



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«МИРЭА – Российский технологический университет»
РТУ МИРЭА

Институт информационных технологий
Кафедра инструментального и прикладного программного обеспечения

Отчет по практическим работам
по дисциплине
«Проектирование информационных систем»

Выполнил: студент группы ИВБО-02-19

К. Ю. Денисов

Принял: ассистент

А. А. Русляков

Работа представлена к защите «___»_____ 202_ г.

Москва 2022

Содержание

1	Формирование требований к системе	4
1.1	РЕФЕРАТ	4
1.2	ОПИСАНИЕ ОБЪЕКТА АВТОМАТИЗАЦИИ	4
1.3	ОБЩИЕ СВЕДЕНИЯ	4
1.3.1	Список требований и определений	4
1.3.2	Описание бизнес-ролей	5
1.4	ТРЕБОВАНИЯ К СИСТЕМЕ	5
1.4.1	Бизнес требования	5
1.4.2	Пользовательские требования	5
1.4.3	Функциональные требования	5
1.4.4	Нефункциональные требования	6
1.4.5	Требования к пользовательскому интерфейсу	6
1.4.6	Требования к защите информации	7
1.4.7	Требования по сохранности информации при авариях	8
1.4.8	Требования по стандартизации и унификации	8
1.4.9	Требования к безопасности	8
1.4.10	Требования к документированию	9
1.4.11	Требования к функциям, выполняемым системой	10
1.5	СЦЕНАРИИ ИСПОЛЬЗОВАНИЯ	10
2	Выбор архитектуры системы	11
2.1	ФУНКЦИОНАЛЬНЫЕ СХЕМЫ СИСТЕМЫ	11
2.1.1	Распределенная архитектура	11
2.1.2	Клиент-серверная архитектура	11
2.2	ВЫБОР АРХИТЕКТУРЫ	13
3	Функциональное проектирование (методология SADT)	14
3.1	ЦЕЛЬ СОЗДАНИЯ СИСТЕМЫ	14
3.2	КРАТКОЕ ОПИСАНИЕ	14
3.3	СРЕДСТВА СОЗДАНИЯ ИС	14
3.4	ПРОЕКТИРОВАНИЕ КОНТЕКСТНОЙ ДИАГРАММЫ	15

4	Функциональное проектирование (нотация IDEF0)	17
4.1	ОПИСАНИЕ ПРОЦЕССОВ ДЕКОМПОЗИЦИИ	17
4.1.1	Процесс регистрации пользователя в системе	18
4.1.2	Процесс публикации файла в системе	18
4.1.3	Процесс получения файла, размещенного в системе .	18
4.1.4	Процесс участия в обсуждении на форуме	19
4.2	ДЕКОМПОЗИЦИЯ ФУНКЦИОНАЛЬНОГО БЛОКА	19
4.2.1	Выбор наименования научной дисциплины	20
4.2.2	Выбор типа научной работ	20
4.2.3	Установка параметров файла	20
4.2.4	Загрузка файла в систему	20
5	Функциональное проектирование (нотация DFD)	21
5.1	ОПРЕДЕЛЕНИЕ ОБЪЕКТА ДЕКОМПОЗИЦИИ	21
5.1.1	Установка критериев поиска	21
5.1.2	Обращение к хранилищу файлов	21
6	Проектирование структуры данных ИС (Entity Relation Diagram)	23
6.1	МОДЕЛИРОВАНИЕ СИСТЕМЫ	23
6.2	КРАТКАЯ ПОСТАНОВКА ЗАДАЧИ	23
6.3	ПОСТРОЕНИЕ ER ДИАГРАММЫ	25
7	СОЗДАНИЕ ДИАГРАММЫ СОСТОЯНИЙ	27
8	Расчет параметров проектируемой информационной системы	28
8.1	ОПИСАНИЕ ЭСЕ	28
8.2	НАПОЛНЕНИЕ СИСТЕМЫ	28
8.3	МАТЕМАТИЧЕСКИЕ РАСЧЕТЫ	28
8.3.1	Расчет математического ожидания информационного блока системы	29
8.3.2	Расчет дисперсии информационного блока системы .	30
8.3.3	Расчет среднеквадратического отклонения	30
8.3.4	Расчет энтропии системы	30
8.4	РЕЗУЛЬТАТЫ РАСЧЕТОВ	31

1 Формирование требований к системе

1.1 РЕФЕРАТ

Суть данной практической работы заключается в анализе и формировании требований к разрабатываемой информационной системе «Электронный сборник лабораторных работ», в том числе требований к составу и содержанию работ по подготовке объекта автоматизации к вводу системы в действие, а также требований к документированию определению и формализации бизнес-ролей.

Данная практическая работа содержит 31 страницы

1.2 ОПИСАНИЕ ОБЪЕКТА АВТОМАТИЗАЦИИ

Разрабатываемая информационная система «Электронный сборник лабораторных работ» (далее по тексту Система) служит для сбора, хранения и учета письменных работ различного характера, выполненных учащимися средних общеобразовательных, средних специальных и высших учебных заведений. На данной платформе пользователям на безвозмездной основе предоставляется доступ к загруженным документам, топикам и темам форума и медиафайлам.

Система выполняет функции образовательной платформы. Платформа может быть интегрирована в информационную среду учебных заведений, предоставляя инструменты для учета и хранения работ учащихся.

1.3 ОБЩИЕ СВЕДЕНИЯ

1.3.1 Список требований и определений

Подсистема управления доступом (ПУД) — часть Системы, назначающая разрешения конечным пользователям в зависимости от их роли в вашей организации. ПУД обеспечивает гранулярный контроль, предлагая простой и управляемый подход к управлению доступом, который менее подвержен ошибкам, чем индивидуальное назначение разрешений.

1.3.2 Описание бизнес-ролей

Таблица 1 — Описание бизнес-ролей

Бизнес-роль	Описание
Администратор системы	Лицо, имеющее доступ к ПУД и другим инструментам администрирования Системы
Пользователь	Лицо или организация, которая использует Систему для размещения и хранения файлов, имеет личный кабинет на портале, может участвовать в обсуждениях на форуме

1.4 ТРЕБОВАНИЯ К СИСТЕМЕ

1.4.1 Бизнес требования

Система должна предоставлять функционал для удобного удаленного хранения и управления электронными документами.

Использование системы в качестве хранилища должно быть более привлекательно, чем хранение файлов на локальном компьютере или физическом носителе.

Система должна иметь возможность быть интегрированной в информационную платформу общеобразовательных учреждений.

1.4.2 Пользовательские требования

Пользователь должен иметь возможность авторизоваться в системе для загрузки документов в хранилище.

Пользователь должен иметь возможность скачивания файлов без ограничений и необходимости регистрации .

