



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 и системы управления»

---

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

---

# РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

## *К КУРСОВОЙ РАБОТЕ*

### *НА ТЕМУ:*

*«Проектирование БД системы обогащения обучающей  
выборки для автоматизированной проверки отчета на  
соответствие нормативным требованиям»*

Студент ИУ7-64Б  
(Группа)

\_\_\_\_\_  
(Подпись, дата)

Разин А.В.  
(И. О. Фамилия)

Руководитель курсовой работы

\_\_\_\_\_  
(Подпись, дата)

Строганов Ю. В.  
(И. О. Фамилия)

*2024 г.*

# СОДЕРЖАНИЕ

<b>ВВЕДЕНИЕ</b>	<b>5</b>
<b>1 Аналитическая часть</b>	<b>6</b>
1.1 Основные ошибки в отчетах . . . . .	6
1.2 Прием лабораторных работ . . . . .	6
1.2.1 Участники процесса приема работ . . . . .	6
1.2.2 Процесс приема работ . . . . .	6
1.3 Формализация требований к базе данных и приложению . . . . .	8
1.4 Анализ существующих средств автоматизации . . . . .	8
1.5 Формализация информации, подлежащей хранению в проектируемой базе данных . . . . .	10
1.6 Формализация пользователей . . . . .	12
1.7 Анализ существующих баз данных . . . . .	13
1.7.1 Дореляционные модели . . . . .	14
1.7.2 Реляционные модели . . . . .	15
1.7.3 Постреляционные модели . . . . .	15
1.7.4 Хранение временных рядов . . . . .	15
<b>2 Конструкторская часть</b>	<b>17</b>
2.1 Формализация взаимодействия с приложением . . . . .	17
2.2 Формализация работы системы проверки отчета . . . . .	20
2.3 Формализация сущностей базы данных . . . . .	20
2.4 Ограничения целостности базы данных . . . . .	20
2.5 Ролевая модель . . . . .	21
2.6 Определение модели базы данных . . . . .	23
2.7 Используемые триггеры . . . . .	23
<b>3 Технологическая часть</b>	<b>25</b>
3.1 Выбор СУБД . . . . .	25
3.1.1 Доступность лицензии . . . . .	26
3.2 Выбор средств реализации . . . . .	27
3.3 Реализация создания отношений . . . . .	27
3.4 Реализация функций . . . . .	28

3.5	Реализация триггеров . . . . .	28
3.6	Описание приложения . . . . .	28
<b>4</b>	<b>Исследовательская часть</b>	<b>29</b>
	<b>ЗАКЛЮЧЕНИЕ</b>	<b>30</b>
	<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>	<b>32</b>
	<b>ПРИЛОЖЕНИЕ А</b>	<b>33</b>
A.1	Основные ошибки в отчетах . . . . .	33
A.1.1	Общие ошибки . . . . .	33
A.1.2	Ошибки в тексте . . . . .	33
A.1.3	Ошибки в рисунках . . . . .	34
A.1.4	Ошибки в таблицах . . . . .	35
A.1.5	Ошибки в формулах . . . . .	36
A.1.6	Ошибки в списках . . . . .	36
A.1.7	Ошибки в списке литературы . . . . .	37
	<b>ПРИЛОЖЕНИЕ Б</b>	<b>38</b>
	<b>ПРИЛОЖЕНИЕ В</b>	<b>42</b>

## ВВЕДЕНИЕ

Во время обучения студентам регулярно приходится писать отчеты к различным видам работ (курсовые, лабораторные, научно-исследовательские работы и т. д.), при этом оформление работ должно соответствовать ГОСТ, что необходимо своевременно проверить и при необходимости отправить отчет на доработку, однако, количество студентов намного превышает количество нормоконтроллеров. Для ускорения процесса проверки возможно использование автоматических систем.

Целью курсовой работы является разработка базы данных обогащения обучающей выборки для автоматизированной проверки отчета на соответствие нормативным требованиям.

Для достижения цели научно-исследовательской работы требуется решить следующие задачи:

- проанализировать существующие решения;
- формализовать задачу и определить необходимый функционал;
- проанализировать способы хранения данных и системы управления базами данных, выбрать подходящую систему для поставленной цели;
- спроектировать базу данных, описать ее сущности и связи;
- спроектировать и разработать базу данных;
- исследовать зависимость времени выполнения запроса от числа получаемых запросов;

## **1 Аналитическая часть**

В данной части работы будут описаны ошибки в отчетах, которые необходимо обнаружить, а также описаны участники процесса приема лабораторных работ, также будут описаны существующие средства автоматизации.

### **1.1 Основные ошибки в отчетах**

Основные ошибки в отчетах описаны в приложении А.

### **1.2 Прием лабораторных работ**

В не автоматизированной системе проверки отчетов на соответствие ГОСТ и дополнительным требованиям присутствуют две роли: студент, выполняющий некоторую работу, которая подразумевает написание отчета и нормоконтроллер, принимающий экспертное решение о соответствии предоставленного ему отчета необходимым требованиям.

#### **1.2.1 Участники процесса приема работ**

С помощью использования автоматической проверки отчета возможно сократить временные ресурсы, выделяемые нормоконтроллером на проверку отчетов студентов, однако, полностью отказаться от финального контроля результатов человеком невозможно, таким образом существует две роли при проверке отчета на соответствие ГОСТ, а именно: студент и нормоконтроллер.

#### **1.2.2 Процесс приема работ**

Студент отправляет отчет на проверку, а затем получает результат со списком ошибок (если имеются). Нормоконтроллер же анализирует отчет, составленный автоматической системой проверки, и при необходимости может внести необходимые правки. Диаграммы процесса проверки отчетов приведены на картинках 1.1–1.2, также приведена диаграмма BPMN 2.0, на которой представлено взаимодействие системы проверки отчетов, нормоконтроллера и студента (рисунок 1.3).

Использование автоматической проверки отчетов на соответствие ГОСТ и дополнительным требованиям сократит временные затраты на проверку отчетов.

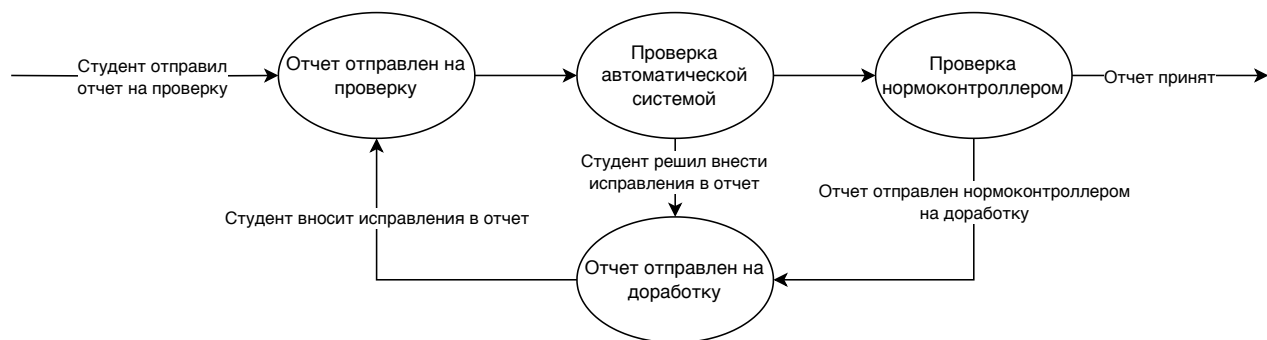


Рисунок 1.1 – Диаграмма состояний проверки отчета



Рисунок 1.2 – Диаграмма последовательности действий

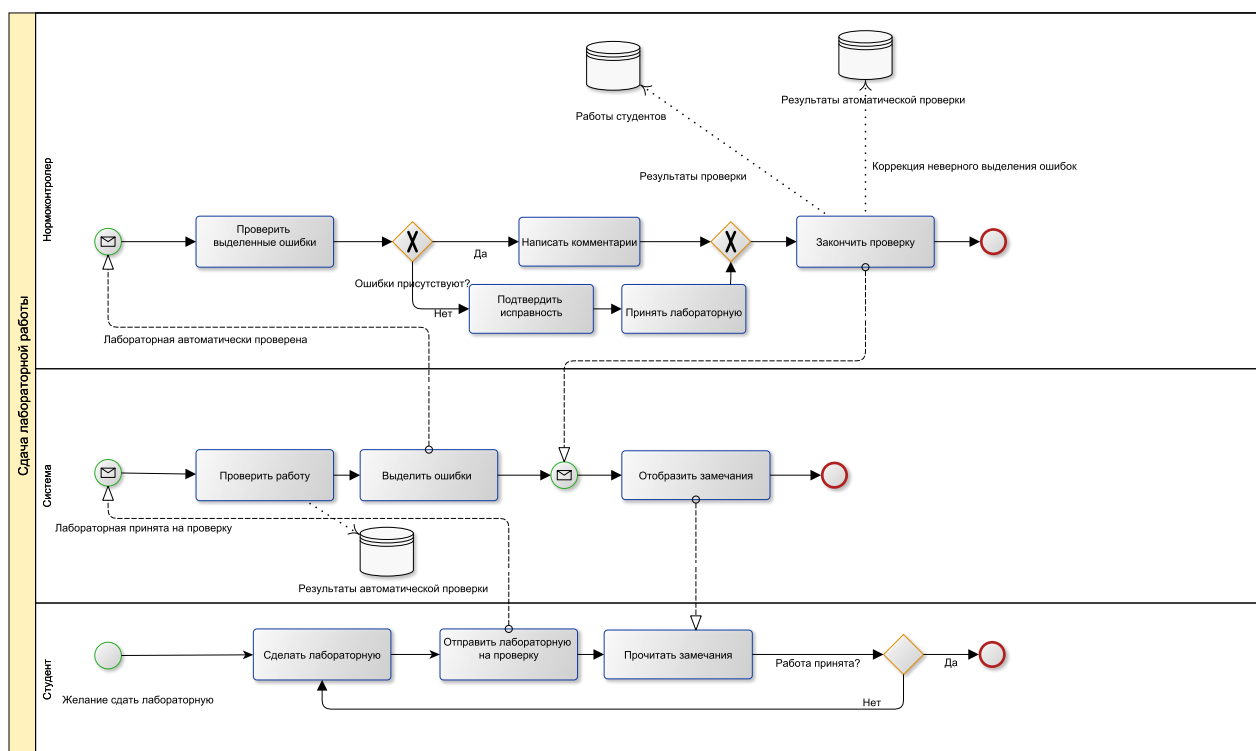


Рисунок 1.3 – BPMN 2.0 диаграмма сдачи лабораторной работы

## 1.3 Формализация требований к базе данных и приложению

В ходе выполнения курсовой работы необходимо разработать базу данных для хранения информации о студентах их работ, достижениях и результатов проверки работ студентов, также стоит хранить разметку основных частей отчетов студентов.

