

**Министерство образования и науки Российской Федерации**  
федеральное государственное автономное образовательное учреждение  
высшего образования  
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

---

Институт кибернетики  
Направление подготовки 09.04.01 Информатика и вычислительная техника  
Кафедра программной инженерии

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

Тема работы
<b>Разработка системы анализа факторов ранжирования web-ресурсов</b>

УДК 004.42:004.738.1-048.57

Студент

Группа	ФИО	Подпись	Дата
8ВМ5В	Бондарчук Наталия Владимировна		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент каф. ПИ	Савельев Алексей Олегович	К.Т.Н.		

**КОНСУЛЬТАНТЫ:**

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. Менеджмента	Конотопский Владимир Юрьевич	К.Э.Н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент каф. ЭБЖ	Акулов Петр Анатольевич			

**ДОПУСТИТЬ К ЗАЩИТЕ:**

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. ПИ	Иванов Максим Анатольевич	К.Т.Н.		

Томск – 2017 г.

<b>Код результата</b>	<b>Результат обучения (выпускник должен быть готов)</b>
	Профессиональные компетенции
P1	Применять базовые и специальные естественно-научные и математические знания в области информатики, экономики, маркетинга и менеджмента, достаточные для комплексной инженерной деятельности.
P2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных и экономических задач.
P3	Ставить и решать задачи комплексного анализа, связанные с созданием новых информационных технологий и информационных систем в экономике, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
P4	Разрабатывать новые и модернизировать уже существующие информационные технологии и системы (в экономике) в соответствии с техническим заданием.
P5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области прикладной информатики. Проводить исследования, связанные с оценкой информационной безопасности проектов.
P6	Внедрять, эксплуатировать и обслуживать современные информационные технологии и системы, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.
	Универсальные компетенции
P7	Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.
P8	Владеть иностранным языком на уровне, позволяющем работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.
P9	Эффективно работать индивидуально и в качестве члена группы, состоящей из специалистов различных направлений и квалификаций, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.
P10	Демонстрировать знания правовых, социальных, экономических и культурных аспектов комплексной инженерной деятельности.
P11	Демонстрировать способность к самостоятельному обучению в течение всей жизни и непрерывному самосовершенствованию в инженерной профессии.

**Министерство образования и науки Российской Федерации**

Федеральное государственное бюджетное образовательное учреждение  
высшего профессионального образования  
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Кибернетики

Направление подготовки (специальность) 09.04.01 Информатика и вычислительная техника

Кафедра Программной инженерии

УТВЕРЖДАЮ:

Зав. кафедрой

М.А. Иванов

\_\_\_\_\_  
(Подпись)      \_\_\_\_\_  
(Дата)

(Ф.И.О.)

**ЗАДАНИЕ**

**на выполнение выпускной квалификационной работы**

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8ВМ5В	Бондарчук Наталии Владимировне

Тема работы:

Разработка системы анализа факторов ранжирования web-ресурсов

Утверждена приказом директора (дата, номер)

От 14.04.2017 г. №2600/с

Срок сдачи студентом выполненной работы:

8.06.2017 г.

**ТЕХНИЧЕСКОЕ ЗАДАНИЕ:**

**Исходные данные к работе**

1. Результаты производственной практики.

**Перечень подлежащих исследованию, проектированию и разработке вопросов**

1. Анализ предметной области;
2. Обзор аналогов;
3. Проектирование архитектуры системы;
4. Проектирование базы данных;
5. Проектирование пользовательского интерфейса;
6. Программная реализация информационной системы и ее результаты.

**Перечень графического материала**

1. Архитектура системы
2. Схема базы данных
3. Интерфейс пользователя

**Консультанты по разделам выпускной квалификационной работы**

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Конотопский В.Ю.
Социальная	Акулов П.А.

<b>ответственность</b>	
<b>Названия разделов, которые должны быть написаны на русском и иностранном языках:</b>	
<b>Анализ предметной области</b>	

<b>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</b>	
---	--

**Задание выдал руководитель:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент кафедры ПИ	Савельев А.О.	к.т.н		

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ВМ5В	Бондарчук Наталия Владимировна		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА  
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И  
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

<b>Группа</b>	<b>ФИО</b>
8ВМ5В	Бондарчук Наталии Владимировне

<b>Институт</b>	<b>Кибернетики</b>	<b>Кафедра</b>	<b>ПИ</b>
<b>Уровень образования</b>	<b>Магистр</b>	<b>Направление/специальность</b>	09.04.01 Информатика и вычислительная техника

**Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:**

- 1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих*
- 2. Нормы и нормативы расходования ресурсов*
- 3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования*

**Перечень вопросов, подлежащих исследованию, проектированию и разработке:**

- 1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения*
- 2. Планирование и формирование бюджета научных исследований*
- 3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования*

**Перечень графического материала (с точным указанием обязательных чертежей):**

- 1. Перечень работ и продолжительность их выполнения*
- 2. Трудозатраты на выполнение проекта*
- 3. Линейный график работ*
- 4. Нарастание технической готовности работы и удельный вес каждого этапа*
- 5. Смета затрат на разработку проекта*
- 6. Оценки научно-технического уровня НИР*

**Дата выдачи задания для раздела по линейному графику**

**Задание выдал консультант:**

<b>Должность</b>	<b>ФИО</b>	<b>Ученая степень, звание</b>	<b>Подпись</b>	<b>Дата</b>
Доцент	Конотопский Владимир Юрьевич	К. Э. Н.		

**Задание принял к исполнению студент:**

<b>Группа</b>	<b>ФИО</b>	<b>Подпись</b>	<b>Дата</b>
8ВМ5В	Бондарчук Наталия Владимировна		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА  
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

<b>Группа</b>	<b>ФИО</b>
8ВМ5В	Бондарчук Наталии Владимировне

<b>Институт</b>	<b>Кибернетики</b>	<b>Кафедра</b>	<b>ПИ</b>
Уровень образования	Магистр	Направление/специальность	09.04.01 Информатика и вычислительная техника

**Исходные данные к разделу «Социальная ответственность»:**

1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	<i>Система анализа факторов ранжирования web-ресурсов предназначена для автоматизации анализа web-ресурса при поисковом продвижении. Сфера применения продукта – интернет-маркетинг.</i>
--	--

**Перечень вопросов, подлежащих исследованию, проектированию и разработке:**

<p><b>1. Производственная безопасность</b></p> <p>1.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения.</p> <p>1.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения.</p>	<p><i>Основными вредными факторами являются:</i></p> <p><i>-недостаточное освещение</i></p> <p><i>-неблагоприятные условия микроклимата.</i></p> <p><i>Основным опасным фактором является поражение электрическим током.</i></p>
<p><b>2. Экологическая безопасность:</b></p> <ul style="list-style-type: none"> <li>- анализ воздействия объекта на литосферу</li> <li>- мероприятия по снижению</li> </ul>	<p><i>В рассматриваемой работе на первый план выдвигается загрязнение литосферы. В связи с этим необходимо рассмотреть источники загрязнения, такие как бумажные отходы, составные части ПК.</i></p>
<p><b>3. Безопасность в чрезвычайных ситуациях:</b></p> <ul style="list-style-type: none"> <li>- перечень возможных ЧС при разработке и эксплуатации проектируемого решения;</li> <li>- выбор наиболее типичной ЧС;</li> <li>- разработка превентивных мер по предупреждению ЧС;</li> <li>- разработка действий в результате возникшей ЧС и мер по ликвидации её последствий.</li> </ul>	<p><i>Наиболее возможные чрезвычайные ситуации техногенного и стихийного характера:</i></p> <p><i>- Пожары</i></p> <p><i>-Кибертерроризм</i></p> <p><i>Мероприятия по предотвращению наиболее вероятной ЧС- пожар.</i></p>
<p><b>4. Правовые и организационные вопросы обеспечения безопасности:</b></p> <ul style="list-style-type: none"> <li>- специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства;</li> <li>- организационные мероприятия при компоновке рабочей зоны.</li> </ul>	<p><i>Рассмотреть специальные нормы и требования трудового законодательства при работе за компьютером.</i></p> <p><i>Влияние использования программного продукта на рабочий процесс работника..</i></p>

<b>Дата выдачи задания для раздела по линейному графику</b>	
---	--

**Задание выдал консультант:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент каф. ЭБЖ	Акулов П. А.			

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ВМ5В	Бондарчук Наталия Владимировна		

Министерство образования и науки Российской Федерации



Федеральное государственное автономное образовательное учреждение  
высшего образования  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Институт Кибернетики

Направление подготовки (специальность) 09.04.01 Информатика и вычислительная техника

Кафедра Программной инженерии

**КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН  
выполнения выпускной квалификационной работы**

Срок сдачи студентом выполненной работы:	5.06.2017
--	-----------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
5.06.2017	Основная часть.	70
30.05.2017	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.	15
29.05.2017	Социальная ответственность.	15

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент каф. ПИ	Савельев А.О.	К.Т.Н.		

**СОГЛАСОВАНО:**

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. ПИ	Иванов М.А.	К.Т.Н.		

## Реферат

Выпускная квалификационная работа 98 с., 20 рис., 19 табл., 32 источника, 4 прил.

Ключевые слова: поисковая оптимизация сайтов, анализ сайтов, факторы ранжирования, поисковые системы, разработка программного продукта.

Объектом исследования являются web-сайты, а предметом исследования – факторы ранжирования, влияющие на выход в ТОП-10 в поисковых системах, таких как Яндекс и Google.

Цель работы – разработать систему анализа факторов ранжирования web-ресурсов.

В качестве методов исследования использовалось изучение тематической литературы, специализированных журналов, ресурсов Интернет.

В работе проведен анализ предметной области связанной анализом сайтов.

Результатом выполнения выпускной квалификационной работы является система анализа факторов ранжирования web-ресурсов, влияющих на его продвижение в поисковых системах

Разработанная система может быть использована для анализа факторов ранжирования web-ресурсов при проведении работ по поисковой оптимизации.

Пояснительная записка к дипломному проекту выполнена в текстовом редакторе Microsoft Word 2013, информационная система разработана с использованием среды разработки Microsoft Visual Studio 2012 и языка программирования C#. Для создания БД использовалась среда Microsoft SQL Server 2012 Management Studio.

## Содержание

Введение.....	14
1 Анализ предметной области .....	16
1.1 Выбор факторов ранжирования, влияющих на продвижение web-ресурсов в поисковых системах.....	16
2.3.1 Мета-теги: Title, Description, Keywords .....	18
2.3.2 Текст на странице сайта .....	19
2.3.3 Изображения на странице сайта.....	20
2.3.4 Заголовки .....	21
2.3.5 Favicon.....	22
2.3.6 Файл Robots.txt.....	23
2.3.7 Карта сайта .....	23
2.3.8 Время загрузки страницы.....	23
2.3.9 Человекопонятные URL .....	24
1.2 Определение эталонных характеристик параметров .....	25
1.3 Обзор аналогичных систем .....	31
2 Разработка информационной системы.....	33
2.1 Функциональные требования к системе.....	33
2.2 Выбор платформы для реализации информационной системы....	34
2.3 Архитектура системы .....	35
2.3.1 Диаграмма потоков данных .....	35
2.3.2 Взаимодействие системы с источниками данных .....	37
2.3.3 Обработка данных .....	40
2.3.4 Структура базы данных.....	42

2.4	Разработка пользовательского интерфейса .....	48
2.5	Анализ соответствия системы сформулированным требованиям	52
3	Результаты внедрения системы .....	54
4	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение .....	60
4.1	Организация и планирование работ .....	60
4.1.1	Продолжительность этапов работ .....	61
4.1.2	Расчет накопления готовности проекта.....	65
4.2	Расчет сметы затрат на выполнение проекта .....	66
4.2.1	Расчет затрат на материалы .....	67
4.2.2	Расчет заработной платы.....	67
4.2.3	Расчет отчислений от заработной платы .....	68
4.2.4	Расчет затрат на электроэнергию .....	68
4.2.6	Расчет расходов, учитываемых непосредственно на основе платежных (расчетных) документов (кроме суточных).....	70
4.2.5	Расчет амортизационных расходов .....	70
4.2.6	Расчет прочих расходов .....	71
4.2.7	Расчет общей себестоимости разработки.....	72
4.2.8	Расчет прибыли .....	72
4.2.9	Расчет НДС .....	72
4.2.10	Цена разработки НИР .....	73
4.3	Оценка эффективности проекта .....	73
4.3.1	Оценка научно-технического уровня НИР .....	73
5	Социальная ответственность.....	77

5.1	Производственная безопасность .....	77
5.3.1	Вредные факторы при разработке и эксплуатации проектируемого решения .....	78
5.3.2	Опасные факторы при разработке и эксплуатации проектируемого решения .....	81
5.2	Экологическая безопасность .....	82
5.3.1	Источники загрязнения .....	82
5.3.2	Мероприятия по предотвращению загрязнения .....	83
5.3	Безопасность в чрезвычайных ситуациях .....	84
5.3.1	Кибертерроризм .....	84
5.3.2	Пожарная безопасность.....	85
5.4	Правовые и организационные вопросы обеспечения безопасности .....	87
5.4.1	Организационные мероприятия при компоновке рабочей зоны .....	87
5.4.2	Правовые нормы трудового законодательства для рабочей зоны оператора ПЭВМ.....	92
5.4.3	Предотвращение нервно-эмоционального напряжения .....	93
	Заключение .....	96
	Список публикаций.....	98
	Список использованных источников: .....	99
	Приложение А .....	103
	Приложение Б.....	104
	Приложение В .....	107
	Приложение Г .....	108

## **Термины и определения**

**мета-теги:** (англ. meta tags) HTML-теги, предназначенные для предоставления структурированных метаданных о веб-странице.