Пользователь должен иметь возможность оставлять отзыв и оценку конкретному документу, размещенному в системе.

Пользователь должен иметь возможность создать тему на портале.

1.4.3 Функциональные требования

Система должна отвечать требованиям по хранению, обработке и защите персональных данных в соответствии с 152-ФЗ «О персональных

данных».

Система должна иметь упорядоченную структуру файлов с разделением по языкам, научным дисциплинам, видам работ.

Данные должны храниться в течение 20 лет.

В системе должен сочетаться функционал хранилища данных и образовательного форума.

Платформа должна реализовывать систему управления (авторизации) доступом на основе ролей.

Система должна реализовывать функционал по аутентификации пользователя по логину (почтовому адресу) и паролю.

Персональные данные пользователя и хеш-строки паролей должны храниться в зашифрованном виде.

К загрузке должны допускаться файлы любого формата.

1.4.4 Нефункциональные требования

В качестве алгоритма хеширования должен использоваться SHA-256.

Система должна представлять собой интернет-ресурс, разработанный в среде разработки Bitrix Framework.

Каждому пользователю должно предоставляться персональное хранилище данных в размере 5 ГБ с возможностью расширения.

Для хранения сведений о пользователях системы, записях образовательного форума и метаданных используется база данных под управлением PostgreSQL.

Для хранения пользовательских файлов должно использоваться хранилище блочного типа. Файлы в таком хранилище делятся на блоки равного размера и размещаются в памяти сервера. По запросу платформы система хранения собирает файл из блоков, используя метаданные.

1.4.5 Требования к пользовательскому интерфейсу

Пользовательский интерфейс должен содержать нейтральные цвета и контрастный, хорошо читаемый текст. Основные цвета — белый, синий.

Обязательными элементами пользовательская являются:

Наличие функции поиска.

Является обязательным условием каждого крупного веб-проекта, состоящего более чем из 10-и страниц.

Реализация простого поиска, позволяющий находить нужную информацию по ключевым запросам. Поисковое окно располагается в верхней части сайта и проходит через все страницы сайта.

Меню

Список доступных разделов и категорий должен быть всегда заметен пользователю, независимо от его местонахождения на сайте. Следовательно, пользователь всегда должен видеть возможные варианты переходов. Кроме того, подобное «сквозное» меню упрощает индексацию сайта для поисковых машин, способствует равномерному распределению «веса» между всеми страницами сайта.

Веб-интерфейс должен содержать элементы управления и манипуляции документами, вкладки и разделы, отведенные под форум.

Адаптивный дизайн

Дизайн веб-страниц должен обеспечивать правильное отображение сайта на различных устройствах, подключённых к интернету и динамически подстраиваться под заданные размеры окна браузера.

Сайт может работать на смартфоне, планшете, ноутбуке и телевизоре с выходом в интернет, то есть практически на всем спектре устройств.

1.4.6 Требования к защите информации

Для обеспечения защиты информации от несанкционированного доступа требуется реализация следующих функций:

В части управления доступом:

- должна осуществляться идентификация и проверка подлинности субъектов доступа при входе в Систему по идентификатору (коду) и паролю условно-постоянного действия длиной не менее шести символов;
- должна осуществляться идентификация АРМ, серверов, каналов связи, внешних устройств ЭВМ по логическим именам;
- должен осуществляться контроль доступа субъектов к защищаемым ресурсам в соответствии с ролевой моделью.

В части регуляции и учета:

- должна осуществляться регистрация входа/выхода субъектов доступа в систему/из системы.
- В параметрах регистрации должны указываться:
 - время и дата входа/выхода субъекта доступа в систему/из системы
 - идентификатор (код или фамилия) субъекта, предъявленный при попытке доступа;

1.4.7 Требования по сохранности информации при авариях

Сохранность информации в Системе должна обеспечиваться при следующих аварийных ситуациях:

- импульсные помехи, сбои и перерывы в электропитании;
- нарушение или выход из строя каналов связи локальной сети;
- сбой общего или специального программного обеспечения (сервера);
- ошибки в работе персонала

1.4.8 Требования по стандартизации и унификации

Для работы с БД должен использоваться язык SQL в рамках стандарта ANSI SQL-92.

Для разработки пользовательского интерфейса и средств генерации отчетов (любых твердых копий) должны использоваться языки 4-го поколения.

При создании Системы должно использоваться стандартное общее программное обеспечение, включающее лицензионные ОС, СУБД, сетевую операционную систему.

1.4.9 Требования к безопасности

Система не должна выдавать варианты, которые потенциально могут усугубить состояние оборудования.

Система должна выдавать пользователю правила технической безопасности, которые ему необходимо принять, чтобы пользоваться системой.

Системой должен осуществляться контроль доступа к конфиденциальным данным.

Должны предприниматься шаги, направленные на предотвращение утечек персональных данных пользователей Системы.

Система должна быть защищена от внешних атак.

1.4.10 Требования к документированию

Проектная документация должна быть разработана в соответствии с ГОСТ 34.201-89 и ГОСТ 7.32-2017.

Отчетные материалы должны включать в себя текстовые материалы (представленные в виде бумажной копии и на цифровом носителе) и графические материалы.

Предоставить следующие документы:

- 1) Схема функциональной структуры автоматизируемой деятельности.
- 2) Описание технологического процесса обработки данных.
- 3) Описание информационного обеспечения.
- 4) Описание программного обеспечения АС.
- 5) Схема логической структуры БД.
- 6) Руководство пользователя.
- 7) Описание контрольного примера (по ГОСТ 24.102).
- 8) Протокол испытаний (по ГОСТ 24.102).

1.4.11 Требования к функциям, выполняемым системой

Таблица 2 — Требования к функциям, выполняемым системой.

Функция	Задача
Работа с пользователями	Регистрация пользователя
	Авторизация пользователя
Добавление новых документов в хранилище данных	Заполнение сведений о документе
	Загрузка документа в объектное хранилище данных
	Обновление сведений о документах пользователя
	Проведение проверки и корректировку связей в хранилище метаданных
Загрузка документов из хранилища данных	Запрос сведений о документе
	Выгрузка документа из объектного хранилища данных
	Выдача результата пользователю
	Проведение проверки и корректировку связей в хранилище метаданных

1.5 СЦЕНАРИИ ИСПОЛЬЗОВАНИЯ

1) Пользователь заходит на веб ресурс с целью ознакомления с содержанием или скачивания интересующего его документа. Он может оценить размещенный документ, оставив отметку «Нравится». Также пользователь может получить цифровую копию документа не проходя аутентификацию в системе.

2) Пользователь заходит на веб-ресурс с целью загрузки документа. Для этого ему необходимо пройти регистрацию в Системе. После этого пользователь получит доступ к личному кабинету и персональному хранилищу данных размером 5 ГБ.