Необходимо спроектировать и разработать приложение, которое позволит проверять отчеты студентов, сохранять их, а также сохранять новые разметки основных частей отчета.

## 1.4 Анализ существующих средств автоматизации

Ввиду распространенности решаемой проблемы уже были созданы приложения для автоматизации проверки документов на соответствие стандартам.

Наиболее популярными из них являются:

1. ВКР-СМАРТ [1];
2. TestVkr [2];

### 3. Applitools visual testing [3];

Система ВКР-СМАРТ, предназначенная для проверки выпускных квалификационных работ (ВКР) студентов, представляет собой универсальную платформу, разработанную для системного хранения и проверки на заимствования ВКР и других работ обучающихся. Также система проверяет предложенные работы на выполнение всех требований ФГОС ВО и СПО, а также соответствие ГОСТам. После выполнения проверок, будет получен отчет о проценте заимствования и замечания по нарушенным стандартам [1].

Система ТЕСТ ВКР (Технический регламент проверки выпускных квалификационных работ) предназначена для проверки выпускных квалификационных работ студентов на объем заимствования и их размещения в электронно-библиотечной системе (ЭБС) университета. Система обеспечивает централизованное хранение и контроль за академическими работами студентов, а также их проверку на оригинальность и уникальность контента, также данную систему (без использования хранилища) возможно запускать локально, для проверки работы на нарушение ГОСТ [2].

Платформа Applitools позволяет использовать «визуальное тестирование» предназначенное для сравнения получаемого изображения с реальным. Этот метод особенно эффективен при выявлении ошибок во внешнем виде страницы или экрана, которые могут остаться незамеченными при традиционном функциональном тестировании. С использованием Applitools Eyes разработчики могут легко интегрировать визуальные тесты, которые могут быть использованы для выявления отклонений от стандартов в PDF [3].

Таблица 1.1 – Сравнение существующих средств автоматизации

Критерий	ВКР СМАРТ	TestVkr	Applitools
Проверка текстов	да	да	нет
Проверка элементов отчета	нет	нет	да
Наличие общего хранилища работ	да	да	нет
Возможность запуска локально	нет	*да	*да

В таблице 1.1 под «элементами отчета» подразумеваются таблицы, рисунки, схема алгоритмов, формулы, использование символа \*, означает, что в



этом случае проверка плагиата не производится.

## **1.5 Формализация информации, подлежащей хранению в проектируемой базе данных**

Разрабатываемая база данных, должна содержать информацию о следующих сущностях:

- студент;
- нормоконтроллер;
- отчет студента;
- тип фрагмента отчета;
- фрагмент отчета;
- комментарий студента;
- достижение студента.

Сведения о каждой категории данных содержится в таблице 1.2.

Таблица 1.2 – Категории и сведения о данных

Категория	Сведения
Студент	ID источника данных, ID студента, никнейм, имя, фамилия, дата регистрации, число сданных лабораторных
Отчет	ID отчета, номер попытки сдачи, файл отчета, время последнего обновления отчета, число проверок отчета, ID студента, значение того что отчет сдан
Тип фрагмента отчета	ID фрагмента, описание типа фрагмента, ID контроллера, создавшего фрагмент отчета
Выделенный фрагмент отчета	ID фрагмента отчета, изображение страницы отчета, данные фрагмента отчета (разметка), значение была ли ошибка проверена разметчиком, ID отчета, ID типа фрагмента отчета, ID проверяющего отчет
Комментарий	ID ошибки, на которую создается комментарий, ID комментария, данные комментария, ID студента, создавшего комментарий
Достижение	ID достижения, ID контроллера, создавшего достижение, данные достижения, описание достижения

На основе описанной информации была получена диаграмма сущность—связь, представленная на рисунке 1.4.

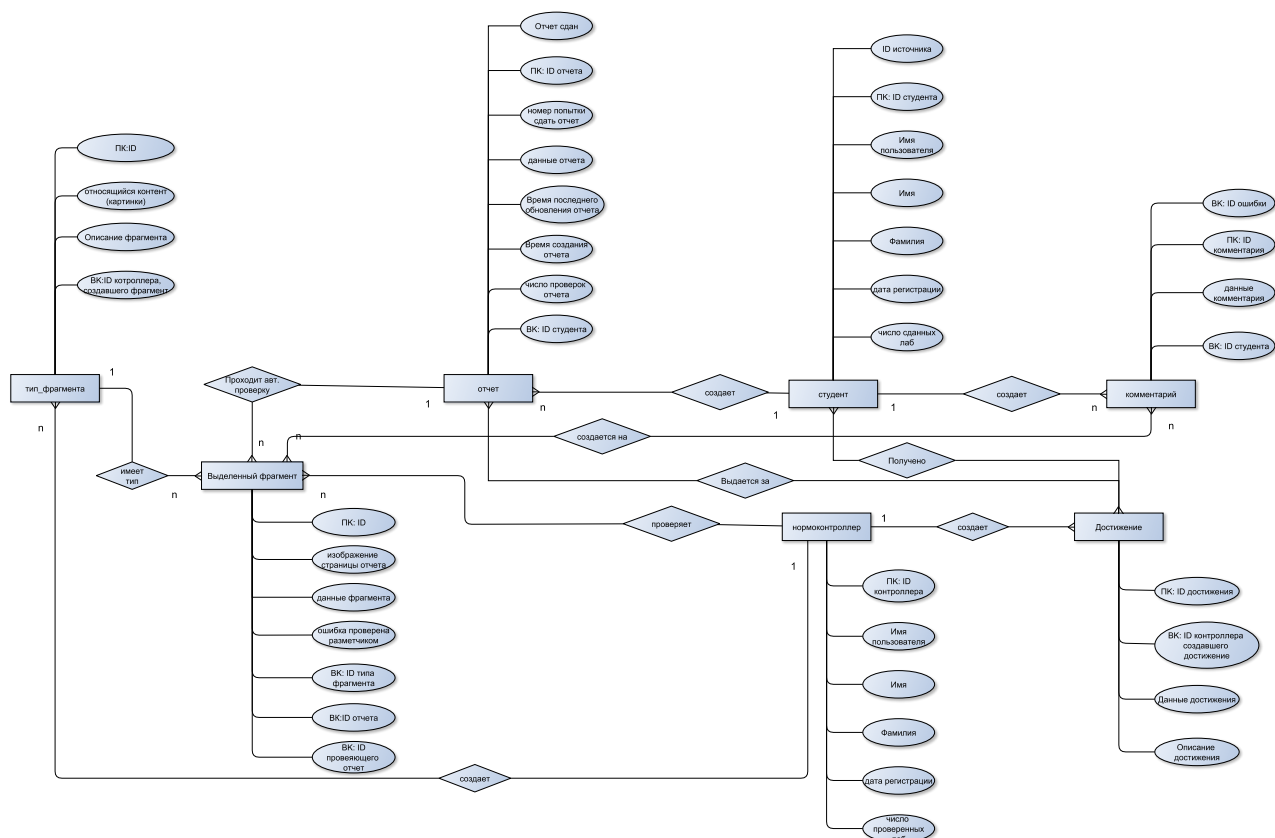


Рисунок 1.4 – Диаграмма сущность—связь

## 1.6 Формализация пользователей

При анализе предметной области были выделены следующие роли в системе:

1. Гость — данная роль представляет не авторизованного пользователя, соответственно он может авторизоваться и войти в систему.
2. Студента — данная роль позволяет создавать комментарии к отчетам, просматривать отправленные документы и ошибки в них.
3. Нормоконтроллер — данная роль позволяет создавать типы разметок, разметки элементов отчета, просматривать отправленные документы и ошибки в них, а также создавать, удалять, модифицировать достижения и создавать комментарии.
4. Админ — возможности данной роли аналогичны роли нормоконтроллера, однако данная роль также позволяет удалять типы элементов отчета и изменять роль пользователей (со студента на нормоконтроллера и наоборот).

Для описания сценариев взаимодействия пользователей с системой была создана диаграмма прецедентов, представленная на рисунке 1.5.