**релевантность:** это семантическое соответствие поискового запроса. Это некая оценка степени соответствия страницы сайта заданному поисковому запросу.

**SEO:** (Search Engine Optimization) совокупность работ, направленная на улучшение позиций сайта в результатах выдачи поисковых систем для увеличения посещаемости сайта.

**сниппет:** фрагмент исходного текста или кода программы, применяемый в: поисковых системах, текстовых редакторах и средах разработки.

**главное зеркало сайта:** основная версия сайта. Зеркалом именуют полную или практически полную копию сайта, размещённую по другому интернет-адресу.

**кликабельность (CTR):** это показатель, выражающий отношение числа кликов по рекламным объявлениям или баннерам к числу показов, представленное в процентах.

## **Введение**

В современном мире трудно себе представить жизнь без интернета, каждый день миллионы людей заходят в Интернет и ищут необходимую им информацию, товары и услуги. Для любого бизнеса Интернет, а в частности поисковые системы, является одним из эффективных способов привлечения клиентов. Поэтому происходит борьба между сайтами-конкурентами за право быть в ТОП-10 поисковых систем, а соответственно за потенциальных клиентов.

Чтобы вывести сайт в ТОП-10 используется поисковая оптимизация. Самым основным этапом поисковой оптимизации является проведение анализа сайта, который достаточно сложный и требует большого объема знаний и времени. Сложность процесса анализа сайтов заключается в большом количестве факторов ранжирования, таких как мета-теги, тексты, изображения, заголовки, карта сайта и другие [1]. Эти факторы влияют на позиции web-ресурса в поисковых системах.

Для автоматизации процесса анализа сайтов было решено разработать систему, которая поможет владельцам бизнеса, маркетологам, рг-менеджерам и seo-специалистам проанализировать внутренние параметры сайта, влияющие на его выход в ТОП-10.

Цель работы – разработать систему анализа факторов ранжирования web-ресурсов, влияющих на его продвижение в поисковых системах.

Актуальность работы обусловлена необходимостью автоматизировать процесс анализа сайтов, для повышения эффективности и уменьшения времени процесса. Разработанная система может быть использована для анализа факторов ранжирования web-ресурсов при проведении работ по поисковой оптимизации.

Объектом исследования являются web-сайты, а предметом исследования – факторы ранжирования, влияющие на выход в ТОП-10 в поисковых системах, таких как Яндекс и Google.

Разработка выполнена в среде Visual Studio 2012, язык программирования – C#. Для создания БД использовалась среда Microsoft SQL Server 2012. Пояснительная записка оформлена в Microsoft Word 2013.

## 1 Обзор предметной области

### 1.1 Выбор факторов ранжирования, влияющих на продвижение web-ресурсов в поисковых системах

Перед началом разработки системы анализа факторов ранжирования web-ресурсов необходимо выявить, какие параметры влияют на продвижение в поисковых системах, таких как Яндекс и Google. По этим параметрам и будет производиться анализ.

На производственной практике на основе изученной специализированной литературы [2-4] по продвижению сайтов в поисковых системах был проведен анализ факторов ранжирования, которые в большей степени влияют на поисковую выдачу. В результате чего были протестированы факторы ранжирования и выделены те, которые более эффективно работают при продвижении сайтов в поисковых системах (Рисунок 1.1).

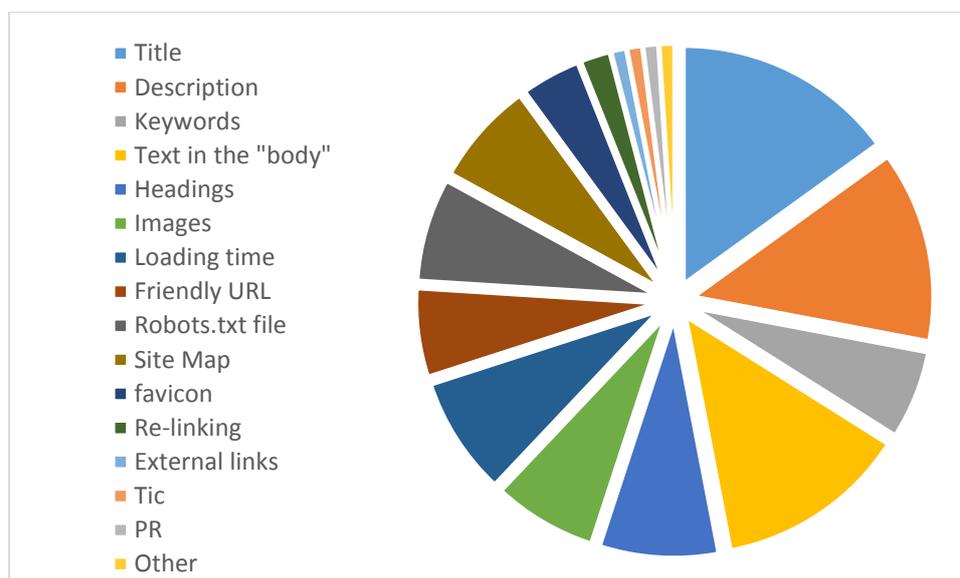


Рисунок 1.1 – Влияние факторов ранжирования

Главным принципом работы поисковых систем является определение релевантности содержимого страницы сайта ключевым словам, поэтому были выбраны факторы, которые на нее влияют. Кроме этого будут рассмотрены и другие важные факторы ранжирования.

Когда поисковые системы индексируют контент сайтов, они распределяют их по базам данных с разными ключевыми словами. Ключевые слова – представляют собой слова (или их набор), которые описывают тему сайта или текст, и используются поисковыми системами для обеспечения точного соответствия (релевантности) результатов поиска. Другими словами, вместо того, чтобы хранить 25 миллиардов страниц в одной базе данных, поисковые системы используют миллионы меньших баз, каждая из которых содержит сайты с определенным ключевым словом. Это позволяет поисковым системам находить релевантные страницы и выдавать результаты поиска за доли секунды [5].

Ключевые слова играют первостепенную роль в Интернет-поиске. Когда пользователь вводит запрос, поисковые системы подбирают страницы с соответствующими ключевыми словами, принцип работы поисковых систем изображен на рисунке 1.2.

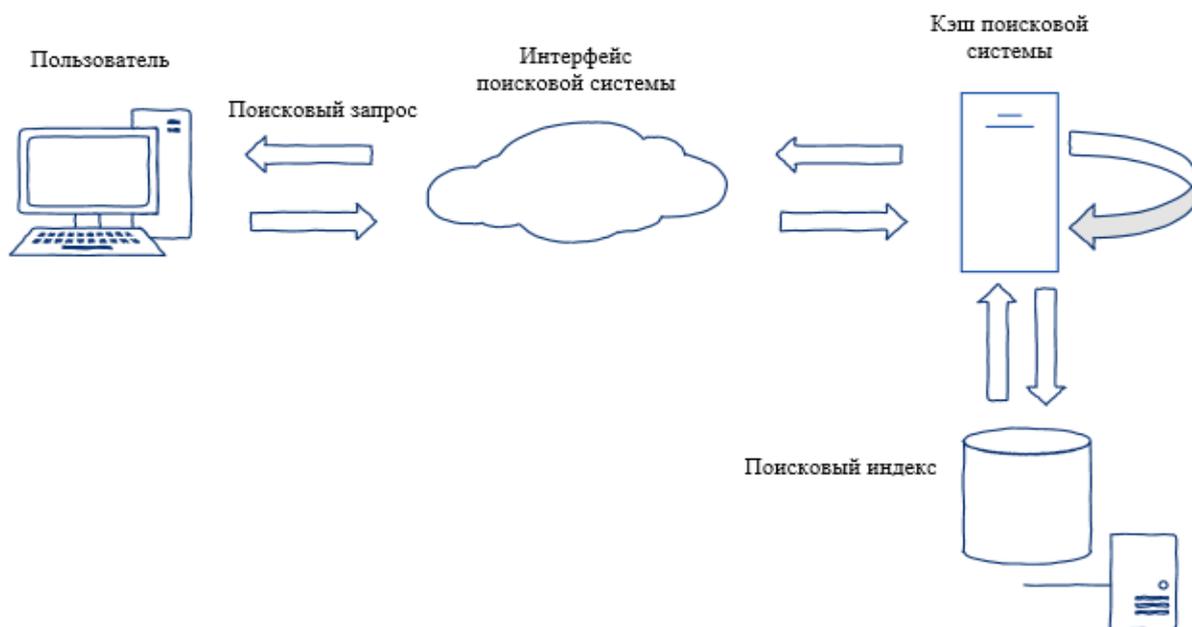


Рисунок 1.2 - Принцип работы поисковых систем

Пользователь запрашивает необходимую ему информацию, затем запрос отправляется в кэш поисковой системы, если ответ есть, он отдается

пользователю, если нет, то запрос посылается в основную базу данных, откуда поднимается ответ и отдается пользователю.

Порядок слов в запросе, пунктуация, использование дополнительных символов – все это помогает поисковым системам подобрать наиболее релевантные ответы. Чтобы определить релевантность страницы, поисковые системы анализируют, как ключевые слова использованы в тексте, так и другие составляющие страницы. Для оптимизации страницы под определенное ключевое слово необходимо использовать его в мета-тегах.

### 2.3.1 Мета-теги: Title, Description, Keywords

**Title** — самый важный мета-тег для внутренней оптимизации html-кода страниц. Поисковые системы придают очень большое значение тексту, заключенному в title, поэтому следует использовать этот фактор максимально эффективно.

Пример: `<title>Название страницы</title>`

Физически, содержание тега title – это название страницы, которое отображается в самом верхнем поле браузера. Также содержание title отображается в выдаче поисковых систем по запросам пользователей (это названия ссылок в результатах поиска). Поэтому заголовок страницы должен максимально ёмко отражать её суть и обязательно содержать ключевые слова. Для каждой страницы содержание *title* должно быть уникальным [6].

Поисковые системы могут учитывать не все слова, содержащиеся в названии страницы, а только определённое количество первых слов (остальные могут просто отсекаются). Для разных поисковых систем эти показатели отличаются, и к тому же они могут просто изменяться со временем при корректировке алгоритма, поэтому настоятельно рекомендуется располагать ключевые слова и фразы в начале заголовка.

Мета-тег **Description** предназначен для создания краткого описания страницы, которое выводится в качестве аннотации к ссылке на страницу в

поисковой выдаче. Однако в настоящее время это правило работает главным образом в поисковой системе Google, и то не во всех случаях. Довольно часто поисковые системы в качестве аннотации показывают просто отрывок текста данной страницы, наиболее релевантный запросу [7]. Тем не менее, Google официально рекомендует использовать мета-тег Description для описания страниц.

Информация оказывает ли содержимое тега Description влияние на релевантность страницы относительно ключевых слов, которые в нём содержатся, в какой-либо поисковой системе не подлежит распространению. Можно допустить, что это влияние присутствует. В любом случае наличие ключевых слов в Description желательно по приведенной выше причине.

**Keywords** – это еще один мета-тег, играющий важную роль в SEO. Пример мета-тега Keywords в html документе:

```
<meta name="keywords" content="Ключевые слово, ключевые слово...">
```

Мета-тег Keywords поисковые системы используют для определения релевантности, или соответствия, ссылки. Для правильного использования данного мета - тега необходимо применять только те слова, которые содержатся в самом документе. Использование тех слов, которых нет на странице, не рекомендуется. Список релевантных ключевых слов мета-тега Keywords может быть разделен запятыми или же вы можете использовать пробел [8]. Эти слова должны отражать наиболее важные темы сайта.

Мета-тег Keywords становится все менее и менее важным с точки зрения поисковой системы. На самом деле, можно получить высокие позиции, вообще ничего не включив в мета-тега Keywords. Тем не менее, чтобы обезопасить себя, желательно включить некоторые ключевые слова в данный мета-тег.

### 2.3.2 Текст на странице сайта

Текстовый контент сайта – основа любого качественного web-ресурса. Его читают посетители и поисковые роботы, он делает продажи и предлагает

услуги. С помощью текстового контента можно управлять действиями пользователей. Текст должен быть интересен и полезен пользователю, при этом количество вхождений и расстановка ключевых слов должны соответствовать требованиям поисковой системы.