3) Пользователь заходит на веб-ресурс с целью заведения темы на форуме, входящем в состав Системы. Для этого ему необходимо пройти регистрацию в Системе. После этого пользователь получит доступ к личному кабинету и получит возможность оставлять комментарии под записями на форуме, создавать новые темы для обсуждения.

2 Выбор архитектуры системы

2.1 ФУНКЦИОНАЛЬНЫЕ СХЕМЫ СИСТЕМЫ

В ходе данной практической работы были разработаны три различные функциональные схемы для реализации информационной системы «Электронный сборник лабораторных работ».

Архитектура информационной системы — концепция, определяющая модель, структуру, выполняемые функции и взаимосвязь компонентов информационной системы. Можно сформулировать проще: Информационная система — это совокупность программного обеспечения, решающего определенную прикладную задачу.

Архитектура информационной системы — абстрактное понятие, определяющее, из каких составных частей (элементов, компонент) состоит приложение и как эти части между собой взаимодействуют.

2.1.1 Распределенная архитектура

Ключевое отличие данной архитектуры — абстрагирование от физической схемы данных и манипулирование данными клиентскими программами на уровне логической схемы. Это позволило создавать надежные многопользовательские информационные системы с централизованной базой данных, независимые от аппаратной (а часто и программной) части сервера базы данных и поддерживающие графический интерфейс пользователя на клиентских станциях, связанных локальной сетью.

2.1.2 Клиент-серверная архитектура

На рисунке 2.3 представлена схема клиент-серверной архитектуры разрабатываемой системы. Все подсистемы можно разбить на 2 группы: подсистемы серверной части и подсистемы клиентской части.

Серверная часть системы включает в себя подсистему БД и подсистему обработки запросов сетевого протокола HTTPS. Подсистема БД содержит СУБД, выполняющую запросы к БД, саму БД, также система должна выполнять резервное копирование данных. Подсистема обработки запросов

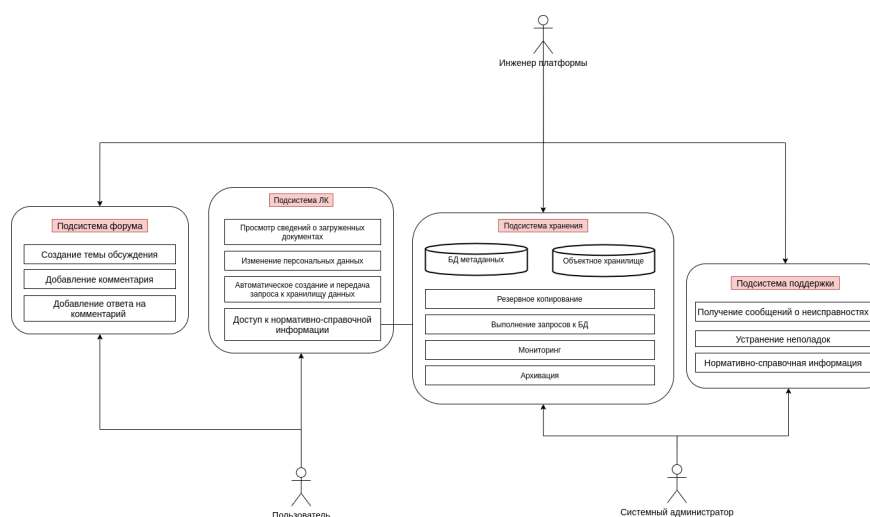


Рисунок 1 — Распределенная архитектура системы

призвана обрабатывать HTTPS запросы, поступающие от клиентов, и формировать запросы к БД в соответствии с целью запроса. Подсистемы ЛКС и контроля успеваемости предназначены для обеспечения диалога между системой и пользователями 2 групп: студентов и сотрудников деканата. Подсистемы имеют графический интерфейс, что облегчает взаимодействие пользователей с системой в целом. Подсистемы также автоматически собирают и отправляют HTTPS запросы в соответствии с нуждами и действиями пользователя.

Особенности:

- 1) клиентская программа работает с данными через запросы к серверному ПО; 14
- 2) базовые функции приложения разделены между клиентом и сервером. Положительные стороны:
 - 3) полная поддержка многопользовательской работы;
 - 4) гарантия целостности данных. Отрицательные стороны:
 - 5) Бизнес логика приложений осталась в клиентском ПО. При любом изменении алгоритмов, надо обновлять пользовательское ПО на каждом клиенте.
 - 6) Высокие требования к пропускной способности коммуникационных каналов с сервером.
 - 7) Слабая защита данных от взлома, в особенности от недобросовестных пользователей системы.
 - 8) Высокая сложность администрирования и настройки рабочих

мест пользователей системы.

9) Необходимость использовать мощные ПК на клиентских местах.

10) Высокая сложность разработки системы из-за необходимости выполнять бизнес-логику и обеспечивать пользовательский интерфейс в одной программе.

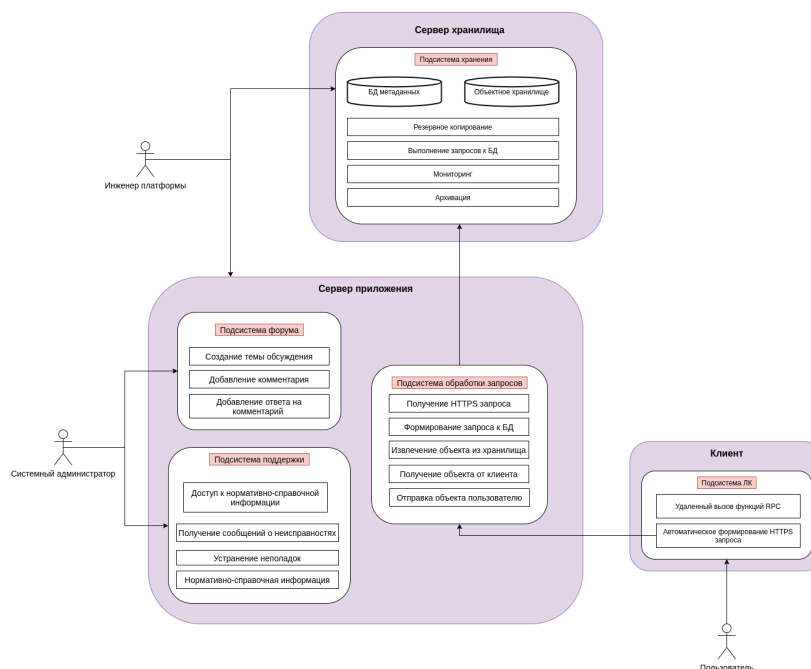


Рисунок 2 — Клиент-серверная архитектура системы

2.2 ВЫБОР АРХИТЕКТУРЫ

В ходе выполнения данной практической работы были предложены две различные функциональные схемы системы. Проанализировав каждую из них, можно сделать вывод, что оптимальным вариантом является система с клиент-серверной архитектурой. Основанием для такого заключения являются следующие характеристики:

- наличие графического интерфейса пользователя;
- зашифрованная передача запросов по протоколу HTTPS.