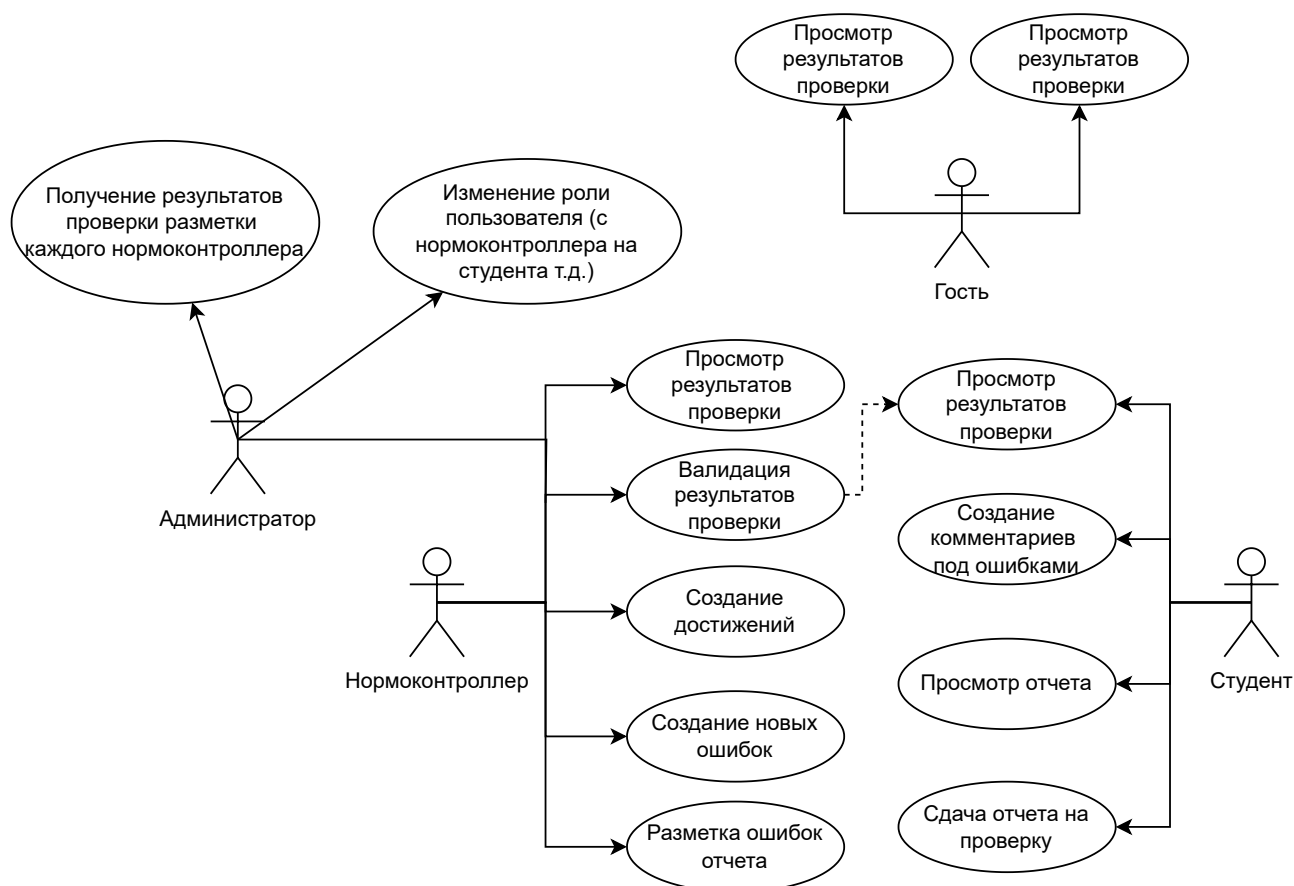


Рисунок 1.5 – Диаграмма прецедентов

## 1.7 Анализ существующих баз данных

База данных — это некоторый набор перманентных (постоянно хранимых) данных, используемых прикладными программными системами какого-либо предприятия [4].

Между физической базой данных (т.е. данными, которые реально хранятся на компьютере) и пользователями системы располагается уровень программного обеспечения, который можно называть по-разному: диспетчер базы данных (database manager), сервер базы данных (database server) или, что более привычно, система управления базами данных, СУБД (DataBase Management System — DBMS). Основная задача СУБД — дать пользователю базы данных возможность работать с ней, не вникая во все подробности работы на уровне аппаратного обеспечения [4].

Модель данных — это абстрактное, самодостаточное, логическое определение объектов, операторов и прочих элементов, в совокупности составляющих

абстрактную машину доступа к данным, с которой взаимодействует пользователь [4]. Рассмотрим классификацию баз данных по различным моделям данных:

1. дореляционные;
2. реляционные;
3. постреляционные.

### **1.7.1 Дореляционные модели**

Дореляционные системы можно разделить на три большие категории:

1. системы с инвертированными списками;
2. иерархические;
3. сетевые [4].

Иерархическая база данных представляется множеством деревьев: в вершинах дерева помещаются записи, состоящие из поименованных полей и представляющие экземпляры некоторого объекта предметной области. Записи связаны строгими иерархическими отношениями — у записи-«потомка» не должно быть более одной записи-«предка» [5].

Сетевая модель данных представляет собой логическую модель данных, которая расширяет иерархический подход. В иерархических структурах каждая запись-потомок имеет ровно одного предка, в то время как в сетевой структуре данных потомок может иметь несколько предков [5].

Инвертированный список в общем случае — это двухуровневая индексная структура. На первом уровне находится файл или часть файла, в которой упорядоченно расположены значения вторичных ключей. Каждая запись с вторичным ключом имеет ссылку на номер первого блока в цепочке блоков, содержащих номера записей с данным значением вторичного ключа. На втором уровне находится цепочка блоков, содержащих номера записей, содержащих одно и то же значение вторичного ключа. При этом блоки второго уровня упорядочены по значениям вторичного ключа, на третьем уровне находится основной файл [6].

Механизм доступа к записям по вторичному ключу при подобной организации записей весьма прост. На первом шаге происходит поиск в области первого уровня заданное значение вторичного ключа, а затем по ссылке считываются блоки второго уровня, содержащие номера записей с заданным значением вторичного ключа, а далее уже прямым доступом загружается в рабочую область пользователя содержимое всех записей, содержащих заданное значение вторичного ключа [6].

### **1.7.2 Реляционные модели**

Реляционная база данных — это такая база данных, которая воспринимается ее пользователями как множество переменных (т.е. переменных отношения), значениями которых являются отношения или, менее формально, таблицы [4].

Реляционная система, поддерживает реляционные базы данных и осуществляет операции над ними, включая RESTRICT (также известную как выборка или англ. SELECT), проекцию (англ. PROJECT) и соединение (англ. JOIN). Эти и подобные операции реляционной алгебры выполняются на уровне множеств [4].

### **1.7.3 Постреляционные модели**

Современный (постреляционный) этап развития связан с использованием объектно-ориентированных технологий разработки программных систем и созданием СУБД нового поколения, унаследовавших все лучшее от дореляционных и реляционных систем. Постреляционные СУБД поддерживают объектные и объектно-реляционные модели данных и обеспечивают разработчикам возможность использовать объектно-ориентированные языки программирования, что дает таким системам технологические преимущества по сравнению с реляционными СУБД [5].

### **1.7.4 Хранение временных рядов**

Отдельно стоит отметить набирающие популярность СУБД временных рядов. Временные ряды (англ. Time series) — это данные, претерпевающие некоторые изменения с течением времени, и фиксируемые в конкретные промежутки времени. Данные СУБД оптимизированы под частую запись данных

и скорость получения доступа к данным не настолько сильно зависит от числа хранимых данных, что характерно для реляционных баз данных, скорость работы которых уменьшается ввиду необходимости индексирования новых элементов. Однако во временных СУБД отсутствует механизм изменения записанных значений, так как считается что записанные метрики (данные) являются фактом в прошлом [7].

## **Вывод**

В данном разделе были описаны основные ошибки студентов в отчетах по лабораторным работам, были выделены участники процесса приема лабораторных работ, формализован процесс приема лабораторных работ, а также рассмотрены существующие средства автоматизации проверки работ на соответствие стандартам.

## 2 Конструкторская часть

В данной части работы будет описана реализация базы данных, описана ER диаграмма БД и принципиальная схема БД, также будут описаны схемы триггеров и хранимые процедуры.

### 2.1 Формализация взаимодействия с приложением

Системе автоматической проверки необходимо проверить на правильность составные части отчета, что подразумевает детекцию рисунков, графиков, схем алгоритмов, списка используемых источников, а также формул.

На рисунках 2.1–2.2 представлены результаты разработки системы детекции составных частей.

