

Государственное учреждение образования
“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ”

Кафедра: Интеллектуальных информационных технологий
Дисциплина: Объектное моделирование интеллектуальных систем

Отчет по лабораторной работе №2

Выполнил:
студент гр.321701
Семков А.А.

Проверил:
Зотов Н.В.

Минск 2025

1. ER-диаграмма Подсистемы загрузки и предобработки текстов

1.1 Сущности и их атрибуты

1.1.1 Пользователь (Сущность, представляющая пользователя системы)

User_ID (PK): Уникальный идентификатор пользователя

Имя: ФИО пользователя

Роль: Роль в системе (Студент, Преподаватель, Администратор)

1.1.2 Исходный документ (Сущность, представляющая загружаемый документ)

Doc_ID (PK): Уникальный идентификатор документа

Имя файла: Название загруженного файла

Формат: Формат документа (PDF, DOCX, DOC, TXT)

Размер: Размер файла в байтах

1.1.3 Обработанный текст (Сущность, представляющая нормализованный текст)

Text_ID (PK): Уникальный идентификатор обработанного текста

Извлеченный текст: Текст, извлеченный из документа

Статус обработки: Статус предобработки (В процессе, Завершена, Ошибка)

1.2 Связи между сущностями

Загружает: Связь между Пользователь и Исходный документ. Один пользователь может загружать множество документов. Каждый документ загружается одним пользователем. (1:N)

Преобразуется в: Связь между Исходный документ и Обработанный текст. Один исходный документ преобразуется в один обработанный текст. Каждый обработанный текст соответствует одному исходному документу. (1:1)

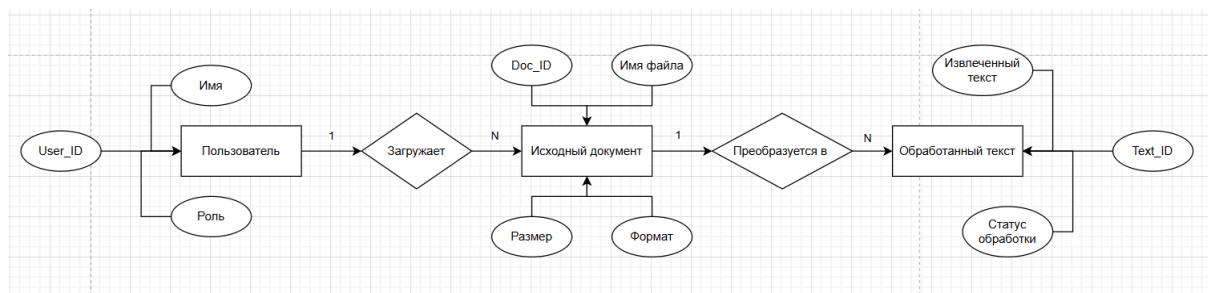


рис. 1 - ER-диаграмма Подсистемы загрузки и предобработки текстов

2. ER-диаграмма Подсистемы сравнения и анализа

2.1 Сущности и их атрибуты

2.1.1 Обработанный текст (Сущность-связка с предыдущей подсистемой)

Text_ID (PK): Уникальный идентификатор текста

2.1.2 Проверка (Сущность, представляющая акт проверки документа)

Check_ID (PK): Уникальный идентификатор проверки

Дата и время проверки: Временная метка начала проверки

2.1.3 Совпадение (Сущность, представляющая найденное заимствование)

Match_ID (PK): Уникальный идентификатор совпадения

Позиция в тексте: Местоположение совпадения в документе

Процент совпадения: Уровень соответствия фрагмента

2.1.4 Источник (Сущность, представляющая источник заимствования)

Source_ID (PK): Уникальный идентификатор источника

Название источника: Наименование источника

Автор: Автор исходного материала

URL источника: Ссылка на источник

2.2 Связи между сущностями

Анализирует: Связь между Проверка и Обработанный текст. Одна проверка анализирует один обработанный текст. Один текст может анализироваться в нескольких проверках. (N:1)

Выявляет: Связь между Проверка и Совпадение. Одна проверка выявляет множество совпадений. Каждое совпадение выявляется одной проверкой. (1:N)