В механизме ранжирования основой является оценка релевантности текста страницы запросу, который ввел пользователь. Для составления индексных баз и определения релевантности поисковые системы используют контент, текст запроса и некоторые элементы html-разметки. В качестве примера, описывающего этот процесс, можно рассмотреть следующую формулу (1.1):

$$R=PR*(T+L) \quad (1.1)$$

где R – значение релевантности;

T – степень соответствия внутренних критериев заданным поисковыми системами требованиям (текстовая релевантность);

L – уровень ссылочного ранжирования - степень соответствия текстов входящих ссылок на документ поисковому запросу (ссылочная релевантность);

PR – показатель внешних критериев документа, не зависящих от запроса (авторитетность ресурса).

Данная формула не является точным отображением работы современных алгоритмов ранжирования поисковых систем, однако даёт представление о значимости и видах критериев, из которых и складывается релевантность страниц [9].

### **2.3.3 Изображения на странице сайта**

Практически на каждом сайте есть картинки, фотографии, какие-то графические изображения или даже галереи изображений. Как и тексты, изображения требуют оптимизации и организации в дополнение к прочим элементам страницы. В результате, размещение и поисковая оптимизация изображений может дать плюс не только к релевантности страниц сайта и

повышению позиций в поисковой выдаче, но и плюс к трафику за счет появления изображений в выдаче сервисов картинок Google и Яндекс [10].

#### 2.3.4 Заголовки

Основные две цели заголовков в HTML-тегах H1-H6 - оформление и структуризация контента на страницах ресурса с целью сделать его более читаемым и легко потребляемым посетителями сайта. Именно по заголовкам Google и Яндекс определяют, насколько тематика страницы совпадает с запросами, которые вводят пользователи. Иными словами, насколько релевантна данная страница. [11]

В идеале на странице должен присутствовать всего один заголовок первого уровня, который как бы содержит в себе два-три заголовка второго уровня в которых в свою очередь "вложены" заголовки рангом поменьше и так далее. Структура документа должна иметь примерно такой вид (Рисунок 1.3):

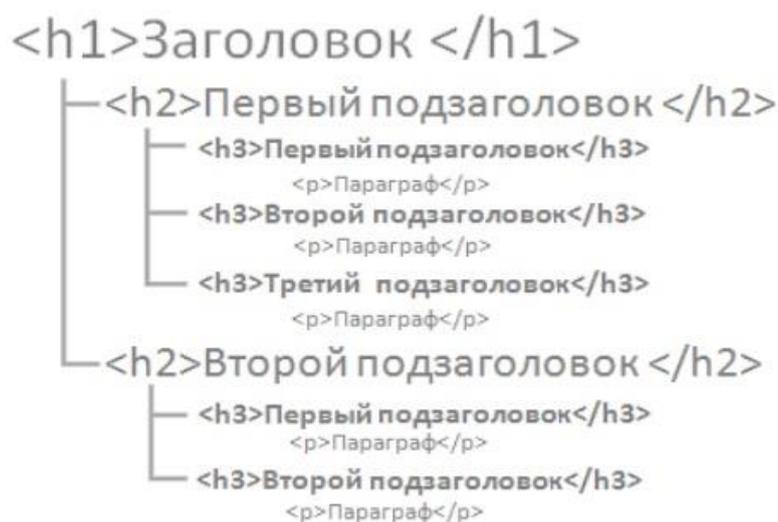


Рисунок 1.3- Иерархия заголовков

То есть заголовки должны быть, как бы вложены друг в друга по степени важности и иметь древовидный вид. Сам же текст желательно разбить на параграфы с помощью тега <p> - где текст каждого абзаца несет в себе одну логическую часть текста.

Такая разметка текста легко "усваивается" поисковыми системами и другими программами, работающими с HTML документами, например, позволяет выхватить один из параграфов для описания сайта в результатах поиска, не ломая голову где начинается и где заканчивается определённая логическая часть текста или учитывая уровни заголовка автоматически составить содержание документа.

Так как текст в заголовках <h1>- <h6> имеет куда более большее значение для поисковых систем, да и просто для человека, просматривающего страницу, нежели чем просто текст то их следует наполнять ключевыми словами.

### **2.3.5 Favicon**

Favicon (англ. Favorite Icon) - это иконка в специальном формате, которая отображается в адресной строке, при посещении какого-либо сайта, например, этого, вкладке браузера или в его закладках ("Избранном"). Обычно она хранится в корневой папке сайта и называется favicon.ico. [12]

Использование этой иконки имеет множество положительных сторон. Во-первых, эта иконка запоминается посетителям порою лучше, чем название сайта. Во-вторых, favicon делает сайт уникальным, выделяя сайт среди прочих сайтов, которые не успели обзавестись иконкой.

Некоторые поисковые системы отображают этот favicon для каждого сайта в результатах поиска. Таким образом, это один из рычагов управления кликабельности сниппета. Изображение всегда привлекает внимание пользователя и, даже если сайт будет не на первых позициях в результатах поиска, к вам перейдут больше посетителей, чем на сайт без favicon.

### **2.3.6 Файл Robots.txt**

Файл robots.txt предоставляет важную информацию для поисковых роботов, которые сканируют интернет. Перед тем как пройти по страницам сайта, поисковые роботы проверяют данный файл.

Это позволяет им с большей эффективностью сканировать сайт, так как вы помогаете роботам сразу приступать к индексации действительно важной информации на сайте (это при условии, что вы правильно настроили robots.txt). [13]

### **2.3.7 Карта сайта**

**Карта сайта** (Sitemap) – инструмент пользовательской навигации по сайту. Карта сайта представляет собой полный каталог разделов ресурса. Карта сайта помещается, как правило, на главной странице сайта с тем, чтобы пользователь в короткий срок и без затруднений перешёл к нужному ему разделу сайта. [14]

Карта сайта нужна для реализации двух целей. Во-первых, сайт, снабжённый простой и информативной картой, пользуется популярностью среди пользователей. В Интернете очень важным параметром ресурса является уровень пользовательского доверия. Сайт, на котором размещена карта, без сомнений, доверие такое вызывает. Во-вторых, чем может эффективно воспользоваться оптимизатор, наличие на сайте карты высоко котируется поисковыми машинами. Удобная пользовательская навигация – путь к успешному продвижению сайта в поисковых системах.

### **2.3.8 Время загрузки страницы**

Поисковики обращают внимание на время загрузки страницы сайта по одной довольно простой причине: чем больше медленных сайтов попадает в

ТОП выдачи конкретной поисковой системы, тем реже пользователи будут ее использовать.[15]

Ускорение работы сайта приводит к увеличению количества просмотров страниц посетителем ресурса в течение одной сессии. Практически восемьдесят процентов пользователей утверждают, что в случае, если страница загружается больше четырех секунд, вероятность того, что они вернутся на данный ресурс, ничтожно мала. [16] Средняя продолжительность загрузки страницы по мнению интернет сообществ должна составлять не более двух секунд.

Получается, чтобы не уступить конкурентной поисковой системе, нужно на первой странице выдачи отображать только те результаты, которые имеют оптимальное время загрузки сайта и не вынуждают посетителей необходимостью долго ждать.

### **2.3.9 Человекопонятные URL**

Человекопонятный URL (от англ. friendly url) — уникальный адрес веб-страницы в глобальной сети, кратко отображающий ее содержание, понятный и удобный для восприятия пользователя. ЧПУ являются одним из инструментов SEO-оптимизации и положительно влияют на ранжирование страниц в поисковой выдаче. Это обусловлено тем, что, как правило, человекопонятные URL содержат ключевые слова запроса, к чему положительно относятся поисковики. Например, введя в Google запрос «натяжные потолки», одним из первых результатов который вы увидите, будет ссылка на страницу с человекопонятным URL-ключевиком.

Грамотно организованная структура веб-адресов с человекопонятными URL значительно помогает людям находить интересующие их страницы, и чаще всего именно на информацию с таких сайтов пользователи оставляют ссылки. В то время как длинные URL с малосимпатичным содержанием отпугивают посетителей. [17]

Стоит отметить, что делать ссылки человекопонятными следует в первые дни создания сайта. Последующее изменение внешнего вида ссылок чревато вылетом страниц из поиска на определенный период времени — ведь поисковым машинам нужно будет заново проиндексировать измененные ссылки, старые удалить, а новые записать.

## 1.2 Определение эталонных характеристик параметров

По каждому фактору ранжирования существуют свои рекомендации от поисковых систем. Для Яндекса они располагаются на сервисе Яндекс.Помощь [19], для поисковой системы Google - Search Console [20]. Рассмотрим все способы и выберем основные рекомендации, которые будут использоваться в разрабатываемой системе для анализа web-ресурсов.

Рекомендации по составлению title:

1. Title каждой страницы должен быть уникальным.
2. Располагать ключевую фразу лучше ближе к началу заголовка.
3. Если ключевых слов, по которым продвигается страница, несколько — следует объединить их в Title без ущерба для читабельности.
4. Нельзя использовать в Title простое перечисление ключевых слов — это может стать поводом для наложения санкций поисковых систем.
5. Не рекомендуется употреблять одно и то же слово более 2-х раз. Но использование «ключевика» дважды (в разных словоформах) даже желательно.
6. Рекомендуется не включать в Title стоп-символы / \ | ( ) + =. А если это необходимо, то лучше использовать их после ключевой фразы, а не до неё.
7. Желательно обходиться без стоп-слов (предлогов, местоимений, междометий, вводных словесных конструкций и т.д.).
8. Длина Title должна быть максимум 80 символов с пробелами, но лучше придерживаться оптимальных границ (40–65 знаков).
9. Точка в конце заголовка не ставится.

10. Вопросительный знак в Title лучше не ставить, даже если ключевая фраза требует этого, иначе пользователи могут решить, что вы задаете вопрос, а не даете ответ.

11. Title написанный прописными буквами может быть расценен как спам. Поэтому следует использовать прописные буквы только для первых букв предложений.

Правила и рекомендации по составлению Description:

1. Как и Title, Description каждой страницы должен быть уникальным.

2. Description пишется специально для каждой интернет-статьи, а не копируется из текста страницы.

3. Оптимальный объем описания веб-документа – 130–180 символов с пробелами. Максимальный – 200, хотя иногда встречаются в выдаче описания и по 240 знаков.

4. Если запросов для веб-страницы несколько, лучше использовать в Description самые высокочастотные.

5. Ключевую фразу рекомендуется использовать 2 раза: в точном вхождении и в морфологическом или разбавленном.

6. Вписывайте в Description синонимы слов из ключевой фразы. Посмотрите в Яндексe, какие слова он подсвечивает вместе с теми, что входят в запрос. Например, «купить» – «продажа», «купить» – «заказать» и т.д.

Рекомендации по составлению Keywords:

1. Количество слов в теге Keywords не более 10.

2. Для правильного использования данного мета - тега необходимо применять только те слова, которые содержатся в самом документе. Использование тех слов, которых нет на странице, не рекомендуется.

3. Список ключевых слов должен быть разделен запятыми или пробелами.

4. Ключевые слова не должны повторяться более 3 раз, включая различные варианты написания.

Указания к текстам:

1. Уникальность текста больше 90% Уникальность текста – это характеристика, с помощью которой можно определить, существуют ли в интернете частичные или полные копии заданного документа. Чем меньше в статье отрывков, которые встречаются на других ресурсах, тем выше процент ее уникальности. [18]

2.Текст должен быть объемом около 2000 символов с пробелами.

4.Текст должен содержать как минимум один маркированный или нумерованный список.

5. В тексте должны присутствовать заголовки 2-3 на раздел: должны присутствовать ключевые фразы с точным вхождением по одному разу и разбавленным вхождением по одному разу.

6. Текст должен содержать по два точных вхождения ключевых и по одному разбавленному вхождению.

7.В тексте необходимо использовать словоформы и однокоренные слова ключевых слов.

8. Плотность ключевых слов должна составлять не больше 6%.

Рассмотрим требования к изображениям:

1. Заголовок и описание – текстовое описание и заголовок позволяют гарантировать то, что поисковые роботы поймут, о чем эта картинка.

2. Размер – поисковые системы заботятся о времени загрузки, так что вы должны убедиться в том, что изображение соответствующего размера.

3. Имя файла – это простая и важная деталь. Имя изображения должно иметь непосредственное отношение к нему же и включать ключевые слова.

4. Форматирование – как правило, лучше загружать на сайт фотографии JPG формата. Это поможет поисковым системам находить и индексировать их правильно. Если использование JPG формата не представляется возможным, придерживайтесь таких форматов как GIF или PNG.

5. Ссылка – можно ссылаться на изображение к статье. Подобно тому, как используются ссылки на статью, вы можете сделать то же самое, распространяя ссылку на изображение.

Заголовки h1-h6 должны быть составлены по следующим правилам:

1. На странице сайта должен присутствовать только один заголовок h1 и находится в начале страницы.

2. В заголовках необходимо использовать ключевые слова. Чем выше значимость ключевого слова, тем выше должен быть уровень заголовка.

3. Заголовки должны быть информативными и при этом не очень длинными.

4. Указывать важные ключевые фразы необходимо как можно ближе к началу заголовка.