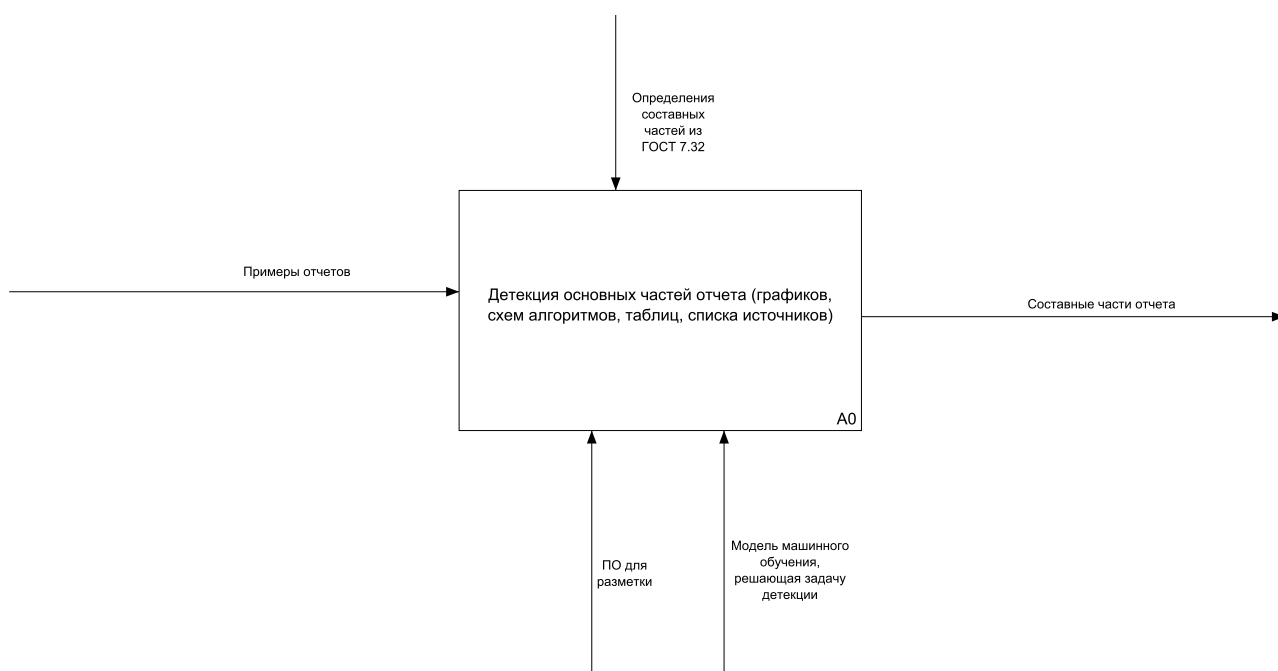


Рисунок 2.1 – IDEF0 обнаружение составных частей отчет 0 уровня



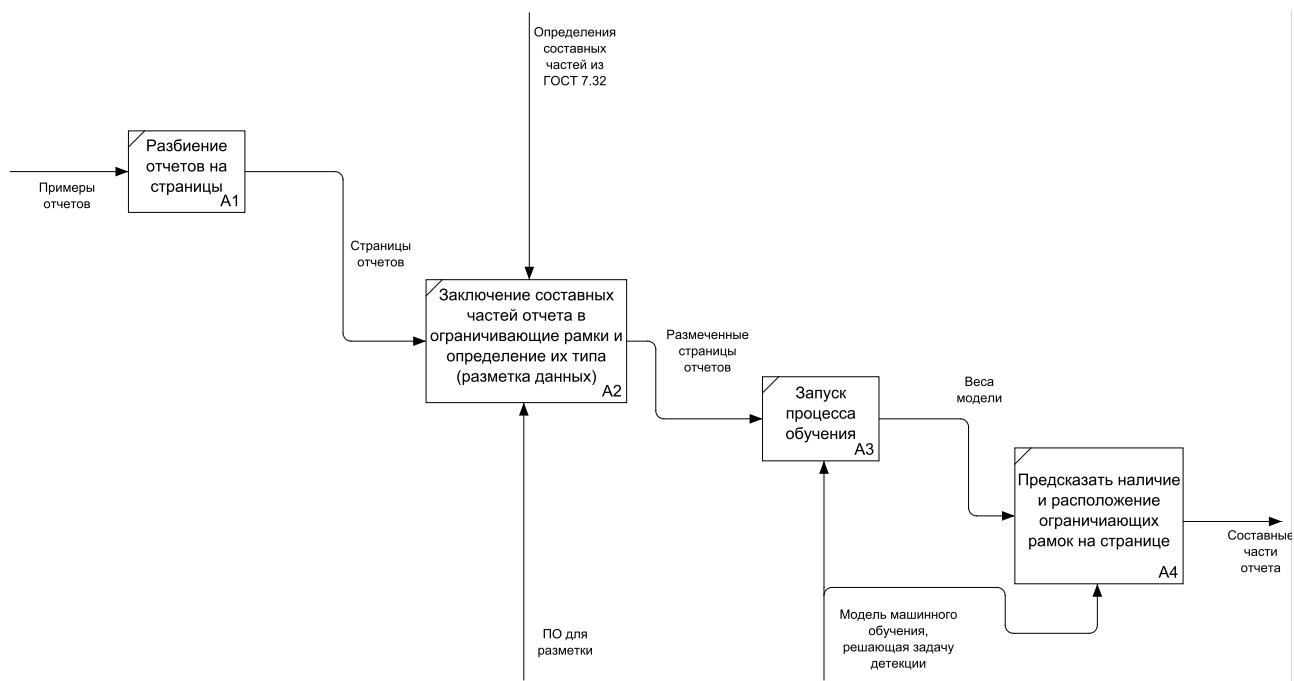


Рисунок 2.2 – IDEF0 обнаружение составных частей отчета 1 уровня

После решения задачи детекции необходимо проверить составные части на соответствие ГОСТ 7.32, однако, финальный вердикт должен выноситься экспертом. На рисунках 2.3–2.4 представлены результаты разработки данной системы.

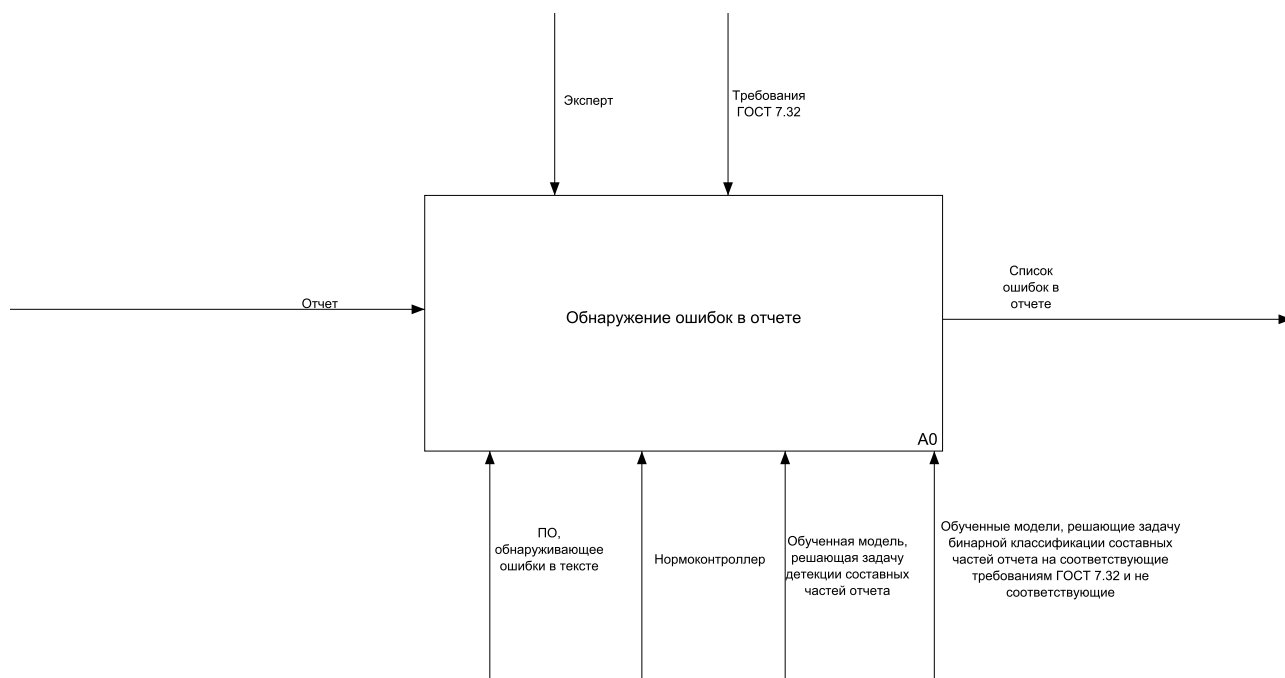


Рисунок 2.3 – IDEF0 проверки составных частей отчета на соответствие ГОСТ 7.32 0 уровня

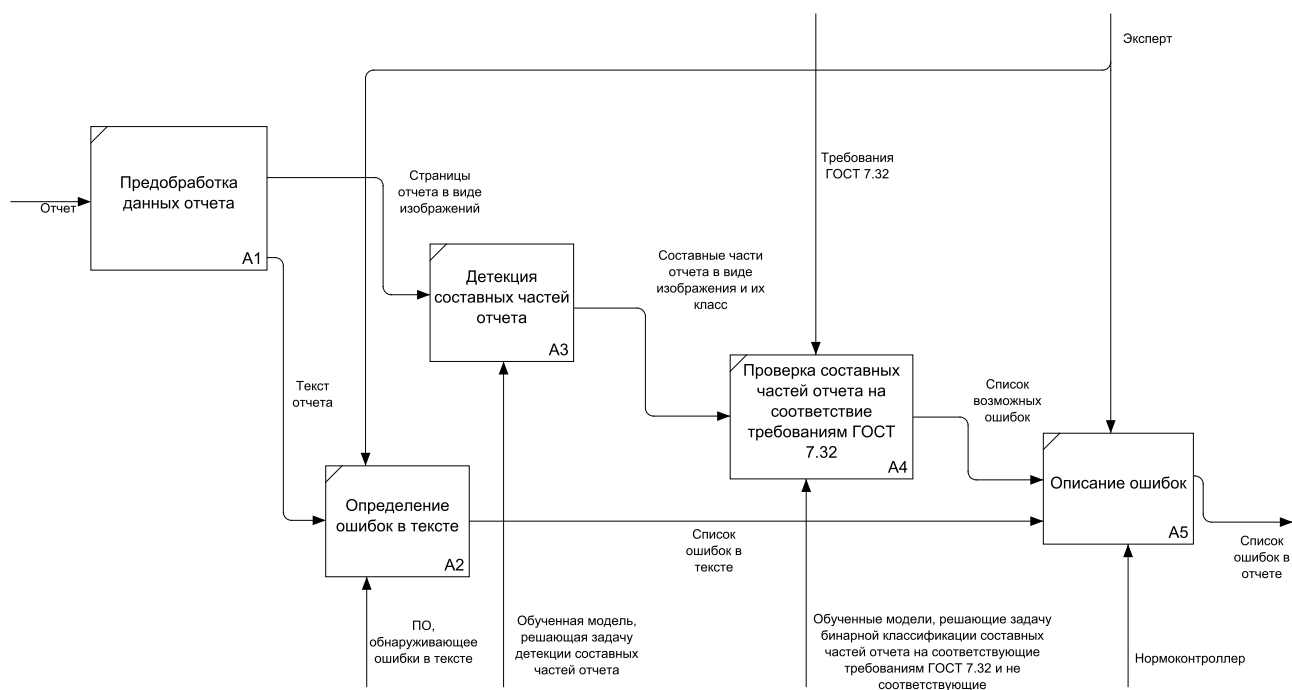


Рисунок 2.4 – IDEF0 проверки составных частей отчета на соответствие ГОСТ 7.32 1 уровня

## 2.2 Формализация работы системы проверки отчета

На рисунке ??, представлена схема BPMN 2.0 работы системы в процессе проверки отчета, а также всех взаимодействующих с ней акторов.

## 2.3 Формализация сущностей базы данных

В соответствие с диаграммами 2.3–2.4 были выделены следующие таблицы:

1. студент — сущность студента (имеет уникальный идентификатор для связи со сданными отчетами и комментариями);
2. нормоконтроллер — сущность нормоконтроллера (имеет уникальный идентификатор для связи с проверенными отчетами и созданными ошибками);
3. выделенный фрагмент — сущность ошибок в отчетах студентов (хранит страницу отчета с ошибкой, а также данные, указывающие на местоположение ошибки, а также тип фрагмента);
4. тип фрагмента — сущность типов ошибок (необходима для введения новых типов ошибок, хранит описание ошибки);
5. отчет — сущность отчета студента (хранит все страницы попытки сдачи отчета, а также время сдачи и номер попытки сдачи);
6. комментарий — сущность комментария студента о ошибке (в случае ошибки преподавателя можно указать на это, кроме данных комментария хранит идентификатор студента, создавшего комментарий);
7. достижение — сущность награды для выдачи студентам, преуспевающим в выполнении лабораторных работ (хранит идентификатор студента, которому выдали награду, а также идентификатор отчета, за который было получено достижение).

## 2.4 Ограничения целостности базы данных

В данной части работы будут описаны описаны поля сущностей баз данных и их ограничения, для обеспечения целостности данных системы.

Ограничение «первичный ключ» подразумевает то что поле сущности является первичным ключом данного отношения в реляционной модели данных.