Графический интерфейс позволит работать с системой «Электронный сборник лабораторных работ» даже пользователям, не имеющим специфических навыков работы с компьютером, т.е. навыки взаимодействия, например, с СУБД не требуются. Защищенное соединение по протоколу HTTPS обеспечит дополнительную защиту данных при передаче их по сети Интернет.

3 Функциональное проектирование (методология SADT)

3.1 ЦЕЛЬ СОЗДАНИЯ СИСТЕМЫ

Разрабатываемая информационная система «Электронный сборник лабораторных работ» (далее по тексту Система) служит для сбора, хранения и учета письменных работ различного характера, выполненных учащимися средних общеобразовательных, средних специальных и высших учебных заведений. На данной платформе пользователям на безвозмездной основе предоставляется доступ к загруженным документам, топикам и темам форума и медиафайлам.

Система выполняет функции образовательной платформы. Платформа может быть интегрирована в информационную среду учебных заведений, предоставляя инструменты для учета и хранения работ учащихся.

3.2 КРАТКОЕ ОПИСАНИЕ

Пользователь взаимодействует с системой посредством веб-сервиса. Сайт является удобным интернет сервисом, представляющим информацию о персональных файлах, перечне загруженных в Систему работ. Веб-интерфейс реализует возможность доступа к форуму для обсуждений. Для комфортного и круглосуточного доступа, сайт так же адаптирован для мобильных устройств.

Кроме того, платформа позволяет пользователю просматривать файлы, основываясь на большом количестве поисковых фильтров. Сортировка возможна по времени загрузки, наименованию научной дисциплины, виду и популярности работы.

3.3 СРЕДСТВА СОЗДАНИЯ ИС

В качестве средств создания ИС был использован язык программирования PHP и JavaScript. Для хранения метаданных и системных сообщений используется база данных под управлением СУБД PostgreSQL. Для хранения пользовательских файлов используется объектное

хранилище S3.

В качестве инструмента веб-аналитики используется Яндекс.Метрика. Данный инструмент помогает получать наглядные отчеты, видеозаписи действий посетителей, отслеживать источники трафика и оценивать эффективность онлайн- и офлайн-рекламы.

Система построена в соответствии с подходом Ajax, заключающийся в «фоновом» обмене данными браузера с веб-сервером. В результате при обновлении данных веб-страница не перезагружается полностью, и веб-приложения становятся быстрее и удобнее.

Для моделирования проектируемой ИС будет использоваться нотация IDEF0 программном обеспечении CASE Ramus Educational edition.

3.4 ПРОЕКТИРОВАНИЕ КОНТЕКСТНОЙ ДИАГРАММЫ

Была спроектирована контекстная диаграмма А-0 в нотации IDEF0. В качестве входа по управлению (стрелка управления) были выбраны следующие нормативные и правовые документы:

- 1) 152-ФЗ "О персональных данных".
- 2) 149-ФЗ "Об информации, информационных технологиях и о защите информации".
- 3) Политика интернет-сайта.
- 4) Пользовательское соглашение.

В качестве входящих информационных потоков, которые подлежат обработке и преобразованию в процессе работы ИС была указана следующая информация:

- 1) Персональные данные пользователя
- 2) Пользовательские файлы

В качестве механизмов (ресурсов, выполняющих работу) были выделены:

- 1) Подсистема хранения.
- 2) Подсистема обработки запросов.
- 3) Подсистема технической поддержки.
- 4) Подсистема личного кабинета.
- 5) Подсистема форума.
- 6) Модератор форума.

7) Пользователь системы.

В качестве выходов после выполнения ИС получены следующие информационные элементы:

- 1) Статистические данные об активности пользователей.
- 2) Сведения о работе подсистем.

Так как механизмы «Подсистема хранения», «Подсистема обработки запросов» и «Пользователь системы» используются во всех выделяемых в ходе дальнейшей декомпозиции процессов, то соответствующие стрелки туннелированы. На схеме оконечности данных стрелок заключены в круглые скобки.

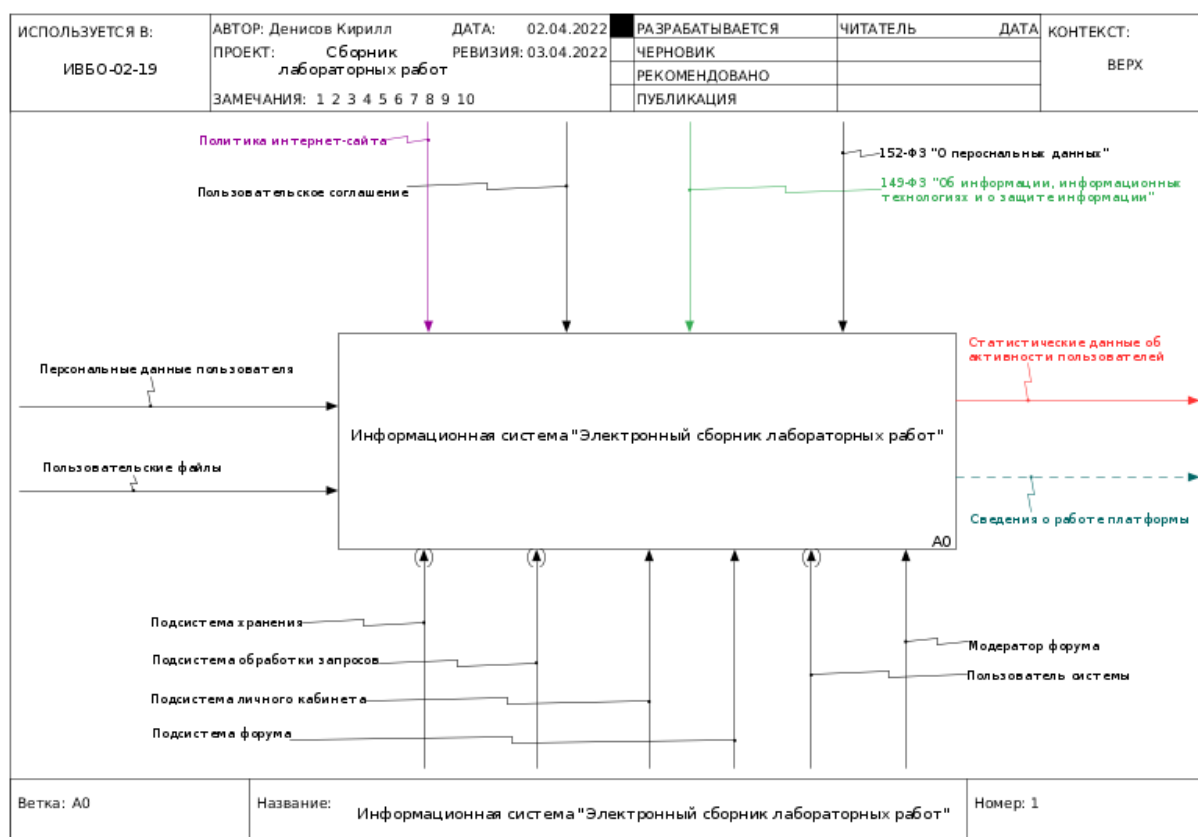


Рисунок 3 — Контекстная диаграмма процесса предоставления услуг сервисом "Сборник лабораторных работ"