5. Ключевые фразы в подзаголовках старайтесь составлять с учетом морфологии и синонимов.

При составлении файла robots.txt следует придерживаться некоторых советов:

1. Название файла необходимо написать только прописными буквами (например, имена Robots.txt или ROBOTS.TXT будут неправильными).

2. Для robots.txt есть ограничение по весу файла у Google это до 500 кб, у Яндекса до 32 кб.

3. Файл robots.txt должен находиться в корневой папке сайта.

Рекомендации по составлению карты сайта:

1. Карта сайта должна показывать иерархию всего сайта. Необходимо разбить карту на удобные разделы, пункты, лучше использовать списки и заголовки.

2. Файл карты сайта должен находиться в корне сайта.

3. В карте сайта не должны присутствовать ссылки, которые закрыты для индексирования в файле robots.txt.

4. В карте сайта указываются приоритеты индексации страниц сайта от 0,1 для самой маловажной страницы до 1 для главной страницы.

Для человеко-понятных URL основные правила использовать ключевые слова, быть читаемыми для посетителя сайта и поискового робота.

Рекомендаций по составлению favicon особых нет, главное его наличие и размер 16 на 16 пикселей.

Скорость загрузки сайта должны составлять 2 секунды. При увеличении времени загрузки увеличивается и шанс, что пользователь покинет страницу до того, как она загрузится. То есть сайт потеряет потенциального клиента, а значит он будет неэффективным.

Изучив рекомендации, определим эталонные значения факторов ранжирования, которые будем использовать для анализа.

Одним из важных составляющих страницы сайта являются текстовые параметры, так как они влияют на релевантность страницы поисковому запросу (Таблица 1.1).

Таблица 1.1 Текстовые параметры

	<b>Количество символов</b>	<b>Вхождение запроса</b>
<b>Title</b>	40-80	1
<b>Description</b>	130-180	1-2
<b>Keywords</b>	до 10 слов	2-3
<b>Текст</b>	2000-3000	3-4
<b>Абзац</b>	500-100	1-2
<b>Заголовки h1-h6</b>	до 10 слов	1-2

В текстовых факторах таких как: Title, Description, Keywords, текстах на страницах, абзацах текста и заголовках необходимо наличие слов, входящих в поисковый запрос. Кроме того, заголовок H1 должен быть один на странице и содержать основной запрос.

На странице сайта должны присутствовать 2-5 изображения, которые должны иметь в названии ключевой запрос (например, если запрос «натяжные потолки», то изображение может называться natyajnie-potolki.jpg). В заголовке и альтернативном тексте изображения также должны быть прописаны ключевые запросы.

Следующим фактором ранжирования сайта было выделено время загрузки сайта. Сайт должен загружаться не более 2 секунд, так как если сайт будет долго загружаться посетители будут закрывать сайт, не дожидаясь полной загрузки.

Для корректной индексации поисковыми роботами на сервере сайта должен быть файл robots.txt, в котором указано главное зеркало сайта, и информация о том, какие страницы открыты для индексации, а какие закрыты. При отсутствии этого файла или при неправильном его заполнении поисковые роботы могут неправильно выбрать зеркало сайта и проиндексировать все страницы сайта. Таким образом в поисковой выдаче могут оказаться административные и приватные страницы сайта.

Так же для более быстрой индексации новых страниц сайта поисковыми роботами необходима карта сайта, в ней не должны находиться закрытые в файле robots.txt страницы сайта. У каждой страницы должен быть указан приоритет от 0,1 до 1.

Для повышения CTR (кликабельности) сайта нужен favicon. Он позволяет улучшить узнаваемость сайта, а значит увеличить лояльность покупателя. Размеры favicon 16 на 16 пикселей.

Следующим пунктом является ЧПУ (человекопонятные URL). Чтобы поисковые системы понимали, о чем страница сайта в них должны присутствовать ключевые запросы.

### **1.3 Обзор существующих систем**

Существует большое количество инструментов для анализа продвижения сайтов в поисковых системах, они предоставляют разнообразные сервисы и программы. Эти системы можно разделить на несколько групп:

1. Анализ текста: [advego.ru](http://advego.ru), [text.ru](http://text.ru), [pr-cy.ru/zypfa/](http://pr-cy.ru/zypfa/)
2. Анализ позиций сайта: [allpositions.ru/](http://allpositions.ru/), [seogadget.ru/serppos](http://seogadget.ru/serppos), [line.pr-cy.ru/](http://line.pr-cy.ru/)
3. Анализ ссылочной массы: [linkpad.ru](http://linkpad.ru), [ru.ahrefs.com](http://ru.ahrefs.com), [rookee.ru](http://rookee.ru)
4. Анализ внутренних параметров, без рекомендаций по оптимизации: [seolib.ru](http://seolib.ru), [SiteAuditor](http://SiteAuditor), [analizsaita.com](http://analizsaita.com)
5. Анализ внутренних и внешних параметров, без рекомендаций по оптимизации: [xtool.ru](http://xtool.ru), [sbup.com](http://sbup.com)
6. Анализ внутренних параметров, с рекомендациями по оптимизации: [WebSite Auditor](http://WebSite Auditor), [iAnalysis](http://iAnalysis)

Разрабатываемая система относится к 6 группе, то есть она может провести анализ внутренних параметров и предоставить информацию о необходимых поправках на сайте.

Рассмотрим подробнее интересующие нас системы:

- Система WebSite Auditor позволяет производить анализ большого количества разных факторов ранжирования, таких как мета-теги, изображения, текст, файл robots.txt. Анализ страниц происходит с учетом конкретных ключевых слов. Данный сервис не анализирует карты сайта, скорость загрузки, наличие ЧПУ и favicon. Программа является платной - стоимость лицензии в России около 124\$.

- Сервис iAnalysis представляет собой сервис, анализирующий PR и ТИЦ, индексацию сайта в поисковых системах, ссылки на сайт, наличие в крупных каталогах, размер страницы, IP и URL сайта. Анализ страниц происходит без учета ключевых слов. Этот сервис проводит анализ не основных параметров, влияющих на продвижение сайта в поисковых системах, но он позволяет сохранять результаты анализа и является бесплатным.

В процессе изучения систем WebSite Auditor и iAnalysis было выявлено, что сохранение анализов происходит на устройство, на котором открыта программа, что не дает возможности открыть отчет из программы на другом устройстве. Это усложняет процесс работы, так как во многих компаниях над оптимизацией сайтов работает несколько человек. Для удобства коллективного использования и использования на разных устройствах необходимо, чтобы отчеты сохранялись в учетной записи пользователя.

Таким образом были выделены слабые стороны аналогичных систем, которые будут учтены при разработке собственной системы. В результате разрабатываемая система должна предоставлять рекомендации по всем основным факторам ранжирования сайта, быть бесплатной, предоставлять возможность создать учетную запись пользователя, сохранять отчеты в системе и производить сравнение отчетов.

## 2 Разработка информационной системы

### 2.1 Функциональные требования к системе

После проведения анализа предметной области были выделены следующие команды и задачи, которые разрабатываемая система должна корректно выполнять:

- по требованию пользователя проводить регистрацию или авторизацию пользователя в системе;
- выводить Title, Description, Keywords анализируемой страницы сайта;
- определять наличие или отсутствие Title, Description, Keywords, заголовков, изображений, тегов Title и Alt изображений, текста, маркированных или нумерованных списков в тексте на странице сайта, файла robots.txt и карты сайта на сайте;
- определять наличие или отсутствие ключевых фраз в Title, Description, Keywords, заголовках, тегах Title и Alt изображений, тексте на странице сайта;
- определять количество символов в Title, Description, Keywords, тексте и абзацах, количество абзацев в тексте на странице сайта;
- определять количество вхождений ключевых слов в Title, Description, Keywords, тексте и абзацах, количество абзацев в тексте на странице сайта;
- определять скорость загрузки сайта;
- определять наличие человеко-понятных URL и favicon;
- определять корректность составления файла robots.txt и карты сайта;
- по запросу пользователя предоставлять подробный отчет об анализе страницы сайта;
- по запросу пользователя предоставлять сравнение подробных отчетов об анализе страницы сайта;
- сохранять подробный отчет об анализе;
- предоставлять справку о работе с программой по запросу пользователя;

- иметь удобный пользовательский интерфейс.

Управление должно осуществляться выбором соответствующих операций в окне программы манипулятором «мышь».

## **2.2 Выбор платформы для реализации информационной системы**

В качестве среды разработки системы сравнительной оценки эффективности web-ресурсов была использована Visual Studio 2012. Выбор данного программного обеспечения обусловлен следующими предоставляемыми возможностями:

- Создание интерфейса с помощью Windows Forms;
- Работа с объектами баз данных в обозревателе объектов SQL Server;
- Запись асинхронного кода простым и интуитивно понятным способом;
- Получение сведений о вызывающем объекте, которые помогают с трассировкой и отладкой;
- Возможность работать с данными из сети Интернет;
- Возможность регистрации и авторизации пользователей.
- В качестве языка программирования использован язык C#, так как он обладает следующими преимуществами:
  - Объектно-ориентированность;
  - Ориентация на безопасность кода;
  - Многофункциональность;
  - Относительная простота.

Кроме того, нет необходимости тратить дополнительное время на обучение работе с ПО, так как оно было изучено в процессе обучения.

Для работы с данными в форматах HTML и PDF была использована программная платформа Microsoft .NET Framework.

Для создания БД использовалась среда Microsoft SQL Server 2012 Management Studio, за счет следующих достоинств:

- Удобный графический интерфейс;
- Богатые возможности;
- Высокая производительность;
- Надежность.

## **2.3 Архитектура системы**

Для лучшего понимания архитектуры разрабатываемой системы рассмотрим ее обобщенный алгоритм работы (Приложение А). После запуска пользователем системы ему предоставляется возможность проводить только экспресс-анализ сайта. Если пользователь зарегистрирован, он может авторизоваться в системе. Если пользователь не зарегистрирован он может зарегистрироваться, а затем авторизоваться. После авторизации пользователю становятся доступны все функции системы: экспресс-анализ, подробный анализ, экспорт результатов подробного анализа, открытие сохраненных в приложении результатов, сравнение результатов подробного анализа.

### **2.3.1 Диаграмма потоков данных**

Проектирование системы факторов ранжирования web-ресурсов было начато с построения диаграммы потоков данных, которая позволяет продемонстрировать, как каждый процесс преобразует свои входные данные в выходные, а также выявить отношения между этими процессами. На рисунке 2.1 представлена диаграмма потоков данных программной системы, на которой изображено взаимодействие компонентов системы. Разрабатываемая система состоит из четырех подсистем:

1. Подсистема сбора данных.
2. Подсистема обработки данных.
3. Подсистема формирования отчетов.
4. Подсистема идентификации пользователей.

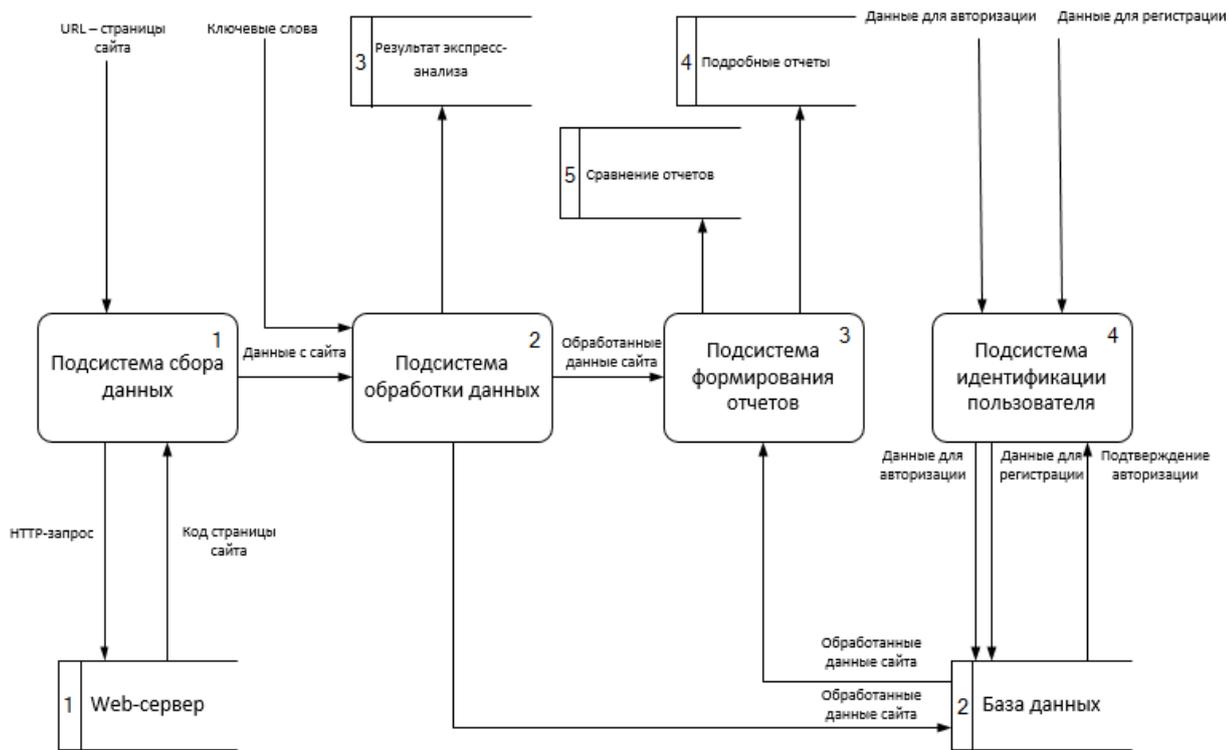


Рисунок 2.1 - Диаграмма потоков данных

Для работы подсистемы сбора данных пользователю требуется ввести URL страницы сайта и ключевые запросы, затем подсистема отправляет HTTP-запрос для получения исходного HTML-кода страницы, web-сервер на запрос подсистемы предоставляет HTML-код страницы сайта. Далее извлекает информацию, которая потребуется для анализа факторов ранжирования, из HTML-кода страницы сайта.