Таблица 2.1 – Ограничение целостности сущности отчета

Поле	Тип	Ограничение
id	целое число	первичный ключ
page_count	целое число	не пустое значение
document_name	строка	не пустое значение
checks_count	целое число	не пустое значение
creator_id	целое число	не пустое значение
creation_time	дата и время	не пустое значение

Таблица 2.2 – Ограничение целостности сущности фрагмента

Поле	Тип	Ограничение
id	целое число	первичный ключ
description	строка	не пустое значение
creator_id	целое число	не пустое значение
class_name	строка	не пустое значение

Таблица 2.3 – Ограничение целостности сущности типа фрагмента

Поле	Тип	Ограничение
id	целое число	первичный ключ
page_data	байты изображения	не пустое значение
error_bb	массив из 4 вещественных чисел	не пустое значение
class_label	целое число	не пустое значение
creator_id	целое число	не пустое значение

## 2.5 Ролевая модель

Была определена следующая ролевая модель:

1. Разметчик — имеет права доступа SELECT к таблице документов и INSERT к таблице выделенных фрагментов.
2. Добавляющий — имеет права доступа INSERT к таблице студентов и нормоконтроллеров

Таблица 2.4 – Ограничение целостности сущности пользователя

Поле	Тип	Ограничение
id	целое число	первичный ключ
login	строка	уникальный, не пустое значение
password	строка	не пустое значение
name	строка	не пустое значение
surname	строка	не пустое значение
role_id	целое число	не пустое значение
role_type	строка	не пустое значение
group	строка	не пустое значение

Таблица 2.5 – Ограничение целостности сущности достижения

Поле	Тип	Ограничение
id	целое число	первичный ключ
picture	байты изображения	не пустое значение
text	строка	не пустое значение
creator_id	целое число	не пустое значение
granted_to_id	целое число	не пустое значение

Таблица 2.6 – Ограничение целостности сущности комментария

Поле	Тип	Ограничение
id	целое число	первичный ключ
description	строка	не пустое значение
creator_id	целое число	не пустое значение
report_id	целое число	не пустое значение

3. Пользователь — имеет права доступа INSERT к таблице комментариев, имеет права доступа SELECT и INSERT к таблице выделенных фрагментов.
4. Нормоконтроллер — является разметчиков, также имеет права доступа INSERT и SELECT к таблице тип\_ошибки и таблице достижений.
5. Администратор имеет все права доступа ко всем таблицам.
6. Также выделяется работник очереди задача которого — работать с отчетами из очереди и помечать их как проверенные, имеет права доступа SELECT и UPDATE к очереди документов.
7. При проверке документов из очереди система работает в роли работника очереди, в иных случаях система использует роль администратора.

## **2.6 Определение модели базы данных**

Необходимо определить модель базы данных, которая будет использоваться для хранения информации в системе. Реляционная модель баз данных позволяет хранить сущности в целостном виде и поддерживает сложные запросы, поэтому она является наилучшим выбором для поставленной задачи. Также необходимо хранить файлы журнала приложения, ввиду частого его пополнения и необходимости частой записи стоит использовать базу данных временных рядов.

## **2.7 Используемые триггеры**

В базе данных присутствует триггер, запускающий процесс обнаружения ошибок в отчете при добавлении отчета в базу данных. При вставке экземпляра отчета в базу данных, в случае, если данный отчет не был до этого проверен и не имеет в имени файла VIP (такие файлы обслуживаются вне очереди), он будет добавлен в очередь отчетов. Очередь отчетов представляет собой таблицу, в которой хранится ID отчета и его статус (не обработан, обработан, обработан с ошибкой).

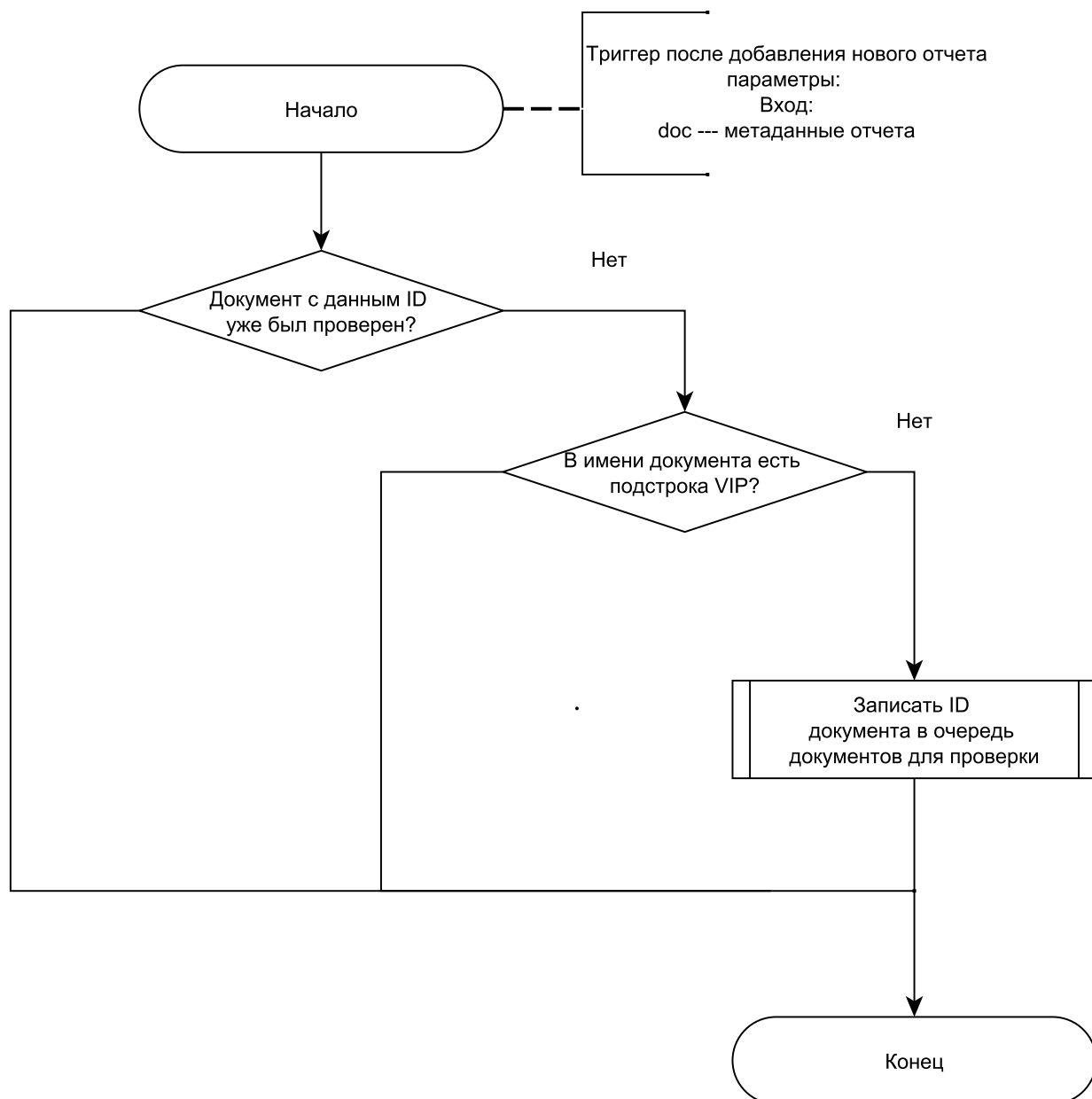


Рисунок 2.6 – Схема алгоритма триггера после добавления в таблицу метаданных отчетов

## Вывод

В данной части были формализованы процессы автоматической проверки, а также выделены различные виды пользователей: нормоконоктроллер, студент, система и администратор, после чего были определены права для каждого из пользователей.

## 3 Технологическая часть

В данной части рассматривается выбор средств реализации, описывается реализация алгоритмов и построенных схем.

### 3.1 Выбор СУБД

В соответствии с приведенной диаграммой сущность—связь необходимо организовать хранение следующих типов данных.

1. Файлы отчетов и результатов проверки работ студентов.
2. Выделенные фрагменты и их описание.
3. Информацию о студентах, достижениях, комментариях.
4. Хранение файлов журнала системы (англ. logs) [8].

Также в выбранной СУБД должна быть возможность создания ролей и выдача прав в соответствии с ними. В конструкторском разделе была определена реляционная модель баз данных для хранения информации в системе, файлы отчетов и результатов проверки будут храниться в файловой системе сервера, мета информация для их получения будет также храниться в базе данных системы.

Рассмотрим наиболее популярные реляционные СУБД [9]:

1. PostgreSQL [10; 11].
2. MySQL [12].
3. MS SQL [13].

Сравнение будет произведено по следующим критериям:

1. Доступность в лицензии Российской Федерации.
2. Популярность среди разработчиков.
3. Наличие ролевой модели.



### 3.1.1 Доступность лицензии

PostgreSQL является ПО с открытым исходным кодом, ввиду чего существуют несколько версий данного продукта [10]. На базе PostgreSQL была разработана СУБД Postgres Pro Certified, которая сертифицирована ФСТЭК РФ [11]. Сертификация позволяет использовать ее для хранения информации:

1. В значимых объектах критической информационной инфраструктуры 1 категории, в государственных информационных системах 1 класса защищенности;
2. В автоматизированных системах управления производственными и технологическими процессами 1 класса защищенности;
3. В информационных системах персональных данных при необходимости обеспечения 1 уровня защищенности персональных данных;
4. В информационных системах общего пользования II класса.

MS SQL (Microsoft SQL Server) является продуктом компании Microsoft, в 2022 году ФСТЭК принял решение об отзыве сертификатов данной компании [14].

MySQL не является сертифицированным продуктом на территории Российской Федерации.

Все 3 СУБД имеют ролевою модель данных [10; 15; 16]. Стоит отметить что наибольший процент использования СУБД имеет MS SQL, затем идет PostgreSQL, затем MySQL [9].

В результате анализа полученных данных, была получена таблица 3.1. В Таблица 3.1 – Сравнение наиболее популярных СУБД

Критерии сравнения \ СУБД	СУБД		
	PostgreSQL	MySQL	MS SQL
Уровень защищенности	1	нет	нет
Частота практического использования	2	3	1
Наличие ролевой модели	да	да	да

таблице 3.1 числа в строке "Частота практического использования" обозначают

порядковый номер СУБД с точки зрения практического внедрения или использования [9].

Таким образом для решения поставленной задачи стоит выбрать СУБД PostgreSQL, так как занимает второе место по популярности и имеет сертификацию 1-го уровня защищенности(см. таблицу 3.1).

Также необходимо реализовать хранения файлов журнала приложения, для этого было решено использовать СУБД временных рядов. Так как InfluxDB поддерживает использование InfluxQL подобного SQL, а также является ПО с открытым исходным кодом стоит использовать ее для хранения файлов журнала [7].