4 Функциональное проектирование (нотация IDEF0)

4.1 ОПИСАНИЕ ПРОЦЕССОВ ДЕКОМПОЗИЦИИ

На диаграмме уровня A0 декомпозиции функционального блока «Информационная система "Электронный сборник лабораторных работ"» обозначены процессы и функциональные блоки, выполняемые в рамках процедуры:

- 1) Процесс регистрации пользователя в системе.
- 2) Процесс публикации файла в системе.
- 3) Процесс получения файла, размещенного в системе.
- 4) Процесс участия в обсуждении на форуме.

Диаграмма декомпозиции блока «Информационная система "Электронный сборник лабораторных работ"» приведена на рисунке 4.

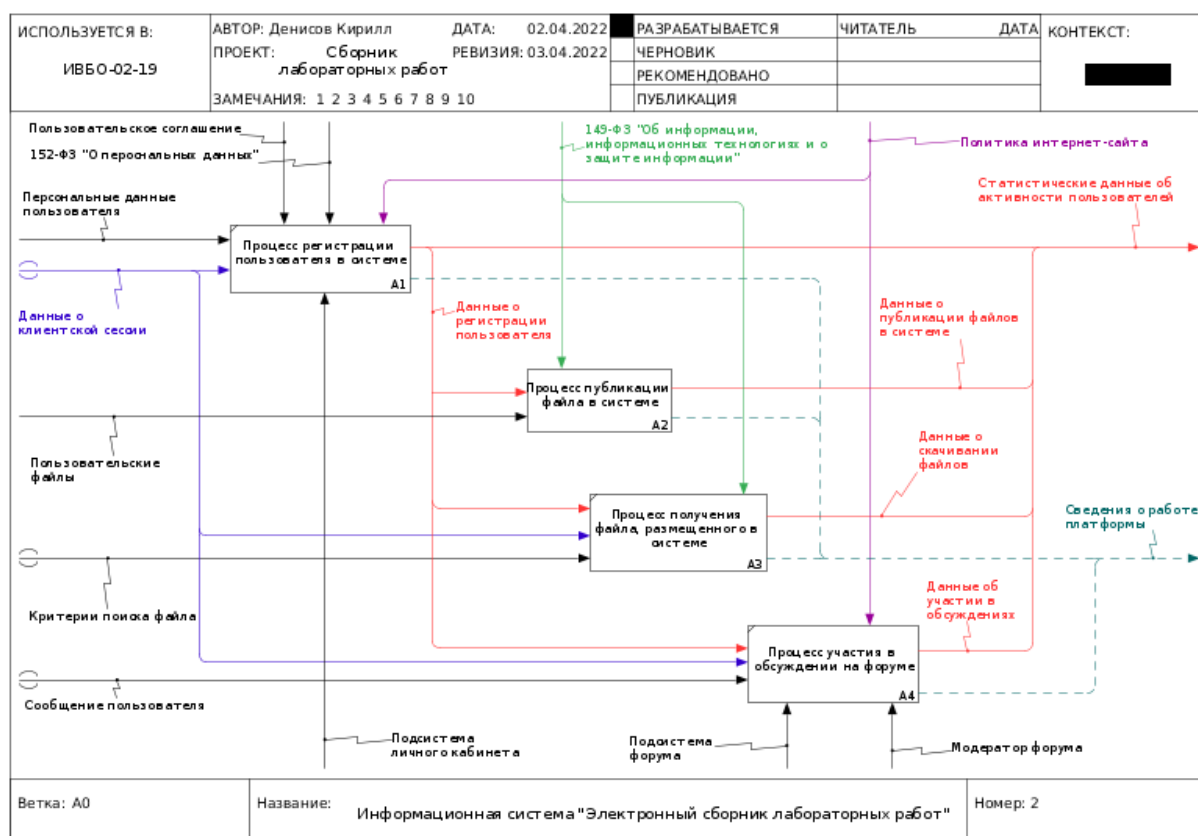


Рисунок 4 — Диаграмма декомпозиции

4.1.1 Процесс регистрации пользователя в системе

Процесс регистрации проходит согласно закону «О персональных данных» и политике интернет-сайта а также условиям, описанным в пользовательском соглашении. В качестве входных потоков выступают персональные данные пользователя и данные о клиентской сессии. После регистрации пользователь может перейти к процессу публикации файлов или процессу получения файлов, размещенных на платформе. В качестве механизмов процесса выступают «Подсистема хранения» и «Пользователь системы», а также «Подсистема личного кабинета». Выходные потоки процесса — данные о регистрации пользователя в системе и сведения о работе платформы (логи и метаданные подсистем).

4.1.2 Процесс публикации файла в системе

Процесс публикации файла в системе происходит в соответствии с федеральным законом «Об информации, информационных технологиях и о защите информации», который оперирует понятиями «объект авторского права» и устанавливает правила распространения загружаемых объектов. В качестве входных потоков выступают данные о регистрации пользователя и данные о клиентской сессии. В качестве механизмов процесса выступают «Подсистема хранения» и «Пользователь системы». Выходные потоки процесса — данные о публикации файлов в системе и сведения о работе платформы (логи и метаданные подсистем).

4.1.3 Процесс получения файла, размещенного в системе

Процесс получения файла, размещенного в системе происходит в соответствии с федеральным законом «Об информации, информационных технологиях и о защите информации», который оперирует понятиями «объект авторского права» и устанавливает правила распространения загружаемых объектов. В качестве входных потоков выступают данные о регистрации пользователя, данные о клиентской сессии и критерии поиска файла, предоставленные пользователем. В качестве механизмов процесса выступают «Подсистема хранения» и «Пользователь системы». Выходные потоки процесса — данные о скачивании файлов и сведения о работе платформы

- 1) Выбор наименования научной дисциплины.
- 2) Выбор типа научной работы.
- 3) Установка параметров файла.
- 4) Загрузка файла в систему.