Ссылается на: Связь между Совпадение и Источник. Одно совпадение ссылается на один источник. Один источник может быть упомянут во многих совпадениях. (N:1)

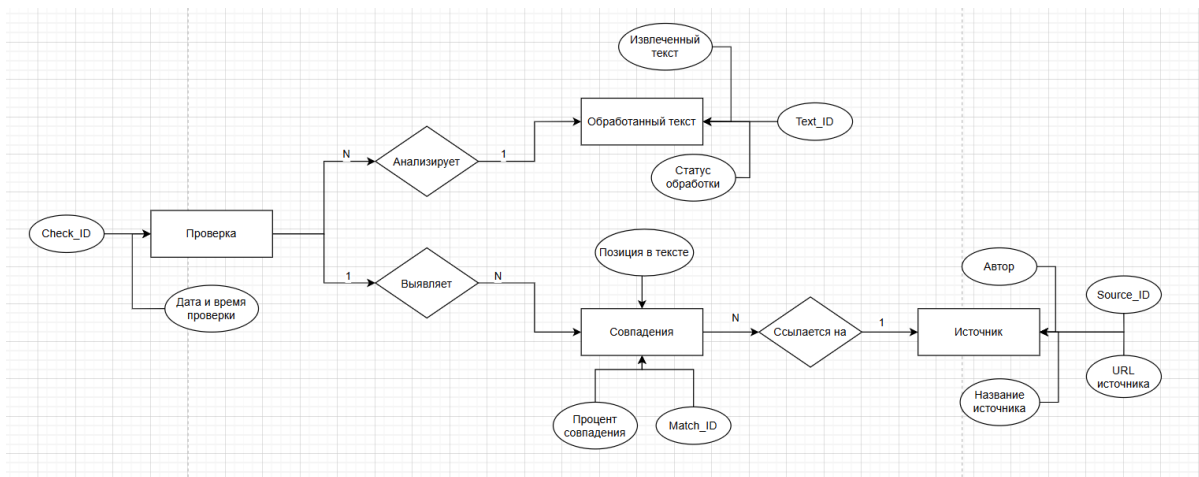


рис. 2 – ER-диаграмма Подсистемы сравнения и анализа

3. ER-диаграмма Подсистемы формирования отчетов

3.1 Сущности и их атрибуты

3.1.1 Пользователь (Сущность из подсистемы загрузки)

- **User_ID (PK):** Уникальный идентификатор пользователя
- **Имя:** ФИО пользователя
- **Роль:** Роль в системе

3.1.2 Проверка (Сущность из подсистемы анализа)

- **Check_ID (PK):** Уникальный идентификатор проверки
- **Дата и время проверки:** Временная метка проверки

3.1.3 Отчет (Сущность, представляющая сгенерированный отчет)

- **Report_ID (PK):** Уникальный идентификатор отчета
- **Дата формирования:** Дата создания отчета
- **Процент уникальности:** Итоговый показатель оригинальности текста

3.2 Связи между сущностями

Лежит в основе: Связь между Проверка и Отчет. Одна проверка лежит в основе одного отчета. Каждый отчет основан на одной проверке. (1:1)

Получает: Связь между Пользователь и Отчет. Один пользователь может получать множество отчетов. Каждый отчет получается одним пользователем. (1:N)