Подсистема обработки данных использует полученную в подсистеме сбора данных информацию для выделения необходимых факторов ранжирования, с учетом ключевых слов, и сравнивает их с эталонными значениями, которые были определены ранее. Итогом работы этой подсистемы является передача данных о факторах ранжирования анализируемой страницы сайта в базу данных и вывод результата экспресс-анализа.

Подсистема формирования отчетов, на основе информации из базы данных о факторах ранжирования, формирует подробный отчет с рекомендациями по их изменению для приведения к эталонным значениям. По

желанию пользователя возможно сравнение сохраненных отчетов и экспорт отчетов в формате PDF. Функции подсистемы формирования отчетов доступны только авторизованному в подсистеме идентификации пользователю.

Подсистема идентификации пользователей позволяет пользователю создать учетную запись в системе для получения расширенных функций системы. Информация о пользователе, его логин и пароль, хранятся в базе данных.

### 2.3.2 Взаимодействие системы с источниками данных

Основная особенность взаимодействия системы анализа факторов ранжирования web-ресурсов с источниками данных заключается в том, что система использует данные для проведения анализа из HTML-кода страницы сайта. Чтобы получить HTML-код страницы, система отправляет HTTP-запрос в web-сервер. Далее система получает HTML-код запрашиваемой страницы сайта (Рисунок 2.2).

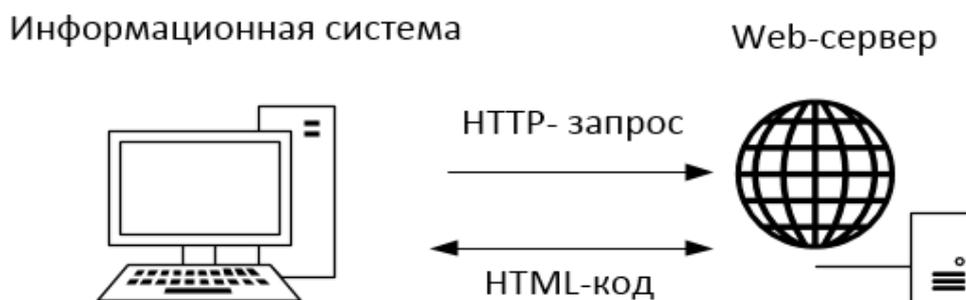


Рисунок 2.2 – Получение HTML-кода страницы с web-сервера

Далее необходимо выделить необходимые данные из HTML-кода страницы для дальнейшего анализа. Для реализации этого процесса была использована программная библиотека HtmlAgilityPack платформы Microsoft .NET Framework. С помощью библиотеки HtmlAgilityPack и классов HttpRequest и HttpResponse производится считывание необходимых данных из HTML-кода страницы сайта для дальнейшей обработки. Класс

HttpRequest позволяет отправлять HTTP-запросы, а класс HttpResponse позволяет обрабатывать полученные ответы. Библиотека HtmlAgilityPack – читает HTML, позволяет делать поиски по сайту, выборки и строит DOM дерево. С их помощью происходит сбор информации с сайтов, и сохранение этой информации в удобном виде - парсинг.

Взаимодействие системы с источниками данных происходит в подсистеме сбора данных. На рисунке 2.3 изображено общее представление этапа сбора данных.

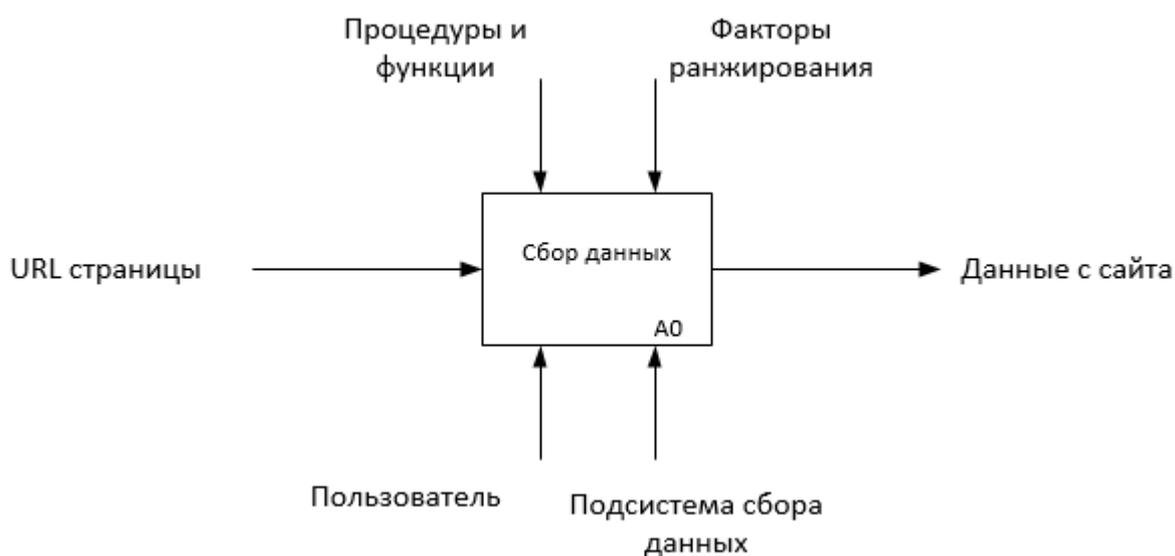


Рисунок 2.3 – Общее представление этапа сбора данных

Рассмотрим детально процессы подсистемы сбора данных (Рисунок 2.4). В этой подсистеме происходит парсинг, в результате которого система получает данные с сайта для дальнейшего проведения анализа.

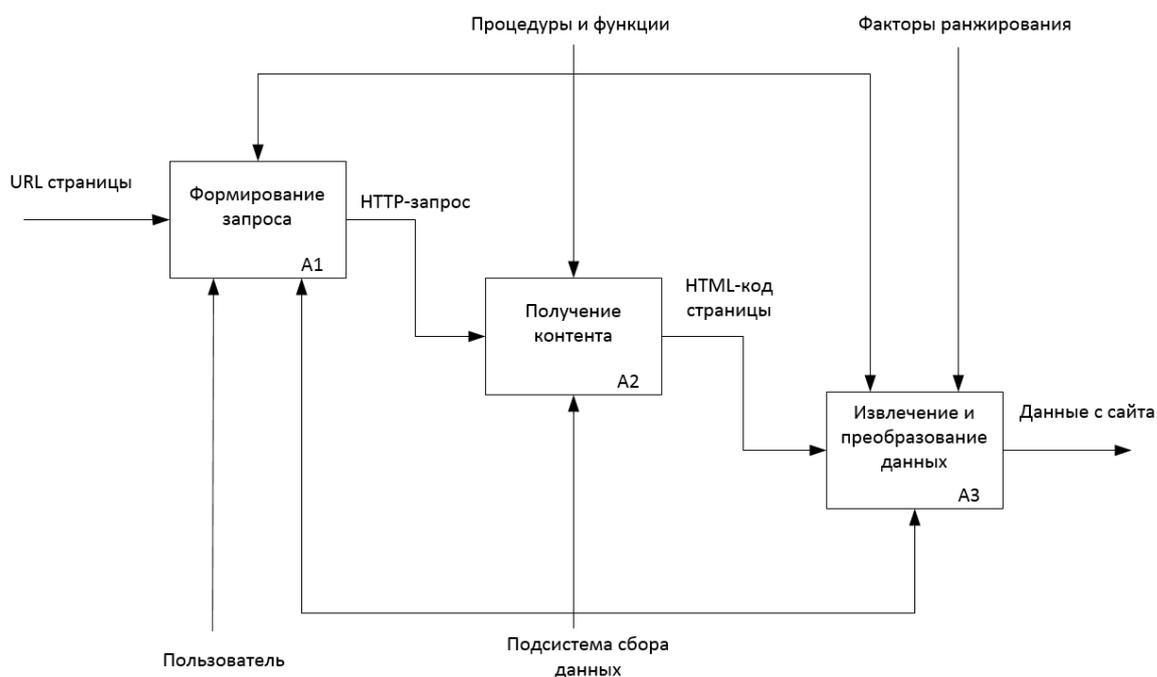


Рисунок 2.4 – Декомпозиция блока «Сбор данных»

Подсистема состоит из трех частей, которые отвечают за три отдельных процесса парсинга:

1. Формирование запроса. На основе полученного от пользователя URL страницы сайта составляется HTTP-запрос, который используется для получения контента.

2. Получение контента в исходном виде. Под получением контента, подразумевается скачивание кода веб-страницы, из которой необходимо извлечь данные или контент;

3. Извлечение и преобразование данных. В этой фазе происходит извлечение требуемых данных из полученного, на первом этапе, кода страницы. Для извлечения использовались регулярные выражения. Также на этом этапе происходит преобразование извлеченных данных к нужному виду. Далее данные отправляются в базу данных и в подсистему обработки данных.

### 2.3.3 Обработка данных

Полученные данные о факторах ранжирования со страницы сайта проходят определение на наличие; подсчет количества символов, вхождений ключевых слов, скорости загрузки; определение корректности. Обработка данных происходит в подсистеме обработки данных, на рисунке 2.5 представлен обобщенный вид процессов, происходящих в этой подсистеме.

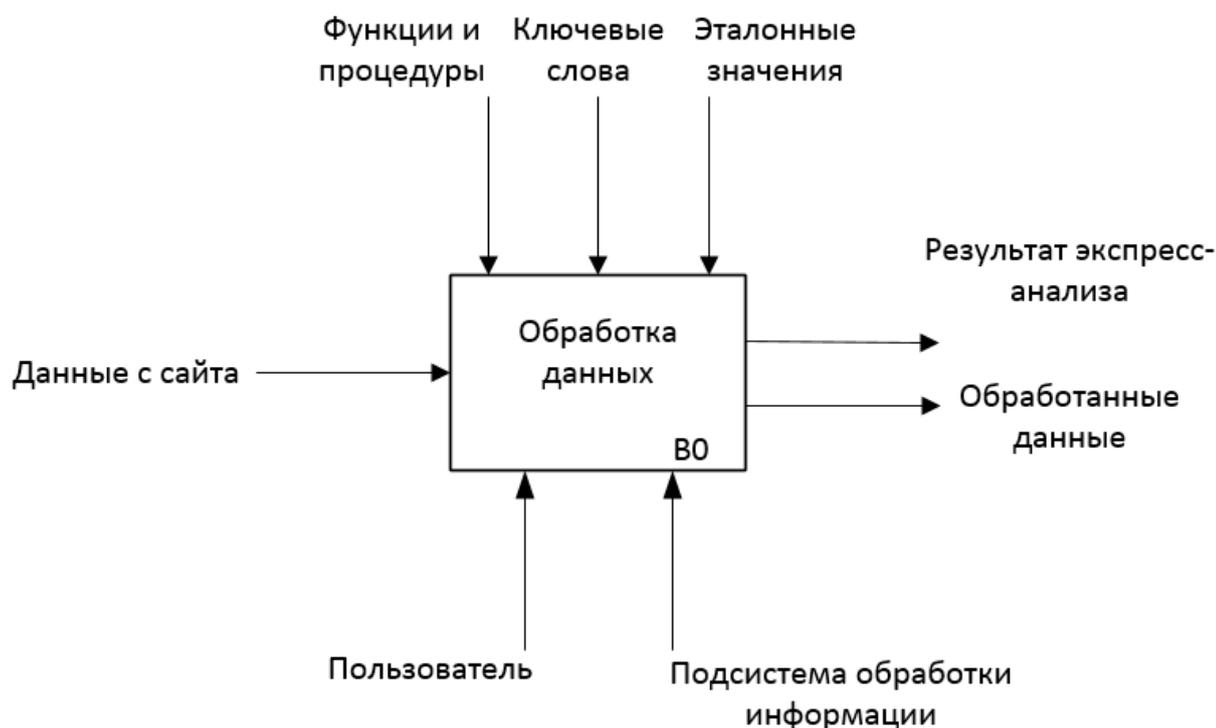


Рисунок 2.5 - Общее представление этапа обработки данных

Для данного этапа входной информацией являются «Данные с сайта», а в качестве механизма служит сама «Подсистема обработки информации» и «Пользователь». Так же были выделены следующие элементы управления: «Функции и процедуры», «Ключевые слова» и «Эталонные значения». Итогом этого этапа является «Результат экспресс-анализа» и «Обработанные данные».