Все процессы осуществляются пользователем с помощью «Подсистемы обработки запросов» системы «Электронного сборника лабораторных работ».

4.2.1 Выбор наименования научной дисциплины

Процесс «Выбор наименования научной дисциплины» принимает в качестве входного потока данные о регистрации пользователя. Выходным потоком являются сведения о наименовании научной дисциплины.

4.2.2 Выбор типа научной работ

Процесс «Выбор типа научной работы» принимает в качестве входного потока данные о регистрации пользователя. Выходным потоком являются сведения о типе научной работы.

4.2.3 Установка параметров файла

Процесс «Установка параметров файла» принимает в качестве входного потока данные о регистрации пользователя. Выходным потоком являются установленные параметры файла.

4.2.4 Загрузка файла в систему

Процесс «Загрузка файла в систему» в дополнение к механизмам «Подсистема обработки запросов» и «Пользователь системы» использует «Подсистему хранения». В качестве входного потока данный процесс принимает сведения о регистрации пользователя, сведения о наименовании научной дисциплины, сведения о типе научной работы, установленные параметры файла и пользовательские файлы. Процесс проходит в соответствии с федеральным законом «Об информации, информационных технологиях и о защите информации», который оперирует понятиями «объект авторского права» и устанавливает правила распространения загружаемых объектов.

5 Функциональное проектирование (нотация DFD)

5.1 ОПРЕДЕЛЕНИЕ ОБЪЕКТА ДЕКОМПОЗИЦИИ

Для построение диаграммы в нотации DFD был выбран «Процесс получения файла, размещенного в системе». Данный процесс был выбран, так как он реализует один из базовых функционалов системы и является технически сложным. В ходе данного процесса осуществляется обращение к нескольким хранилищам данных, а именно к хранилищу метаданных, содержащему сведения о предлагаемых к выбору наименованиях научных дисциплин, типах работ, параметрах файлов, хранящихся в объектном хранилище файлов, и к самому объектному хранилищу файлов.

Было выделено два основных процесса на диаграмме потоков данных:

- 1) Установка критериев поиска в блоке фильтров.
- 2) Обращение к хранилищу файлов.

5.1.1 Установка критериев поиска

В ходе первого процесса устанавливаются критерии поиска и фильтры файлов. В ходе данного процесса подсистема обработки запросов формирует запрос к хранилищу метаданных, получая список файлов, удовлетворяющих требованиям пользователя.

5.1.2 Обращение к хранилищу файлов

В ходе второго процесса осуществляется доступ к хранилищу метаданных с целью получения сведений о расположении требуемого файла, затем инициализируется защищенное FTP соединение и пользователю предлагается скачать файл.

Полученная схема в нотации DFD изображена на рисунке 6.

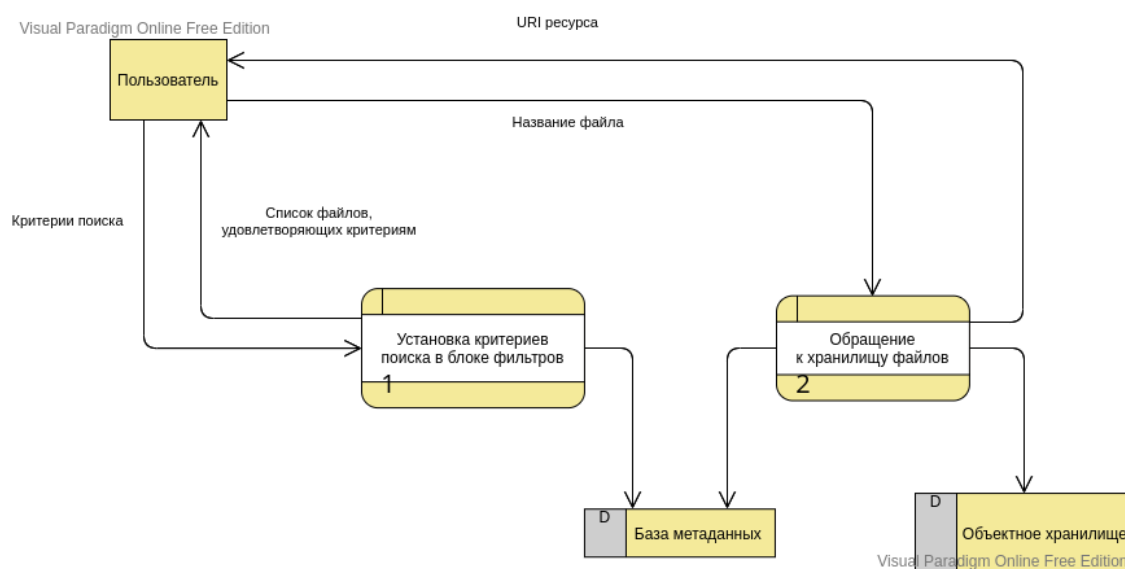


Рисунок 6 — Диаграмма потока данных

6 Проектирование структуры данных ИС (Entity Relation Diagram)

6.1 МОДЕЛИРОВАНИЕ СИСТЕМЫ

Для логического моделирования было выбрано программное обеспечение Draw.io в силу того, что оно предлагает удобный веб-интерфейс для работы а также реализует возможность создания диаграмм в различных нотациях. Данное программное обеспечение позволяет экспортировать файлы в формате, удобном для переноса между платформами и в виде изображений, что удобно при разработке документации, прилагаемой к реализуемой ИС.

6.2 КРАТКАЯ ПОСТАНОВКА ЗАДАЧИ

Главная задача системы — сбор и обработка научных и ученических работ пользователей. Система должна представлять данные о файлах в структурированном виде, предлагать удобный интерфейс взаимодействия с личными файлами и представлять возможность получения файлов, хранящихся в системе.

Основываясь на поставленной задаче, была создана модель данных базы метаданных, обеспечивающей работу и обслуживание объектного хранилища, содержащего пользовательские файлы, загружаемые в систему.