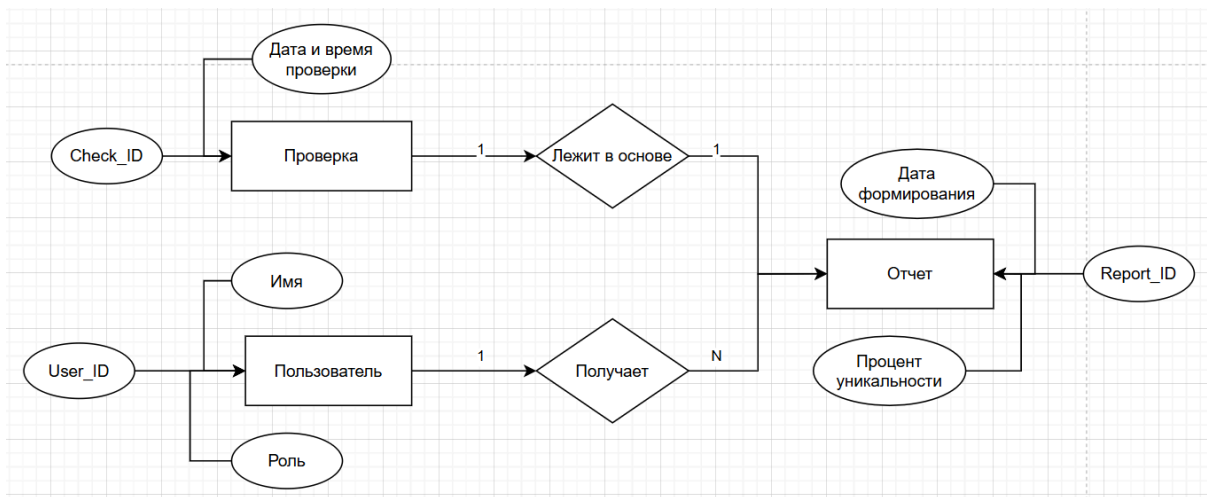


рис. 3 – ER-диаграмма Подсистемы формирования отчетов

4. ER-диаграмма Интерфейсной подсистемы

4.1 Сущности и их атрибуты

4.1.1 Пользователь (Основная сущность системы)

- **User_ID (PK)**: Уникальный идентификатор пользователя
- **Имя**: ФИО пользователя
- **Роль**: Роль в системе

4.1.2 Сессия пользователя (Сущность, представляющая сеанс работы)

- **Session_ID (PK)**: Уникальный идентификатор сессии
- **Время входа**: Время начала сессии
- **Время выхода**: Время завершения сессии

4.1.3 Интерфейсный элемент (Сущность, представляющая элемент UI)

- **Element_ID (PK)**: Уникальный идентификатор элемента
- **Название элемента**: Наименование элемента интерфейса

4.2 Связи между сущностями

Имеет: Связь между Пользователь и Сессия пользователя. Один пользователь может иметь множество сессий. Каждая сессия принадлежит одному пользователю. (1:N)

Взаимодействует с: Связь между Сессия пользователя и Интерфейсный элемент. Одна сессия включает взаимодействие со многими элементами интерфейса. Один элемент интерфейса может использоваться во многих сессиях. (M:N)

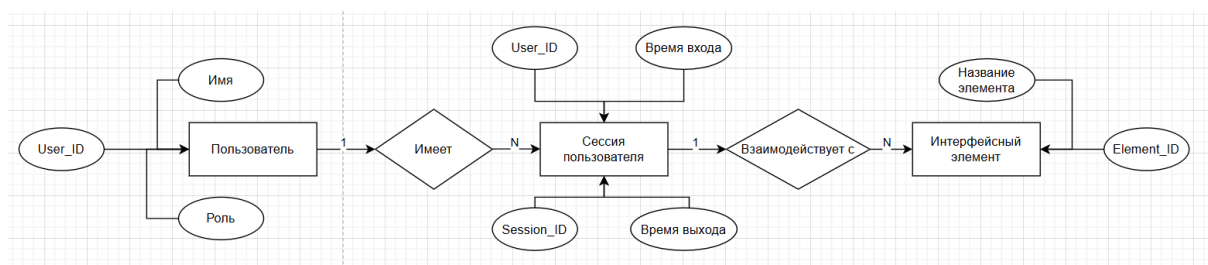


рис. 4 – ER-диаграмма Интерфейсной подсистемы

5. ER-диаграмма Подсистемы хранения данных (интеграционная)

5.1 Сущности и их атрибуты (объединение всех подсистем)

5.1.1 Пользователь

- **User_ID (PK):** Уникальный идентификатор
- **Имя:** ФИО пользователя
- **Роль:** Роль в системе

5.1.2 Исходный документ

- **Doc_ID (PK):** Уникальный идентификатор
- **Имя файла:** Название файла
- **Формат:** Формат документа
- **Размер:** Размер файла

5.1.3 Обработанный текст

- **Text_ID (PK):** Уникальный идентификатор
- **Статус обработки:** Статус предобработки