## 3.2 Выбор средств реализации

Для реализации описанного функционала был выбран язык программирования GO, в силу следующих причин:

1. Существует библиотека для данного языка, позволяющая реализовывать запросы к базам данных на уровне языка [17].
2. Язык поддерживает создание интерфейсов и структур, что позволяет выделить основные объекты и сущности при разработке и использовать их при реализации [18].
3. GO имеет обширную документацию, в которой в каждом описываемом модуле (пакете) приведены примеры реализации и использования объектов модуля [19].

Для реализации нейронной сети был выбран язык Python, так как он поддерживает фреймворк PyTorch, использование которого достаточно для решения задачи детекции основных частей отчета [20].

## 3.3 Реализация создания отношений

В листинге В.2 приведена реализация создания отношений. Отношения обозначают соответственно:

1. documents — сущность документа;
2. markup\_types — сущность типов фрагмента отчета;

3. `markups` — сущность выделенного фрагмента;
4. `users` — сущность пользователя (контроллера, админа или студента);
5. `document_queues` — сущность элемента очереди отчета для проверки;
6. `comments` — сущность комментариев к отчету;
7. `achievements` — сущность достижения.

### 3.4 Реализация функций

В листинге В.3 приведены реализации спроектированных функций. Функция `getTasksReady` получает все документы из очереди, которые готовы к проверке и не были проверены до этого.

### 3.5 Реализация триггеров

В листинге В.4 приведены реализации спроектированных триггеров. Триггер `addToQueue`, при добавлении метаданных документа, в случае если в имени файла нет подстрока `VIP` и он не был уже проверен добавляет документ в очередь для проверки.

### 3.6 Описание приложения

## 4 Исследовательская часть

В данном разделе будет описано исследование зависимости среднего числа генерируемых кадров от числа и типа примитивов на сцене. Также будет описаны технические характеристики устройства, на котором проводились замеры и приведен анализ полученных результатов.

## ЗАКЛЮЧЕНИЕ

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. ВКР ВУЗ [Электронный ресурс]. — URL: <http://www.vkr-vuz.ru/> (дата обращения: 27.11.2023).
2. TestVkr [Электронный ресурс]. — URL: <https://vkr.bmstu.ru/> (дата обращения: 27.11.2023).
3. PdfTest [Электронный ресурс]. — URL: <https://applitools.com/blog/automate-pdf-testing/> (дата обращения: 27.11.2023).
4. Дж. Д. К. Введение в системы баз данных: 8-е издание //. — — «Вильямс», 2006. — С. 1328.
5. К. В. В. Базы данных, проектирование, программирование, управление и администрирование //. — — Издательство «Лань», 2020.
6. Физические модели баз данных [Электронный ресурс]. — URL: <https://intuit.ru/studies/courses/1001/297/lecture/7415?page=5&ysclid=lur73sp2bk237864836> (дата обращения: 08.04.2024).
7. Обзор баз данных временных рядов. — URL: <https://cyberleninka.ru/article/n/obzor-baz-dannyh-vremennyh-ryadov> (дата обращения: 09.05.2024).
8. Д. К. С. Основы современных баз данных //. — — Центр Информационных Технологий, 1998.
9. Анализ популярных реляционных систем управления базами данных (2022 г.) [Электронный ресурс]. — URL: <https://drach.pro/blog/hitech/item/196-popular-relational-dbms-2022> (дата обращения: 09.05.2024).
10. PostgreSQL [Электронный ресурс]. — URL: <https://www.postgresql.org/> (дата обращения: 09.05.2024).
11. Postgres Pro Certified [Электронный ресурс]. — URL: <https://postgrespro.ru/products/postgrespro/certified> (дата обращения: 09.05.2024).
12. SQL Server 2022 [Электронный ресурс]. — URL: <https://www.microsoft.com/ru-ru/sql-server/sql-server-2022> (дата обращения: 09.05.2024).

13. MySQL [Электронный ресурс]. — URL: <https://www.mysql.com/> (дата обращения: 09.05.2024).
14. ФСТЭК России приостановила действие 56 сертификатов [Электронный ресурс]. — URL: <https://www.anti-malware.ru/news/2022-03-24-111332/38397> (дата обращения: 09.05.2024).
15. Роли уровня сервера [Электронный ресурс]. — URL: <https://learn.microsoft.com/ru-ru/sql/relational-databases/security/authentication-access/server-level-roles?view=sql-server-ver16> (дата обращения: 09.05.2024).
16. MySQL Roles – All You Need to Know [Электронный ресурс]. — URL: <https://mysqlcode.com/mysql-roles/> (дата обращения: 09.05.2024).
17. gorm [Электронный ресурс]. — URL: <https://gorm.io/> (дата обращения: 09.05.2024).
18. Interface in Go [Электронный ресурс]. — URL: <https://golangbyexample.com/interface-in-golang/> (дата обращения: 09.05.2024).
19. Discover Packages [Электронный ресурс]. — URL: <https://pkg.go.dev/> (дата обращения: 09.05.2024).
20. PyTorch [Электронный ресурс]. — URL: <https://pytorch.org/> (дата обращения: 09.05.2024).
21. ГОСТ 7.32—2017 Система стандартов по информации, библиотечному и издательскому делу. Отчет о научно-исследовательской работе. Структура и правила оформления. — М.: Стандартинформ, 2017. — 35 с.

# ПРИЛОЖЕНИЕ А

## А.1 Основные ошибки в отчетах

В данном разделе будут рассмотрены наиболее часто встречающиеся ошибки, которые совершают студенты при написании различных отчетов.

В целях выявления наиболее часто встречающихся ошибок были опрошены преподаватели, работа которых непосредственно связана с проверкой отчетов студентов.

### А.1.1 Общие ошибки

В ГОСТ 7.32 указаны следующие размеры полей: левое — 30 мм, правое — 15 мм, верхнее и нижнее — 20 мм [21]. Выход за границы листа является одной из самых распространенных ошибок.

Каждый объект (таблица, рисунок, схема алгоритма, формула) должен быть подписан и пронумерован, однако более подробно подписи к каждому из них будут рассмотрены в следующих подразделах.

Если таблицу или схему не удастся разместить на одной странице, то следует разбить данный объект на несколько частей, каждая из которых должна быть подписана.

### А.1.2 Ошибки в тексте

Слова в тексте должны быть согласованы в роде, числе и падеже.

Страницы отчета должны быть пронумерованы, однако, номер на титульном листе не ставится, но он является первой страницей, что означает, что следующая страница должна иметь номер 2.

Ненумерованный заголовок (введение, список литературы, оглавление и т. п.) должен быть выровнен по центру, при этом он состоит только из прописных букв (пример представлен в приложении Б.1), другие варианты оформления являются не соответствующими стандарту.

Абзацный отступ должен быть одинаковым по всему тексту отчета и равен 1,25 см [21]. Любые другие варианты оформления считаются ошибочными.

Возможна потеря научного стиля и переход к публицистике, что является



ошибкой, текст работы должен быть написан на государственном языке в научном стиле.

### **А.1.3 Ошибки в рисунках**

Частой ошибкой является неправильное оформление рисунков. Каждый рисунок должен быть подписан, при этом подпись должна располагаться строго по центру, внизу рисунка. Другое оформление считается ошибочным.

Использование рисунков низкого разрешения является ошибкой. Все рисунки должны быть выполнены в высоком качестве, если обратное не требуется в самой работе.

Некорректный поворот рисунка считается ошибкой. Если рисунок не удастся разместить на странице, то допускается повернуть его таким образом, чтобы верх рисунка был ближе к левой части страницы (см. рисунок Б.2).

### **Ошибки в графиках**

Для каждого графика должна существовать легенда, для оформления которой существует два варианта:

- в одном из углов графика находится область, в которой указаны все обозначения;
- в подписи к графику описано каждое обозначение;

другое оформление является ошибкой.

Часто на графиках отсутствуют единицы измерения, что является ошибкой. Должны быть подписаны единицы измерения каждой из осей графика, даже в том случае, если на графике оси подписываются словами, например, если измерение идет в штуках или на оси обозначены времена года (см. рисунок Б.3).

Отчеты могут быть напечатаны в черно-белом варианте, поэтому на графиках должны быть маркеры, которые позволяют отличить графики друг от друга даже не в цветном варианте. Отсутствие маркеров считается ошибкой.

При большом количестве графиков на одном рисунке возможна ситуация, при которой невозможно отличить один график от другого, что является ошибкой.

## Ошибки в схемах алгоритмов

Если схему не удастся разместить на одной странице, то она разбивается на несколько частей, каждая из которых должна быть подписана. Для разделения схемы алгоритма на части используется специальный символ-соединитель, который отображает выход в часть схемы и вход из другой части этой схемы, соответствующие символы-соединители должны содержать одно и то же уникальное обозначение, любые другие варианты оформления являются ошибочными.

Часто вместо символа начала или конца алгоритма используют овал, однако в этом случае должен быть использован прямоугольник с закругленными углами (см. рисунок Б.4).

При использовании символа процесса (прямоугольник) часто используют прямоугольник с закругленными углами (см. рисунок Б.5), что является ошибкой.

При соединении символов схемы алгоритмов не нужны стрелки, если они соединяют символы в направлении слево-направо или сверху-вниз, в остальных случаях символы должны соединяться линиями со стрелкой на конце, отсутствие требуемых стрелок считается ошибкой.

При использовании символа процесса-решение как минимум одна из соединительных линий должна быть подписана (см. рисунок Б.6), однако возможен также вариант, когда подписаны обе линии. Отсутствие пояснений к выходам данного символа является ошибкой.

Часто пояснительный текст пересекается с символами, использующимися для составления схем, что является ошибкой.

### А.1.4 Ошибки в таблицах

Каждая таблица должна быть подписана. Наименование следует помещать над таблицей слева, без абзачного отступа в следующем формате: Таблица Номер таблицы - Наименование таблицы. Наименование таблицы приводят с прописной буквы без точки в конце [21]. Другие варианты оформления считаются не соответствующими стандарту.

Таблицу с большим количеством строк допускается переносить на другую страницу. При переносе части таблицы на другую страницу слово «Таб-

лица», ее номер и наименование указывают один раз слева над первой частью таблицы, а над другими частями также слева пишут слова «Продолжение таблицы» и указывают номер таблицы [21]. Любое другое оформление считается ошибочным.