Хранилище метаданных выполнено по схеме «Звезда», где основная таблица фактов связана с несколькими таблицами измерений, организуя удобную для хранения многомерных показателей схему реалационных таблиц. В таблице фактов содержатся следующие данные:

В таблице измерения «Авторы» содержатся следующие данные:

В таблице измерения «Организации» содержатся следующие данные:

Таблица 3 — Таблица фактов *PaperFacts*

Название	Назначение
ID	Индификатор записи
paperID	Индификатор документа
authorID	Индификатор автора
disciplineID	Индификатор дисциплины
work_typeID	Индификатор типа работы
organizationID	Индификатор организации
licenseID	Индификатор лицензии
timestamp	Временная метка
status	Статус для версионирования
name	Название работы

Таблица 4 — Таблица измерения *Authors*

Название	Назначение
authorID	Индификатор автора
firstname	Имя
secondname	Фамилия
degree	Степень (должность)
nickname	Никнейм
email	Адрес электронной почты
country	Страна
city	Город

Таблица 5 — Таблица измерения *Organizations*

Название	Назначение
organizationID	Индификатор организации
country	Страна
city	Город
type	Тип организации
name	Название организации

В таблице измерения «Лицензии» содержатся следующие данные:

Таблица 6 — Таблица измерения *Licenses*

Название	Назначение
licenseID	Индификатор лицензии
name	Название лицензии
supervisor	Контролирующая организация

В таблице измерения «Типы работ» содержатся следующие данные:

Таблица 7 — Таблица измерения WorkTypes

Название	Назначение
work_typeID	Индификатор типа работы
name	Название типа работы

В таблице измерения «Работы» содержатся следующие данные:

Таблица 8 — Таблица измерения Papers

Название	Назначение
paperID	Индификатор файла работы
URI	Универсальный индификатор ресурса
status	Статус для версионирования

В таблице измерения «Дисциплины» содержатся следующие данные:

Таблица 9 — Таблица измерения Disciplines

Название	Назначение
disciplineID	Индификатор дисциплины
name	Название дисциплины

Таблица фактов поддерживает версионирование медленно изменяющихся измерений второго типа (SCD2). Версионирование нужно не для хранения разных версий одного файла в системе, а для отражения в базе данных факта изменения параметров файла. Например, при изменении типа работы запись об этом сохранится в базе, при этом это никак не скажется на доступности файла в объектном хранилище, так как ссылка на файловый ресурс не будет затронута.

6.3 ПОСТРОЕНИЕ ER ДИАГРАММЫ

ER диаграмма приведена на рисунке 7.

Разработанный пример ER-диаграммы является примером концептуальной диаграммы, не учитывающей особенности конкретной СУБД.

На основе данной концептуальной диаграммы можно построить физическую диаграмму, которая будут учитывать такие особенности

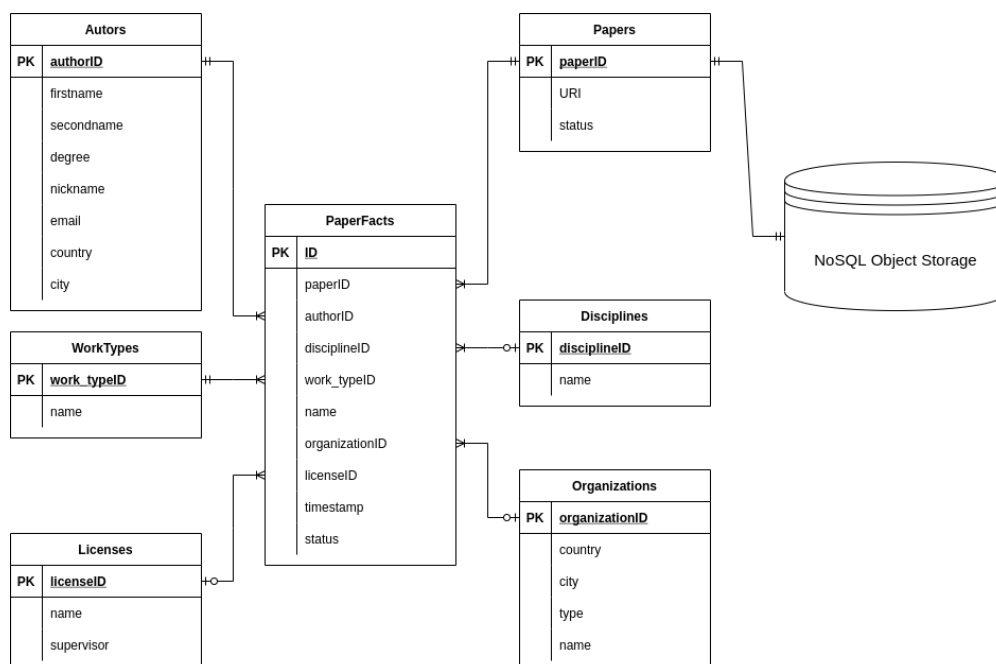


Рисунок 7 — ER диаграмма

СУБД, как допустимые типы, наименования полей и таблиц, ограничения целостности и т.п. Для преобразования концептуальной модели в физическую необходимо знать, что:

- 1) Каждая сущность в ER-диаграмме представляет собой таблицу базы данных;
- 2) Каждый атрибут становится колонкой (полем) соответствующей таблицы;
- 3) В некоторых таблицах необходимо вставить новые атрибуты (поля), которых не было в концептуальной модели — это ключевые атрибуты родительских таблиц, перемещённых в дочерние таблицы для того, чтобы обеспечить связь между таблицами посредством внешних ключей.

7 СОЗДАНИЕ ДИАГРАММЫ СОСТОЯНИЙ

Диаграммы состояний применяются, как правило, для моделирования поведения классов, прецедентов или системы в целом.

Составим диаграмму состояний для объекта, реализующего поиск по файлам.

1) Первым этапом выступает инициализация поискового запроса. При передаче управления этому процессу передаются параметры поиска, а на выходе данные поиска сохраняются.

2) Второй процесс — выдача результатов поиска, в ходе которого пользователю предоставляется список файлов, удовлетворяющих параметрам поиска. При необходимости параметры поиска могут быть уточнены.

3) Последним процессом выступает обращение к объектному хранилищу данных, для того, чтобы предоставить пользователю сведения, необходимые для скачивания файла.

В случае, если был найден хотя бы один файл, процесс завершается успешно. Если ни один файл из объектного хранилища не удовлетворял критериям поиска, процесс завершится не успешно. Пользователю будет предложено изменить критерии поиска.

Диаграмма состояний приведена на рисунке 8.