В качестве примера разберем процесс обработки текста страницы сайта (Рисунок 2.6).

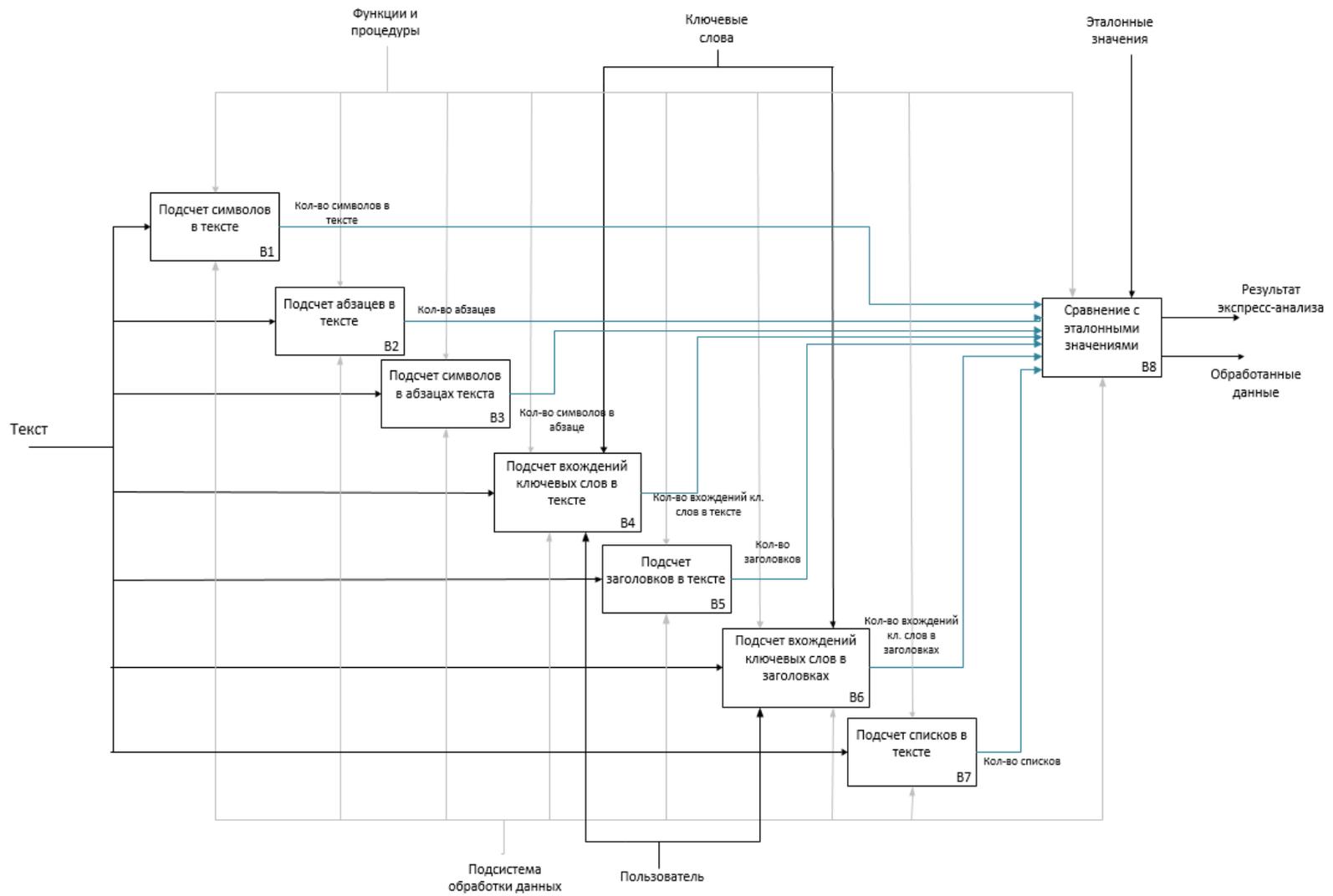


Рисунок 2.6 – Процесс обработки текста

Из текста с помощью соответствующих процессов выделяются необходимые для проведения анализа значения:

- Количество символов в тексте.
- Количество абзацев в тексте.
- Количество символов в абзацах.
- Количество вхождения ключевых слов в тексте.
- Количество заголовков в тексте.
- Количество вхождений ключевых слов в заголовках.
- Количество списков в тексте.

Далее полученные значения факторов ранжирования сравниваются с эталонными значениями этих факторов. Подобным образом происходит обработка остальных данных, полученных с сайта. После завершения процесса обработки всех данных пользователю выводится результат экспресс-анализа, а также происходит передача обработанных данных, используемых в дальнейшем для формирования подробных отчетов и сравнения двух подробных отчетов, в базу данных.

#### **2.3.4 Формирование отчетов**

На основе обработанных данных происходит формирования подробных отчетов с рекомендациями по проведению работ по оптимизации страницы сайта. Обработанные данные – это данные с сайта, которые прошли сравнение с эталонными значениями. Они состоят из исходных данных (такие как: Title, Description, Keywords, время загрузки, файл robots.txt и карта сайта), а также из информации о соответствии или несоответствии их эталонным значениям. Для формирования отчета информация о несоответствии какого-либо фактора ранжирования сопоставляется с рекомендацией по его изменению для приведения к эталонному значению.

Например, если количество символов в Title меньше 40, обработанное значение будет «Мало символов», соответственно этому значению выводится рекомендация «Следует увеличить количество символов до 40». Если в Title нет вхождений ключевых запросов, то в отчете выводится «Добавьте ключевые слова».

Аналогичным образом происходит формирование рекомендаций по остальным обработанным данным (Таблица 2.1), в результате чего пользователю предоставляется полный список советов по приведению страницы сайта к соответствию требованиям поисковых систем.

Таблица 2.1 – Соответствие рекомендаций обработанным данным

Обработанные данные		Рекомендации
Title	Мало символов	Следует увеличить количество символов до 40.
	Много символов	Следует уменьшить количество символов до 80.
	Нет ключевых слов	Добавьте одну ключевую фразу.
	Отсутствует	
	Переоптимизирован	Уменьшите количество вхождений ключевых запросов до одного.
Description	Мало символов	Следует увеличить количество символов до 130.
	Много символов	Следует уменьшить количество символов до 180.
	Нет ключевых слов	Добавьте 1-2 ключевые фразы.
	Отсутствует	
	Переоптимизирован	Уменьшите количество вхождений ключевых запросов до одного.
Keywords	Много слов	Следует уменьшить количество слов до 10.
	Нет ключевых слов	Добавьте 1-2 ключевую фразу. Чтобы количество символов было не больше 10.
	Отсутствует	
	Переоптимизирован	Уменьшите количество вхождений ключевых запросов до 1-2.
	Много символов	Уменьшите количество символов до 2000.

Текст	Мало символов	Увеличьте количество символов до 1000,
	В тексте нет списков	Добавьте минимум один список.
	Мало ключевых фраз	Увеличьте количество ключевых слов до трех.
	Много ключевых фраз	Уменьшите количество ключевых слов до четырех.
Абзац	Много символов	Уменьшите количество символов до 1000.
	Мало символов	Увеличьте количество символов до 5000.
Заголовки	Заголовок Hn отсутствует	Напишите заголовок Hn с использованием одной ключевой фразы.
	Заголовок H1 больше одного	Уменьшите количество заголовков H1 до одного.
	Заголовок Hn не содержит ключевые слова	Добавьте одно вхождение ключевого слова.
Изображения	Title не прописан	Пропишите в Title ключевую фразу.
	Alt не прописан	Пропишите в Alt ключевую фразу.
	В названии изображения нет ключевых слов	Переименуйте изображение с использованием ключевых слов.
	На странице 1 изображение.	Увеличьте количество изображений от 2 до 5.
	В Title нет ключевых слов	Пропишите в Title ключевую фразу.
	В Alt нет ключевых слов	Пропишите в Alt ключевую фразу.
Время загрузки	Большое время загрузки	Уменьшите вес страницы сайта. Уберите тяжелые анимации, уменьшите расширение изображений.
Человеко-понятные URL	ЧПУ не прописаны	Пропишите ЧПУ с использованием одной ключевой фразы.
	ЧПУ не содержит ключевых фраз	Пропишите одну ключевую фразу.
favicon	Отсутствует	Добавьте favicon на сайт.

Файл robots.txt	Отсутствует	Добавьте файл robots.txt.
	Не указано главное зеркало сайта	Укажите главное зеркало сайта.
Карта сайта	Отсутствует	Добавьте карту сайта
	Не корректна	Исключите закрытые в файле robots.txt страницы.

### 2.3.5 Структура базы данных

Все учетные записи и обработанные данные с сайтов сохраняются в базе данных системы (Рисунок 2.7). База данных состоит из 5 таблиц со своими наборами атрибутов:

1. Таблица Client описывает зарегистрированного пользователя и содержит атрибуты логин, пароль и имя.

2. Таблица Site содержит информацию о домене сайта, который имеет URL главной страницы, логин пользователя, favicon, информацию о robots.txt и sitemap.xml.

3. Таблица SitePage определяет страницу сайта, хранит URL главной страницы и анализируемой страницы сайта и набор ключевых слов.

4. – 5. Таблица Report1 и Report2 содержат URL анализируемой страницы сайта, идентификатор сессии и набор ключевых слов, главными ключами являются idSession1 и idSession2 соответственно. Эти сущности необходимы для реализации сравнения двух отчетов, поэтому атрибуты urlPage и serchKeywords должны иметь одинаковые значения.

6. Таблица Report представляет собой набор данных о проведенном анализе: идентификатор сессии, дата и время сессии, ключевые слова; Title, Description, Keywords страницы сайта; информацию о тексте, заголовках, изображениях, находящихся на странице сайта; информацию о наличии человеко-понятных URL и favicon; файл robots.txt и карту сайта, а также время загрузки страницы.

У каждой сущности есть свой уникальный составной идентификатор, который позволяет однозначно определить элементы сущностей, а также помогает предотвратить дублирование данных. Кроме того, все сущности связаны с помощью бинарных связей. В данном случае используется связь один ко многим, например, у одного пользователя может быть несколько сайтов, а у сайта несколько страниц.

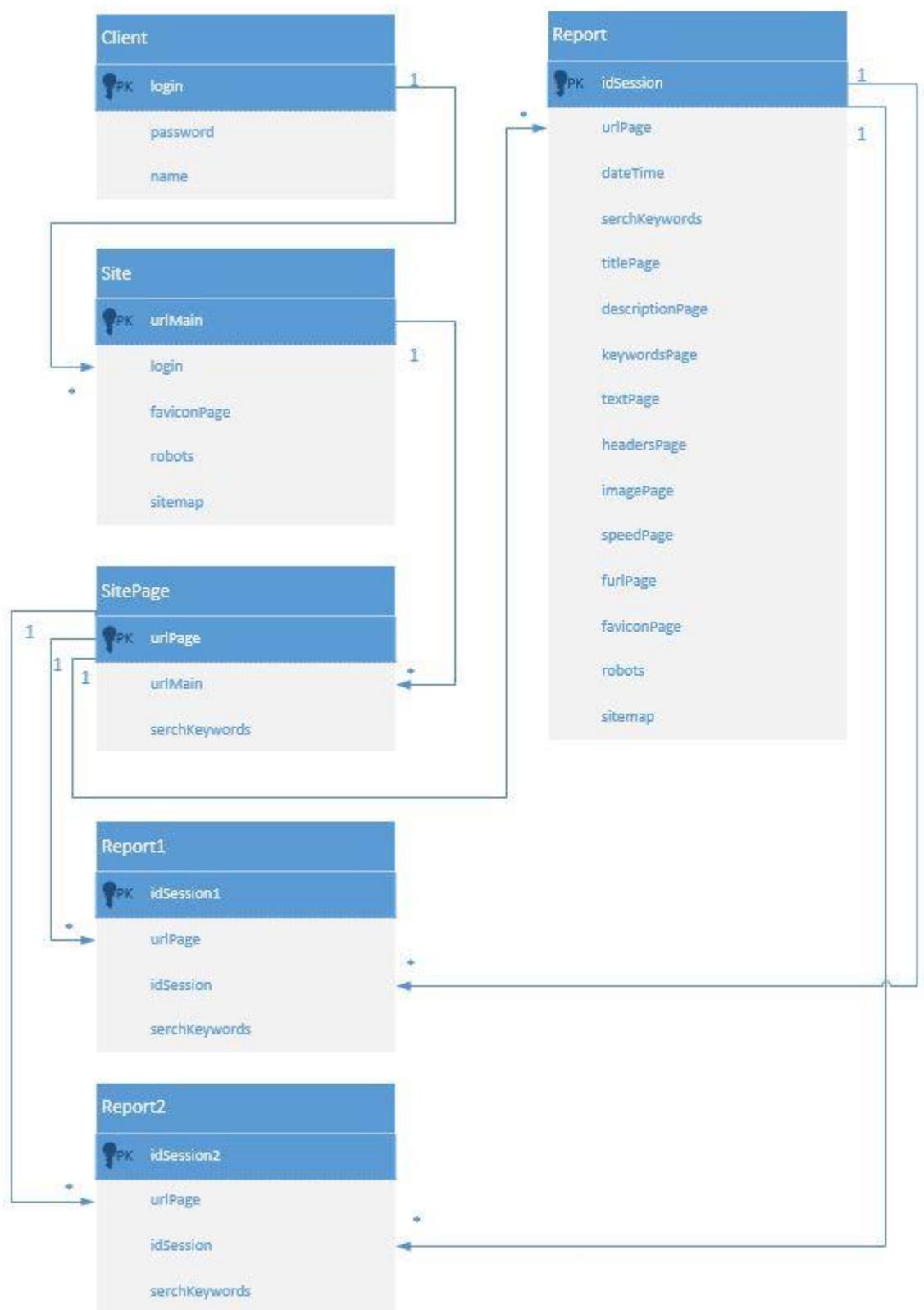


Рисунок 2.7 - Структура базы данных

## 2.4 Разработка пользовательского интерфейса

Закрывающим этапом разработки системы, как и в большинстве случаев, является разработка пользовательского интерфейса. При разработке интерфейса пользователя учитывались следующие требования:

1. Единый стиль оформления для пользовательских интерфейсов;
2. Интерфейс пользователя не должен быть перегружен информацией;
4. Удобная, интуитивно понятная навигация в интерфейсе пользователя;
5. Взаимодействие пользователя с приложением должно осуществляться на русском языке.