### **А.1.5 Ошибки в формулах**

Каждая формула должна быть пронумерована вне зависимости от того, существует ли ссылка на нее. Нумерация может осуществляться в двух вариантах:

- сквозная нумерация (номер формулы не зависит от раздела, в котором она находится);
- нумерация, зависящая от раздела (в том случае номер формулы начинается с номера раздела);

другое оформление считается ошибкой.

Отсутствие знака препинания после формулы является ошибкой. После каждой формулы должен находиться знак препинания (точка, запятая и т. п.), зависящий от контекста. Если в формуле содержится система уравнений, то после каждого из них (за исключением последнего) ставится запятая, а после последнего — точка, либо запятая (см. рисунок Б.7).

Номер формулы должен быть выравнен по правому краю страницы и находиться по центру формулы (в вертикальной плоскости). Другое оформление нумерации формул считается не соответствующим стандарту.

Если формула вставляется в начале страницы, то часто перед ней может присутствовать отступ, которого быть не должно.

### **А.1.6 Ошибки в списках**

Ненумерованные списки должны начинаться с удлиненного тире (см. рисунок Б.8), другое оформление является ошибочным.

В нумерованных списках после номера пункта обязательно должна стоять скобка (см. рисунок Б.10), использование другого знака считается ошибкой.

В конце каждого пункта списка должен быть знак препинания, от которого зависит первая буква первого слова следующего пункта (см. рисунок Б.9):

- если пункт заканчивается на точку, то первое слово следующего пункта должно начинаться на прописную букву;
- если пункт заканчивается запятой или точкой с запятой, то следующий первое слово следующего слова должно начинаться со строчной буквы;

другое оформление является ошибочным.

### **А.1.7 Ошибки в списке литературы**

Часто при описании одного из источников не указывается одна из составных частей (автор, издательство и т. п.), что является ошибкой.

Также нередко встречаются ссылки на так называемые «препринтовские» издательства (статья еще не вышла), однако была использована в отчете, это считается ошибкой.

## ПРИЛОЖЕНИЕ Б

### Введение

Рисунок Б.1 – Пример ошибочного оформления нумерованного заголовка

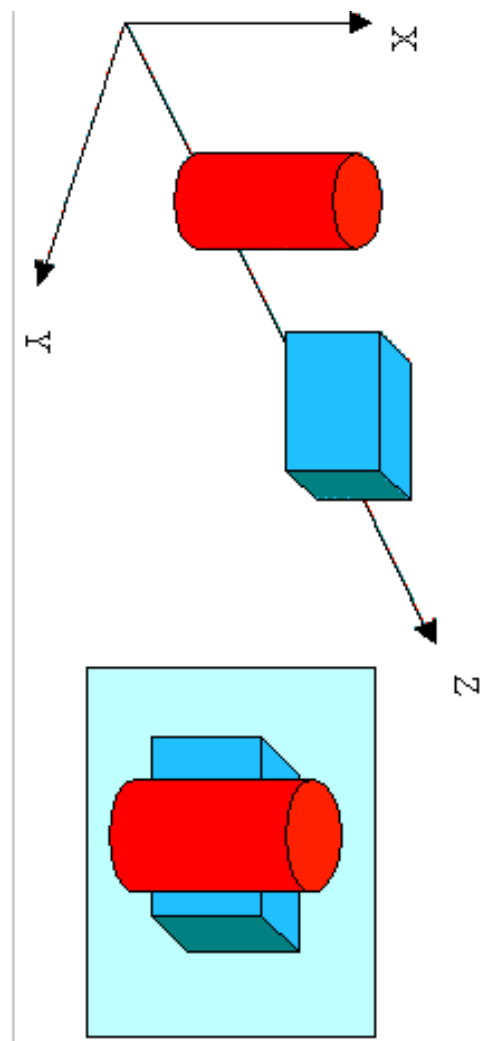


Рисунок 1 - Пример работы Z-буфера

Рисунок Б.2 – Пример ошибочного оформления рисунка — некорректный поворот

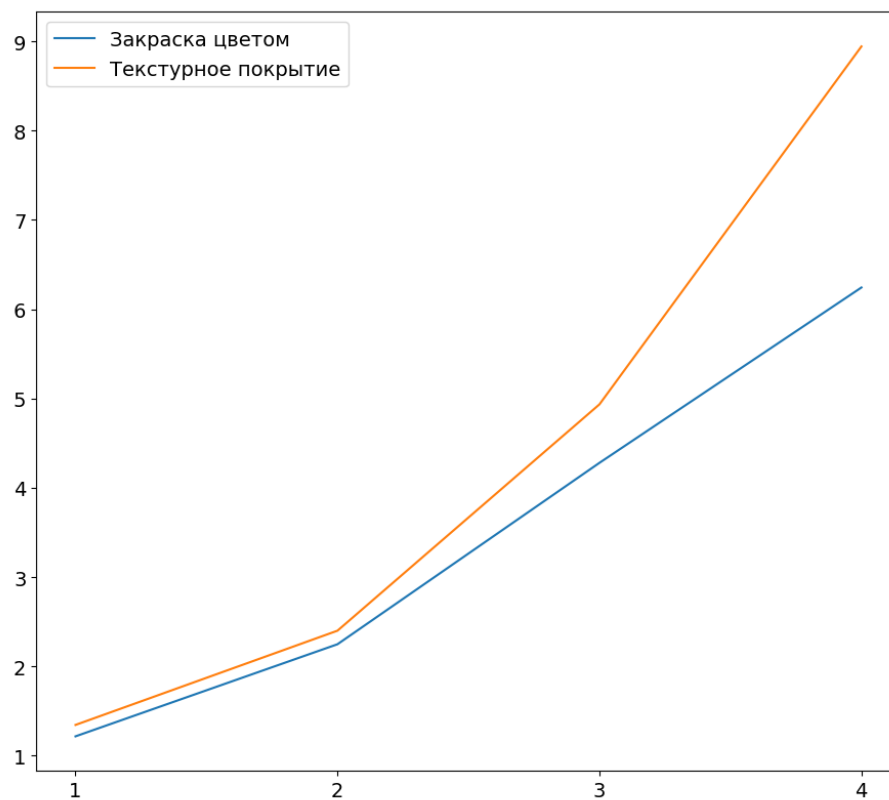


Рисунок Б.3 – Пример ошибочного оформления графика — отсутствуют единицы измерения

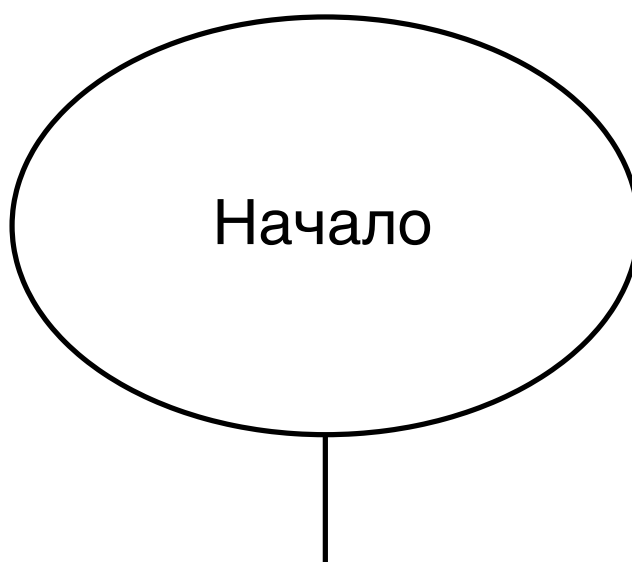


Рисунок Б.4 – Пример ошибочного оформления схемы — некорректный символ начала

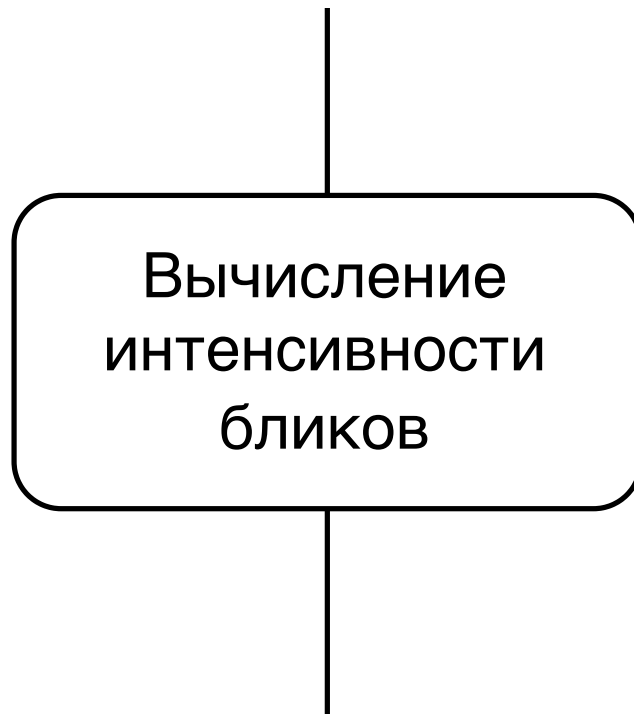


Рисунок Б.5 – Пример ошибочного оформления схемы — некорректный символ процесса

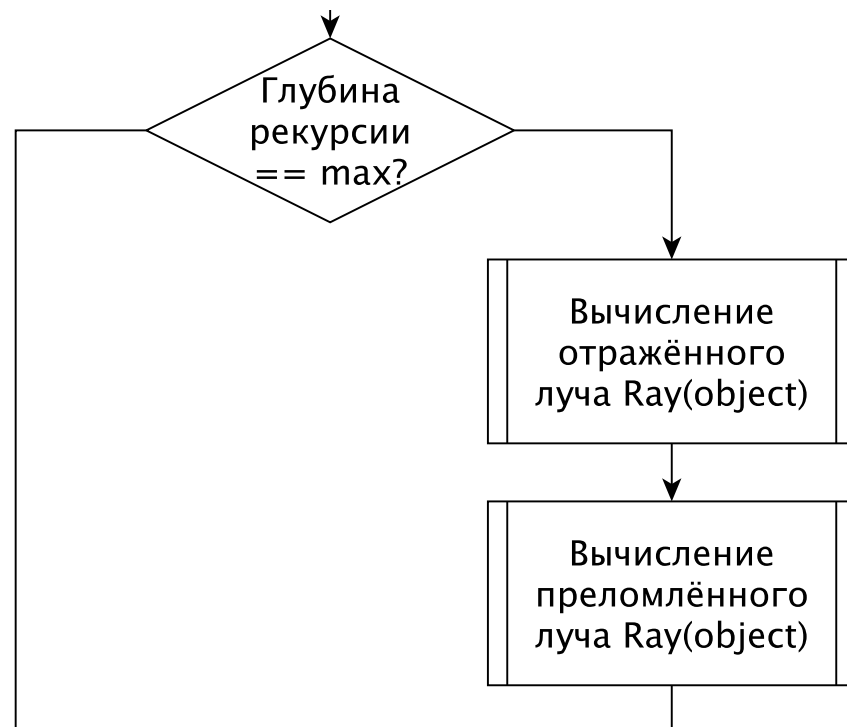


Рисунок Б.6 – Пример ошибочного оформления схемы — не подписана ни одна из веток символа процесса—решение