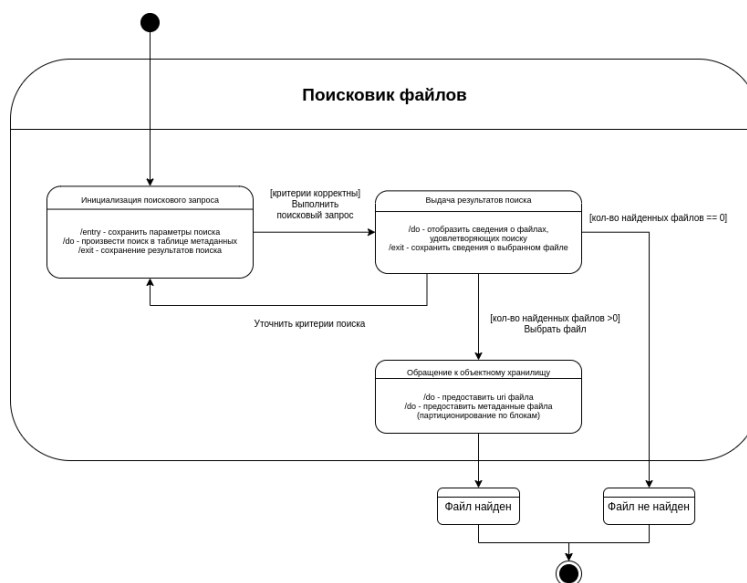


Рисунок 8 — Диаграмма состояний

8 Расчет параметров проектируемой информационной системы

8.1 ОПИСАНИЕ ЭСЕ

Элементарная семантическая единица (ЭСЕ) — неделимая единица информации, используемая в ИС. ЭСЕ представляет собой завершенную контекстную конструкцию, вызываемую в результате поиска по различным атрибутам или в результате тех или иных команд в виде отклика или отчета. В случае исследования настоящей системы за элементарную семантическую единицу была выбрана одна из характеристик поиска, а именно, количество файлов, удовлетворяющим пользовательским критериям поиска. В нашем примере эта величина меняется случайным образом в пределах от 0 до 49999 [файлов].

8.2 НАПОЛНЕНИЕ СИСТЕМЫ

Проектируемая информационная система «Электронный сборник лабораторных работ» может быть наполнена практически любым количеством элементов базы данных. Их количество ограничиваются только параметрами сервера.

В рамках данной практической работы система была наполнена 50000 ЭСЕ. Было проведено сто экспериментов, в ходе которых стало известно количество файлов, удовлетворяющих пользовательским параметрам, заданным в модуле поиска Подсистемы хранения ИС «Электронный сборник лабораторных работ». Список результатов измерений приведен в приложении 1.

8.3 МАТЕМАТИЧЕСКИЕ РАСЧЕТЫ

Для дальнейшего исследования проектируемой ИС необходимо рассчитать вероятности, с которыми ЭСЕ принимает то или иное значение. Для оценки этих вероятностей было принято решение разбить весь диапазон значений на 10 дискретных величин с шагом в 5000. Расчеты вероятности

ведутся с помощью следующей формулы 1.

$$P(\xi) = \frac{n}{N} \quad (1)$$

В данной формуле n — благоприятное число исходов (в данном случае число найденных файлов, попавших в данный диапазон), а N — общее число исходов.

В таблице 10 приведены возможные значения, принимаемые ЭСЕ и их вероятности.

Таблица 10 — Ряд распределения

№	x_i	$P(x)$
1	2499.50	150.15
2	7499.50	100.10
3	12499.50	80.08
4	17499.50	90.09
5	22499.50	110.11
6	27499.50	130.13
7	32499.50	100.10
8	37499.50	90.09
9	42499.50	60.06
10	47499.50	90.09

8.3.1 Расчет математического ожидания информационного блока системы

Математическим ожиданием случайной величины называется сумма произведений всех возможных значений случайной величины на вероятности этих значений.

Рассчитаем математическое ожидание для нашей системы, взяв за случайную величину число файлов. Расчёт математического ожидания информационного блока на примере 10 записей:

$$M_{x_i} = \sum_{i=0}^n [p_i \cdot x_i] \quad (2)$$

Используя данные, полученные в таблице 10, получаем:

$M(10) = 23199.50$ файлов, следовательно, наиболее вероятное количество файлов в ответе на запрос находится в районе 23199.50 [файлов].

8.3.2 Расчет дисперсии информационного блока системы

Рассчитаем дисперсию информационного блока системы по формуле 3.

$$D_{x_i} = \sum_{i=0}^n [p_i \cdot (x_i)^2] - \left[\sum_{i=0}^n (p_i \cdot x_i) \right]^2 \quad (3)$$

Используя данные, полученные в таблице 10, получаем:

$$D(10) = 206010000 \text{ файлов}^2$$

8.3.3 Расчет среднеквадратического отклонения

Рассчитаем среднеквадратическое отклонение по формуле 4.

$$\sigma x_i = \sqrt{Dx_i} = \sqrt{206010000} = 14353.05 \text{ файлов} \quad (4)$$

8.3.4 Расчет энтропии системы

Информационная энтропия — мера неопределённости некоторой системы (в статистической физике или теории информации), в частности, непредсказуемость появления какого-либо символа первичного алфавита. В последнем случае при отсутствии информационных потерь энтропия численно равна количеству информации на символ передаваемого сообщения

Вычислим энтропию системы по формуле

$$H(x) = - \sum_{i=1}^n [p_i \cdot \log_a p_i] \quad (5)$$

За основание логарифма a возьмем двоичную систему счисления (формула Шеннона). Энтропия фрагмента информационного наполнения в размере 10 ЭСЕ: Используя данные, полученные в таблице 10, получаем: $H(10) = 3.28$ бит.

8.4 РЕЗУЛЬТАТЫ РАСЧЕТОВ

В данной главе был осуществлен расчет основных характеристик проектируемой ИС, и получены следующие результаты:

Таблица 11 — Параметры проектируемой ИС

Математическое ожидание информационного блока	$M(10) = 23199.50$ файлов
Допустимый разброс значений смысловых информационных блоков (дисперсия)	$D(10) = 206010000$ файлов ²
Среднеквадратичное отклонение	$\sigma x_i = 14353.05$ файлов
Энтропия информационного наполнения	$H(10) = 3.28$ бит

Приведем таблицу с исходными сгенерированными данными (см. таблицу 12).

Таблица 12 — Исходные значения

№	Значение	№	Значение	№	Значение	№	Значение
1	19956	26	1983	51	3986	76	47044
2	30937	27	26452	52	20469	77	26520
3	49379	28	899	53	14592	78	35967
4	32802	29	41888	54	1352	79	21749
5	3998	30	8182	55	32704	80	18295
6	31012	31	2176	56	40990	81	21813
7	38782	32	28627	57	29975	82	13467
8	15895	33	25433	58	10325	83	23783
9	27847	34	9457	59	3700	84	32821
10	35964	35	2373	60	24083	85	17607
11	3827	36	10010	61	33412	86	36274
12	24247	37	48230	62	43067	87	10648
13	39557	38	26061	63	49735	88	19679
14	2321	39	5810	64	40199	89	49114
15	29616	40	25460	65	20796	90	43305
16	24485	41	30058	66	27099	91	21310
17	46153	42	859	67	36430	92	31711
18	35865	43	46295	68	18874	93	18424
19	44427	44	1274	69	6761	94	6493
20	26134	45	37563	70	8725	95	5029
21	2133	46	47881	71	25199	96	9433
22	3270	47	15273	72	21981	97	38202
23	27499	48	1346	73	48131	98	33052
24	33954	49	11422	74	13900	99	24332
25	17943	50	11215	75	6764	100	5224