На рисунке 2.8 изображена главная форма системы и открыта вкладка «Текстовые и изображения». В этой вкладке пользователь вводит URL страницы сайта и ключевые запросы, после нажимает кнопку «Начать анализ» и получает экспресс-анализ web-ресурса.

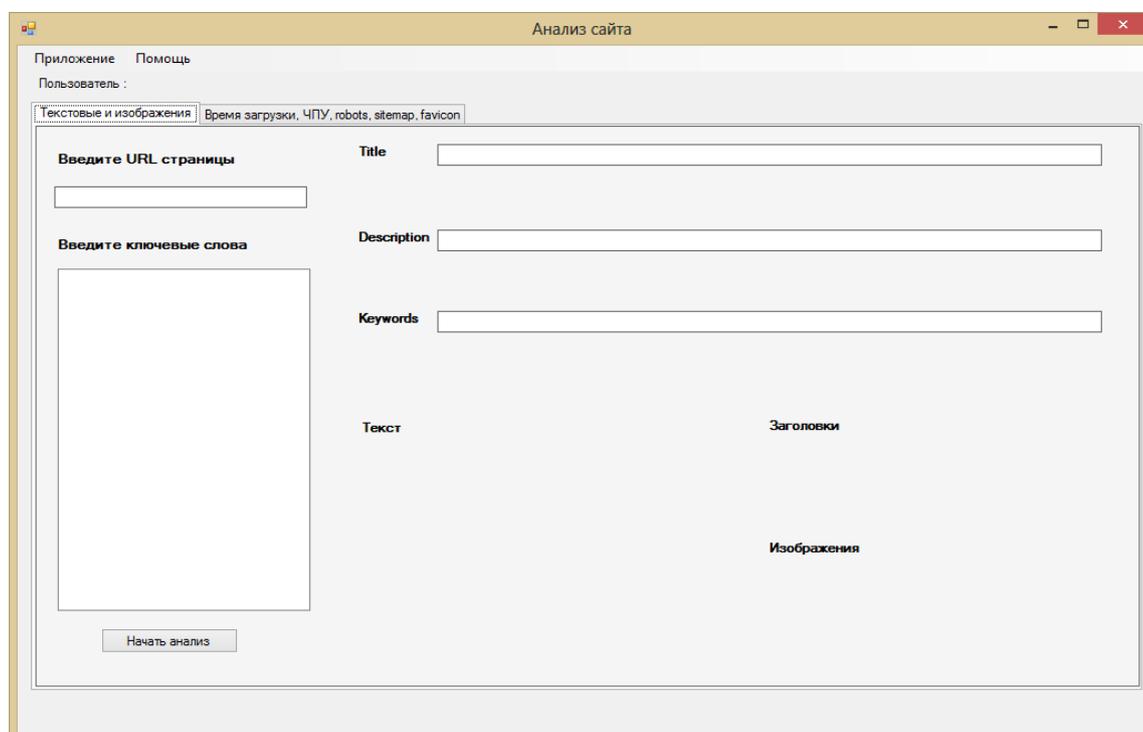


Рисунок 2.8 – Вкладка «Текстовые и изображения»

В результате пользователь получает информацию о внутренних параметрах сайта, такие как Title, Description, Keywords, текст, заголовки и изображения.

При анализе тегов Title, Description и Keywords выводится:

1. Отсутствует. – Если теги не прописаны.
  - 2.1. Много символов – N. – Если символов больше 80 для Title, больше 180 для Description, больше 10 слов для Keywords.
  - 2.2. Мало символов – N. – Если символов меньше 60 для Title, меньше 130 для Description, меньше 2 слов для Keywords.
  - 2.3. N символов. – Количество символов в теге.
- 3.1. Есть ключевые слова. – Если в теге есть ключевые слова.
- 3.1. Нет ключевых слов. – Если в теге нет слов.
4. Переоптимизирован. – Если вхождений запроса больше 1 для Title, больше 2 для Description, больше 3 для Keywords.

При анализе текста в окне приложения выводится:

1. Объем текста N символов. – Выводит количество символов в тексте.
  - 2.1. В тексте нет списков. – Если текст не содержит списков.
  - 2.2. В тексте есть список. – Если текст содержит список.
3. В тексте N абзацев. – Выводит количество абзацев.
3. N. N. K символов. – Выводит количество символов в N-ом абзаце.
  - 4.1. Текст содержит ключевые запросы. – Если в тексте присутствуют ключевые запросы.
  - 4.2. Текст не содержит ключевые запросы. – Если в тексте отсутствуют ключевые запросы.

В результате анализа заголовков система выдает:

- 1.1. Hn отсутствует. – Если на странице отсутствует Hn.
- 1.2. Hn содержит ключевые слова. – Если в HN есть ключевые слова.

1.3. H<sub>n</sub> не содержит ключевые слова. – Если в H<sub>n</sub> нет ключевых слов.

2. H<sub>1</sub> используется больше одного раза. – Если на странице сайта больше одного заголовка H<sub>1</sub>.

n – идентификатор заголовка.

При анализе изображений в окне приложения выводится:

1.1. На странице m изображений. – Количество изображений на странице.

1.2. Отсутствуют. – Если на странице нет изображений.

2.1. В названии изображения нет ключевых слов. – Если название изображения не содержит ключевые слова.

2.2. В названии изображения есть ключевые слова. – Если название изображения содержит ключевые слова.

3.1. Alt не прописан. – Если тег Alt не прописан.

3.2. Alt прописан. – Если тег Alt прописан.

4.1. Title не прописан. – Если тег Title не прописан.

4.2. Title прописан. – Если тег Title прописан.

Во вкладке «Время загрузки, ЧПУ, robots, sitemap, favicon» происходит анализ остальных параметров (Рисунок 2.9).

Здесь выводится время загрузки сайта, прописан или нет ЧПУ, есть или нет favicon. Так же здесь проверяется наличие файлов robots.txt и sitemap.xml и их корректность, и выводится прописано ли в файле robots.txt главное зеркало сайта.

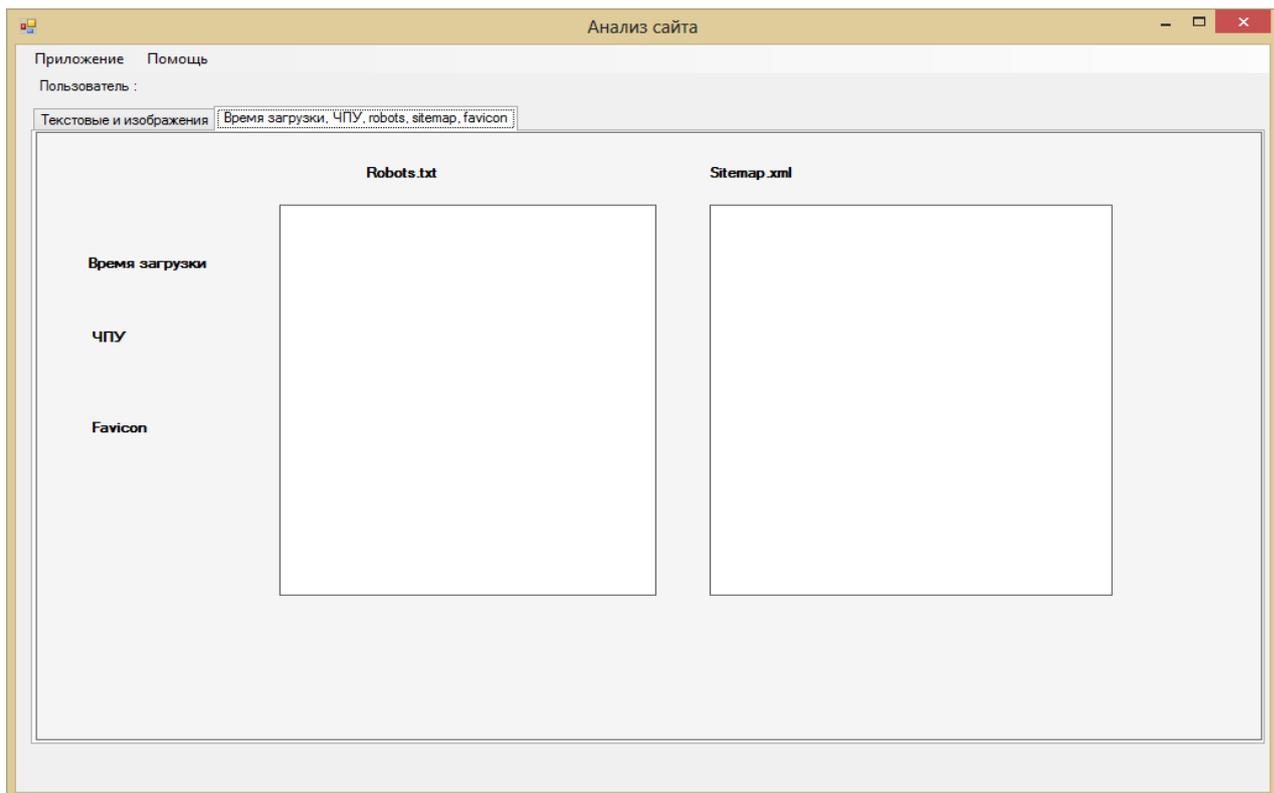


Рисунок 2.9 – Вкладка «Время загрузки, ЧПУ, robots, sitemap, favicon»

В Приложении Б находятся остальные формы системы:

1. Форма для входа в систему, в ней пользователь вводит свой логин и пароль (Рисунок Б.1). Если пользователь еще не зарегистрировался, то он может перейти в форму регистрации.

2. Форма регистрации (Рисунок Б.2) позволяет создать учетную запись пользователя.

3. Зарегистрированному и авторизованному пользователю предоставляется возможность получить подробный отчет об анализе сайта (Рисунок Б.3).

4. А также сравнение двух отчетов с одинаковыми URL и ключевыми словами (Рисунок Б.4).

5. На рисунке Б.5 изображена форма, в которой пользователь может выбрать нужный ему URL сайта, дату и время его анализа для просмотра или сравнения отчетов.

6. На рисунке Б.6 изображена форма “О программе”, в которой находится информация о разрабатываемой системе.

## 2.5 Анализ соответствия системы сформулированным требованиям

По разработанной информационной системе был проведен анализ соответствия сформулированным требованиям. В таблице 2.1 представлены результаты проведения этого анализа, где видно, что система соответствует всем поставленным требованиям.