$$D(i, j) = \begin{cases} 0 & \text{если } i = 0, j = 0 \\ j & \text{если } i = 0, j > 0 \\ i & \text{если } j = 0, i > 0 \\ \min(\min(D(i, j - 1) + 1, \\ D(i - 1, j) + 1) \\ D(i - 1, j - 1) + m(S_1[i], S_2[j])) & \left[ \begin{array}{l} D(i - 2, j - 2) + 1 \quad \text{если } i > 1, j > 1, \\ S_1[i - 1] == S_2[j - 2], \\ S_1[i - 2] == S_2[j - 1] \end{array} \right] \text{ иначе} \end{cases} \quad (1)$$

Рисунок Б.7 – Пример ошибочного оформления системы уравнений — отсутствуют знаки препинания после уравнений

- One
- Two
- Three

Рисунок Б.8 – Пример ошибочного оформления нумерованного списка — некорректный символ перед элементами списка

- Первый,
- Второй,
- Третий.

Рисунок Б.9 – Пример ошибочного оформления нумерованного списка — некорректный регистр буквы следующего пункта после запятой в предыдущем

1. первый,
2. второй,
3. третий.

Рисунок Б.10 – Пример ошибочного оформления нумерованного списка — некорректный символ после номера элемента списка



## ПРИЛОЖЕНИЕ В

Листинг В.1 – Реализация ролевой модели

```
1  --Создание админа
2  CREATE ROLE admin WITH
3  SUPERUSER
4  NOCREATEDB
5  CREATEROLE
6  NOINHERIT
7  NOREPLICATION
8  NOBYPASSRLS
9  CONNECTION LIMIT -1
10 LOGIN
11 PASSWORD 'admin';
12
13 GRANT ALL PRIVILEGES
14 ON ALL TABLES IN SCHEMA public TO admin;
15
16 --Создание добавляющего
17 CREATE ROLE adder WITH
18 NOSUPERUSER
19 NOCREATEDB
20 NOCREATEROLE
21 NOINHERIT
22 NOREPLICATION
23 NOBYPASSRLS
24 CONNECTION LIMIT -1
25 LOGIN
26 PASSWORD 'adder';
27
28 GRANT INSERT
29 ON public.users
30 TO adder;
31
32
33 --Создание разметчика
34 CREATE ROLE annoter WITH
35 NOSUPERUSER
36 NOCREATEDB
37 NOCREATEROLE
```

```

38 NOINHERIT
39 NOREPLICATION
40 NOBYPASSRLS
41 CONNECTION LIMIT -1
42 LOGIN
43 PASSWORD 'anoter';
44
45 GRANT INSERT
46 ON public.markups TO anoter;
47
48 GRANT SELECT
49 ON public.documents
50 TO anoter;
51
52
53
54
55 --Создание пользователя (студента)
56 CREATE ROLE student WITH
57 NOSUPERUSER
58 NOCREATEDB
59 NOCREATEROLE
60 NOINHERIT
61 NOREPLICATION
62 NOBYPASSRLS
63 CONNECTION LIMIT -1
64 LOGIN
65 PASSWORD 'student';
66
67
68
69
70 GRANT INSERT
71 ON public.comments
72 TO student;
73
74 GRANT INSERT,SELECT
75 ON public.markups
76 TO student;
77
78

```

```

79
80 --Создание нормоконтроллера
81 CREATE ROLE controller WITH
82 NOSUPERUSER
83 NOCREATEDB
84 NOCREATEROLE
85 INHERIT
86 NOREPLICATION
87 NOBYPASSRLS
88 CONNECTION LIMIT -1
89 LOGIN
90 PASSWORD 'controller';
91
92
93 GRANT ALL PRIVILEGES
94 ON public.comments
95 TO controller;
96
97 GRANT anoter TO controller;
98
99 GRANT INSERT,SELECT
100 ON public.markup_types
101 TO controller;
102
103 GRANT ALL PRIVILEGES
104 ON public.achievment
105 TO controller;
106
107 --Работник очереди
108 CREATE ROLE queue_worker WITH
109 NOSUPERUSER
110 NOCREATEDB
111 NOCREATEROLE
112 NOINHERIT
113 NOREPLICATION
114 NOBYPASSRLS
115 CONNECTION LIMIT -1
116 LOGIN
117 PASSWORD 'queue_worker';
118
119 GRANT SELECT,UPDATE

```

```
120 ON public.document_queues
121 TO queue_worker;
```

Листинг В.2 – Реализация создания отношений

```
1 CREATE TABLE IF NOT EXISTS documents (
2     id UUID PRIMARY KEY,
3     page_count INT NOT NULL,
4     document_name VARCHAR(255) NOT NULL,
5     checks_count INT NOT NULL,
6     creator_id BIGINT NOT NULL,
7     creation_time TIMESTAMP NOT NULL
8 );
9
10 CREATE TABLE IF NOT EXISTS markup_types (
11     id serial PRIMARY KEY,
12     description VARCHAR(255) NOT NULL,
13     creator_id INT NOT NULL,
14     class_name VARCHAR(255) NOT NULL
15 );
16
17 CREATE TABLE IF NOT EXISTS users (
18     id serial PRIMARY KEY,
19     login VARCHAR(255) UNIQUE NOT NULL,
20     password VARCHAR(255) NOT NULL,
21     name VARCHAR(255),
22     surname VARCHAR(255) NOT NULL,
23     role INT NOT NULL,
24     user_group VARCHAR(255) NOT NULL
25 );
26
27 CREATE TABLE IF NOT EXISTS markups (
28     id serial PRIMARY KEY NOT NULL,
29     page_data BYTEA NOT NULL,
30     error_bb JSONB DEFAULT '[]' NOT NULL,
31     class_label BIGINT NOT NULL,
32     creator_id BIGINT NOT NULL
33 );
34
35
36 CREATE TABLE IF NOT EXISTS document_queues (
37     id serial PRIMARY KEY NOT NULL,
38     doc_id UUID NOT NULL,
```

```

39     status SMALLINT NOT NULL,
40     CONSTRAINT fk_doc_queue_docs FOREIGN KEY (doc_id) REFERENCES
        documents(id) ON DELETE CASCADE
41 );
42
43
44 CREATE TABLE IF NOT EXISTS comments (
45     id serial PRIMARY KEY NOT NULL,
46     doc_id UUID NOT NULL,
47     description text NOT NULL,
48     creator_id INT NOT NULL,
49     CONSTRAINT fk_comment_docs FOREIGN KEY (doc_id) REFERENCES
        documents(id) ON DELETE CASCADE,
50     CONSTRAINT fk_comment_users FOREIGN KEY (creator_id)
        REFERENCES users(id) ON DELETE CASCADE
51 );
52
53 CREATE TABLE IF NOT EXISTS achievements (
54     id serial PRIMARY KEY NOT NULL,
55     description text NOT NULL,
56     creator_id INT NOT NULL,
57     granted_to_id INT NOT NULL,
58     CONSTRAINT fk_comment_creator FOREIGN KEY (creator_id)
        REFERENCES users(id) ON DELETE CASCADE,
59     CONSTRAINT fk_comment_granted FOREIGN KEY (granted_to_id)
        REFERENCES users(id) ON DELETE CASCADE
60 );
61
62
63 ALTER TABLE markups ADD CONSTRAINT fk_markup_markup_type FOREIGN
    KEY ( class_label ) REFERENCES markup_types( id );
64
65 ALTER TABLE documents ADD CONSTRAINT fk_document_user FOREIGN
    KEY ( creator_id ) REFERENCES users( id );
66
67 ALTER TABLE markups ADD CONSTRAINT fk_markup_user FOREIGN KEY (
    creator_id ) REFERENCES users( id );
68
69 ALTER TABLE markup_types ADD CONSTRAINT fk_markup_types_user
    FOREIGN KEY ( creator_id ) REFERENCES users( id );
70

```

```

71 ALTER TABLE document_queues ADD CONSTRAINT fk_doc_queue_docs
    FOREIGN KEY ( doc_id ) REFERENCES documents( id );

```

### Листинг В.3 – Реализация используемых функций

```

1 CREATE OR REPLACE FUNCTION getTasksReady()
2 RETURNS TABLE (
3     id int,
4     doc_id uuid,
5     status SMALLINT
6 )
7 AS $$
8 #variable_conflict use_column
9 BEGIN
10     RETURN QUERY
11     SELECT dq.id, dq.doc_id, dq.status
12     FROM document_queues as dq
13     WHERE status = 0;
14 END;
15 $$ LANGUAGE plpgsql;
16
17 select * from getTasksReady();

```

### Листинг В.4 – Реализация используемых триггеров

```

1
2 CREATE OR REPLACE FUNCTION addToQueue()
3 Returns TRIGGER
4 AS $$
5 BEGIN
6     IF NOT EXISTS (
7         SELECT
8             *
9         FROM
10             document_queues
11         WHERE
12             doc_id = NEW.ID AND status = 2
13     ) AND NEW.document_name NOT LIKE '%VIP%'
14 THEN
15     INSERT INTO
16     document_queues(doc_id,status)
17     VALUES
18     (NEW.ID,0); -- status zero means unchecked
19 END IF;

```

```
20         RETURN NEW;
21 END;
22 $$ LANGUAGE plpgsql;
23
24 CREATE TRIGGER insertQueueTrigger AFTER INSERT ON
    public.documents
25 FOR EACH ROW EXECUTE PROCEDURE addToQueue();
```