5.1.4 Проверка

- **Check_ID (PK):** Уникальный идентификатор
- **Дата и время проверки:** Временная метка

5.1.5 Совпадение

- **Match_ID (PK):** Уникальный идентификатор
- **Позиция в тексте:** Местоположение совпадения

5.1.6 Источник

- **Source_ID (PK):** Уникальный идентификатор
- **Название источника:** Наименование источника
- **Автор:** Автор материала

5.1.7 Отчет

- **Report_ID (PK):** Уникальный идентификатор
- **Дата формирования:** Дата создания
- **Процент уникальности:** Показатель оригинальности

5.2 Связи между сущностями (интеграционные)

Загружает: Пользователь → Исходный документ (1:N)

Преобразуется в: Исходный документ → Обработанный текст (1:1)

Анализирует: Проверка → Обработанный текст (N:1)

Выявляет: Проверка → Совпадение (1:N)

Ссылается на: Совпадение → Источник (N:1)

Лежит в основе: Проверка → Отчет (1:1)

Получает: Пользователь → Отчет (1:N)

Имеет: Пользователь → Сессия пользователя (1:N)

Взаимодействует: Сессия пользователя → Интерфейсный элемент (M:N)

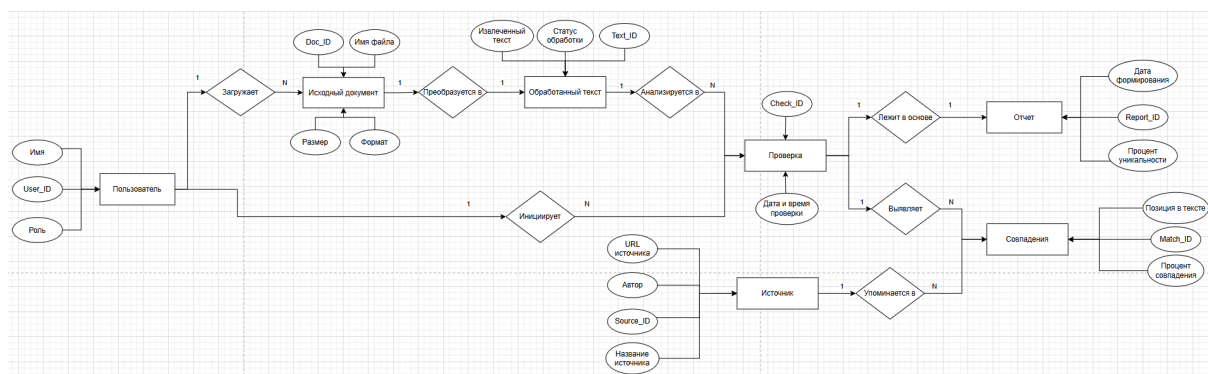


рис. 5 – ER-диаграмма Подсистемы хранения данных

2. BPMN-диаграммы.

2.1 BPMN-диаграмма Процесса загрузки и подготовки документа:

Пользователь начинает процесс проверки документа на плагиат. Он выбирает файл для анализа и отправляет запрос на загрузку в систему. Система принимает документ и проверяет его формат.

Если формат документа не поддерживается системой, пользователь получает уведомление об ошибке и может выбрать новый документ для загрузки. Если формат корректен, система начинает подготовку текста - извлекает содержимое из файла, нормализует его, удаляет лишние элементы.

После успешной обработки текст готов к проверке на плагиат. Система уведомляет пользователя о том, что документ подготовлен и можно переходить к этапу анализа. На этом процесс загрузки и предобработки завершается.

