Memrise is authentic, useful & personalised. [Электронный Ресурс]. — Режим Доступа: <https://www.memrise.com/> (дата обращения: 14.04.2023).
20. LingoDeer: Learn Japanese, Korean, Chinese and more. [Электронный Ресурс]. — Режим Доступа: <https://www.lingodeer.com/> (дата обращения: 14.04.2023).
21. WaniKani, a kanji learning application by Tofugu. [Электронный Ресурс]. — Режим Доступа: <https://www.wanikani.com/> (дата обращения: 14.04.2023).
22. *Ohta A. S.* Second Language Acquisition Processes In The Classroom: Learning Japanese. — London : Lawrence Erlbaum Associates, 2001.

ПРИЛОЖЕНИЕ А