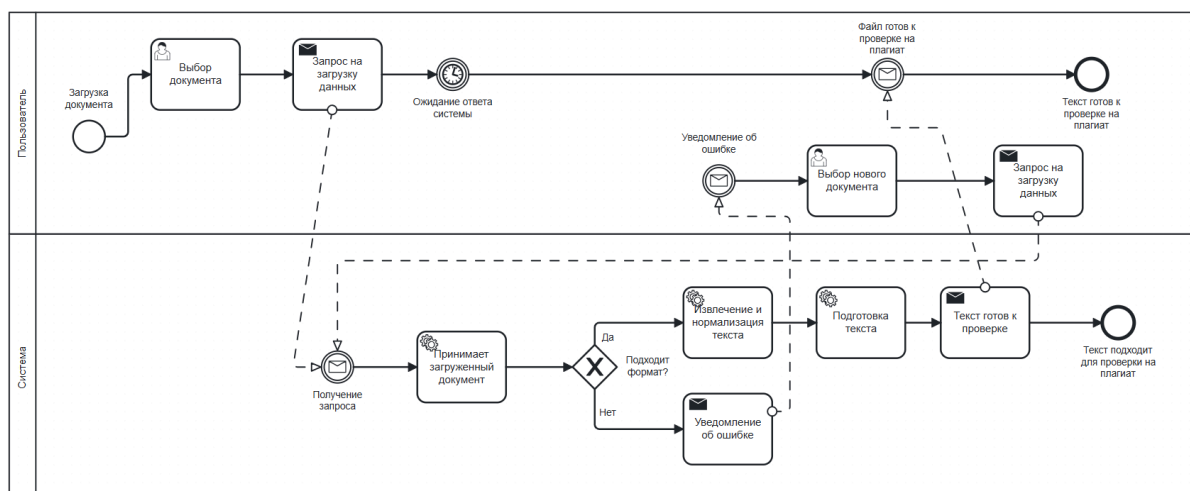


рис. 6 – BPMN-диаграмма подсистемы загрузки и предобработки текстов

2.2 BPMN-диаграмма Процесса анализа текста на плагиат:

Система получает подготовленный текст для анализа. Начинается процесс сравнения с различными источниками. Система выполняет направленный поиск совпадений в академических базах данных, интернет-источниках и публикациях.

Происходит сбор и обработка данных из multiple источников. Система анализирует полученные результаты, выявляя дословные совпадения и перефразирования. Для сложных случаев используется углубленный анализ с применением семантических и синтаксических алгоритмов.

На основе собранных данных система рассчитывает процент уникальности текста. Формируются детальные данные для отчета: список найденных заимствований, источники, цитаты. После завершения анализа система уведомляет о готовности результатов проверки.

Процесс сравнения и анализа завершается передачей подготовленных данных в подсистему формирования отчетов для финального оформления результатов.

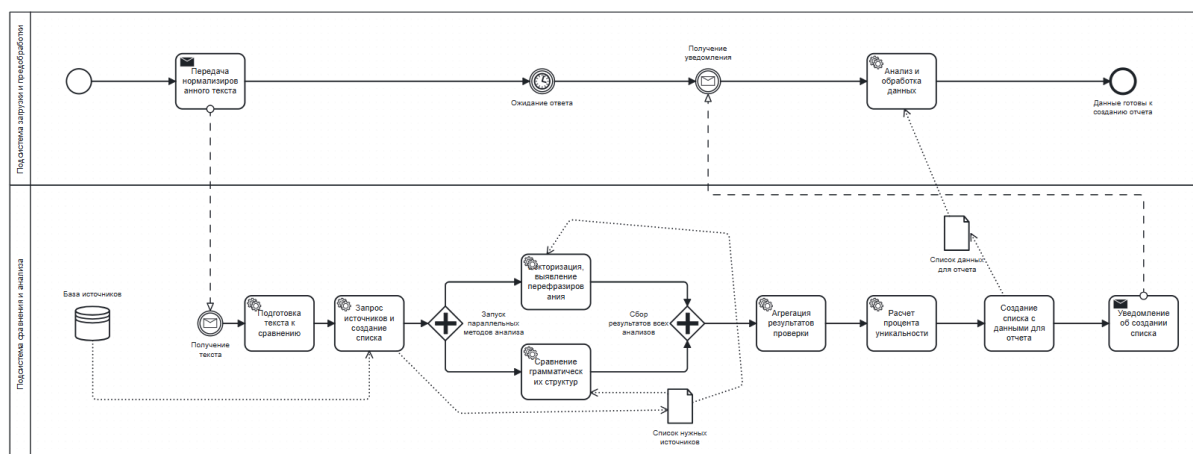


рис. 7 – BPMN-диаграмма Процесса анализа текста на плагиат