Таблица 2.1 - Анализ соответствия системы сформулированным требованиям

Формулировка требования	Описание реализации требования	Отметка о выполнении требования
По требованию пользователя проводить регистрацию или авторизацию пользователя в системе	Разработана подсистема идентификации пользователей. Организовано хранение информации об учетной записи в базе данных.	Выполнено
Выводить Title, Description, Keywords анализируемой страницы сайта	Выделение Title, Description, Keywords из HTML-кода страницы сайта	Выполнено
Определять наличие или отсутствие Title, Description, Keywords, заголовков, изображений, тегов Title и Alt изображений, текста, маркированных или нумерованных списков в тексте на странице сайта, файла robots.txt и карты сайта	Проверка на наличие в HTML-коде страницы: Title, Description, Keywords, заголовков, изображений, тегов Title и Alt изображений, текста, маркированных или нумерованных списков в тексте на странице сайта, Проверка на наличие на сервере сайта файла robots.txt и карты сайта.	Выполнено
Определять наличие или отсутствие ключевых фраз в Title, Description, Keywords, заголовках, тегах Title и Alt изображений, тексте на странице сайта	Проверка на наличие ключевой фразы в Title, Description, Keywords, заголовках, тегах Title и Alt изображений, тексте	Выполнено

Определять количество символов в Title, Description, Keywords, тексте и абзацах, количество абзацев в тексте на странице сайта	Подсчет символов в Title, Description, Keywords, тексте и абзацах. Подсчет количества абзацев в тексте	Выполнено
Определять количество вхождений ключевых слов в Title, Description, Keywords, тексте и абзацах, количество абзацев в тексте на странице сайта	Подсчет количества вхождений ключевых слов в Title, Description, Keywords, тексте и абзацах, количество абзацев в тексте на странице сайта	Выполнено
Определять скорость загрузки сайта	Создан контроллер для генерации страницы системой	Выполнено
Определять наличие человеко-понятных URL и favicon	Проверка на наличие человеко-понятных URL и favicon	Выполнено
Определять корректность составления файла robots.txt и карты сайта	Проверка наличия главного зеркала в файле robots.txt. Проверка карты сайта на страницы, которые закрыты в файле robots.txt.	Выполнено
По запросу пользователя предоставлять подробный отчет об анализе страницы сайта	Сравнение данных страницы с эталонными значениями факторов ранжирования и формирование рекомендаций	Выполнено
По запросу пользователя предоставлять сравнение подробных отчетов об анализе страницы сайта	Вывод результатов двух подробных отчетов одной страницы с одинаковыми ключевыми фразами	Выполнено
Экспортировать подробный отчет об анализе	Конвертация подробного отчета в PDF-документ и предоставление пользователю	Выполнено
Предоставлять справку о работе с программой по запросу пользователя	Разработана форма справки с основной информацией по работе системы и определениями.	Выполнено
Иметь удобный пользовательский интерфейс	Разработан интуитивно-понятный пользовательский интерфейс	Выполнено

### 3 Результаты внедрения системы

Целевой аудиторией разрабатываемого продукта являются веб-студии и организации, у которых есть отдел веб-аналитики или интернет-маркетинга.

Разработанная система используется в отделе интернет-маркетинга ООО «Нева-Томь Консалт». Для демонстрации работы системы были проведены анализы факторов ранжирования сайта [potolkikaluga.ru](http://potolkikaluga.ru) до и после проведения работ по поисковой оптимизации, которые основывались на рекомендациях, полученных в результате анализа. Внедрение системы выполнялось на основании акта внедрения (Приложение В).

Результат анализа главной страницы сайта <http://potolkikaluga.ru/> изображен на рисунках 3.1 и 3.2, где видно какие проблемы следует устранить на этой странице в процессе оптимизации сайта.

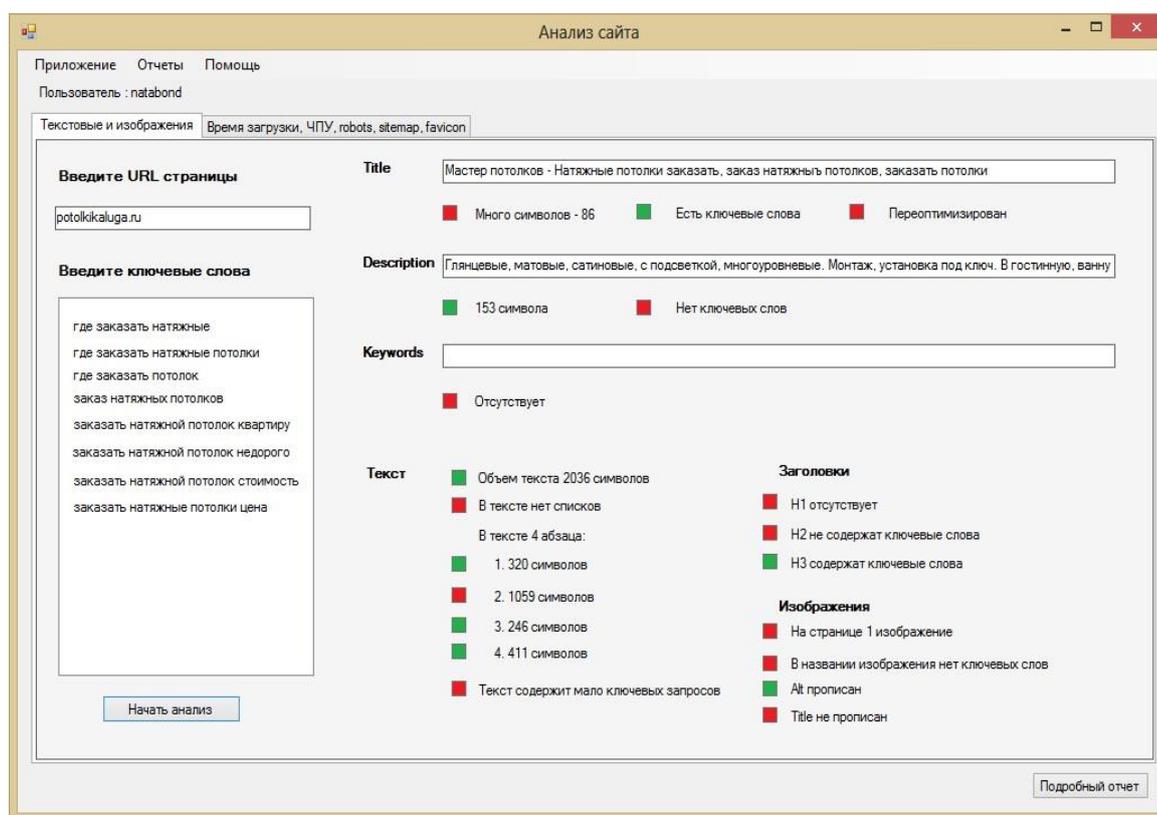


Рисунок 3.1- Результат анализа страницы <http://potolkikaluga.ru/>, вкладка «Текстовые и изображения»

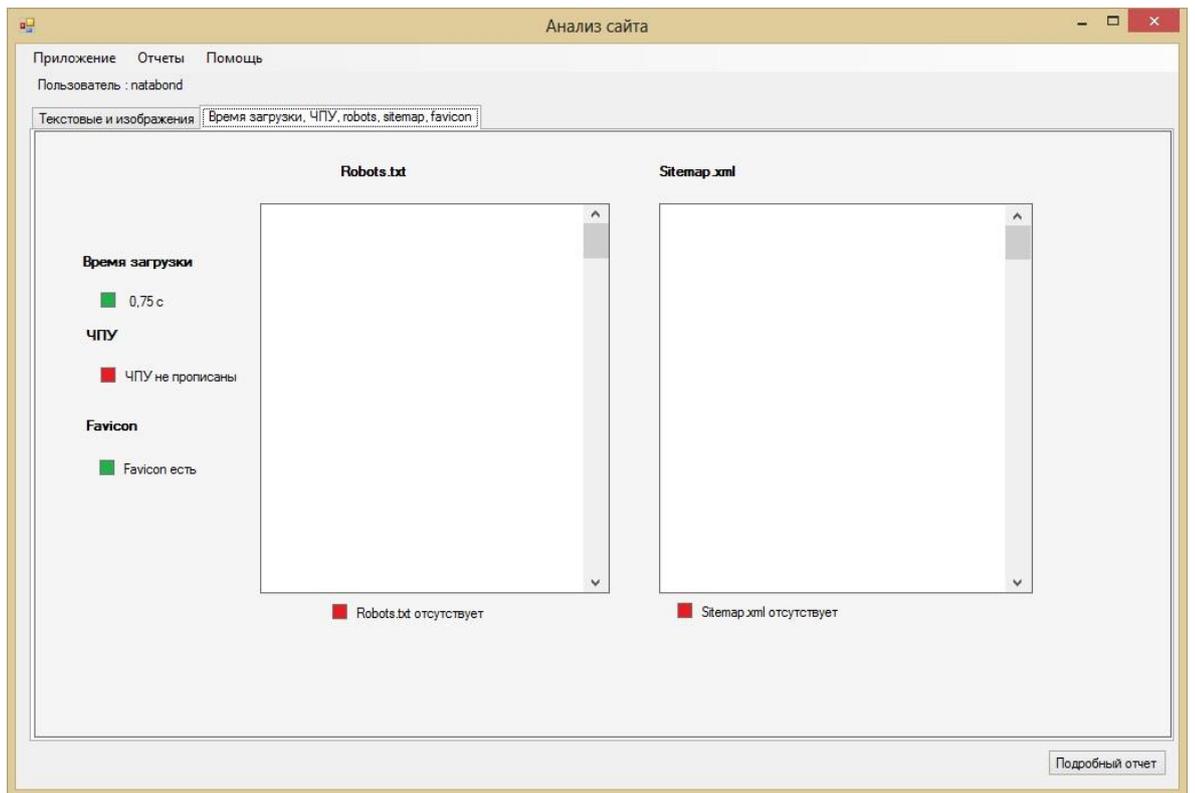


Рисунок 3.2- Результат анализа страницы <http://potolkikaluga.ru/>, вкладка «Время загрузки, человеко-понятные URL, robots, sitemap, favicon»

На основе экспресс-анализа система сформировала подробный отчет с рекомендациями для проведения работ по оптимизации страницы сайта, чтобы она соответствовала основным требованиям поисковых систем.

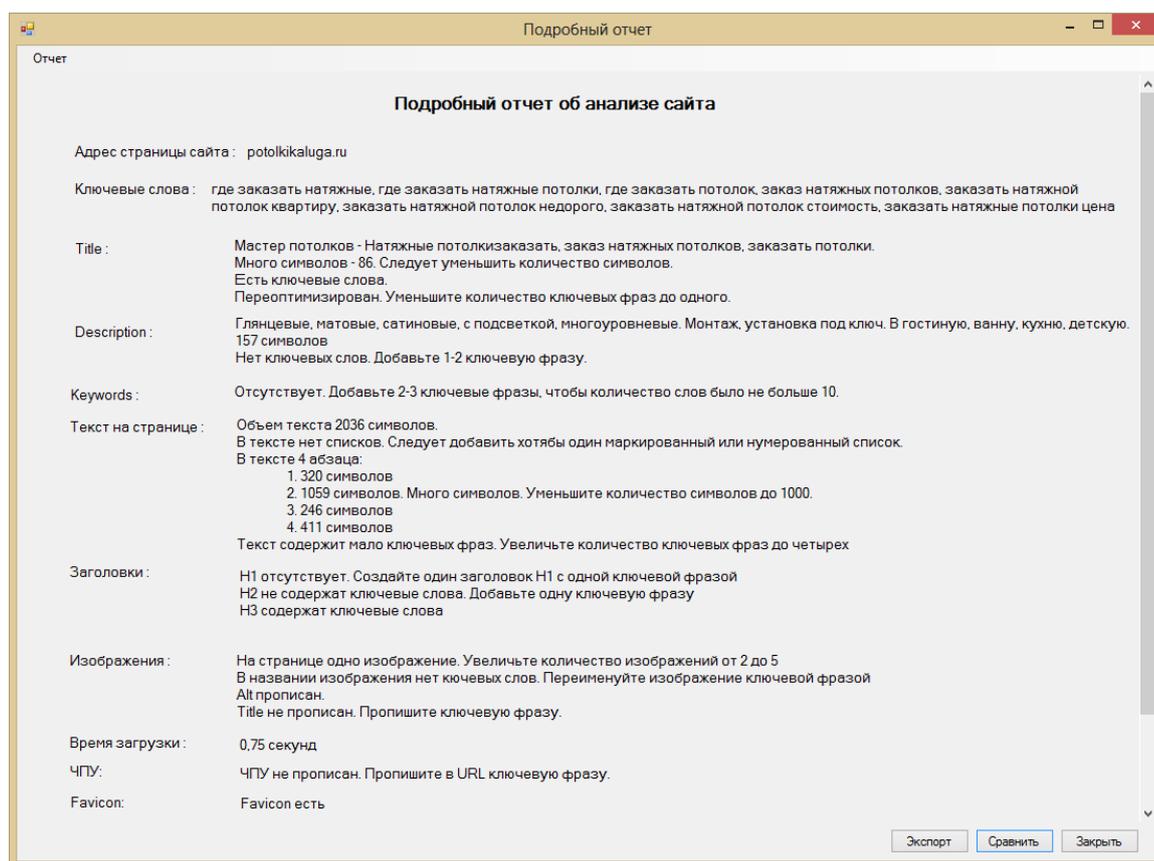


Рисунок 3.3 – Подробный отчет

После проведения всех работ по оптимизации сайта, на основе подробного отчета с рекомендациями, был произведен повторный анализ страницы сайта. На рисунках 3.4 и 3.5 изображены результаты повторного анализа страницы сайта, на которых можно увидеть, что после оптимизации сайт стал соответствовать требованиям поисковых систем.

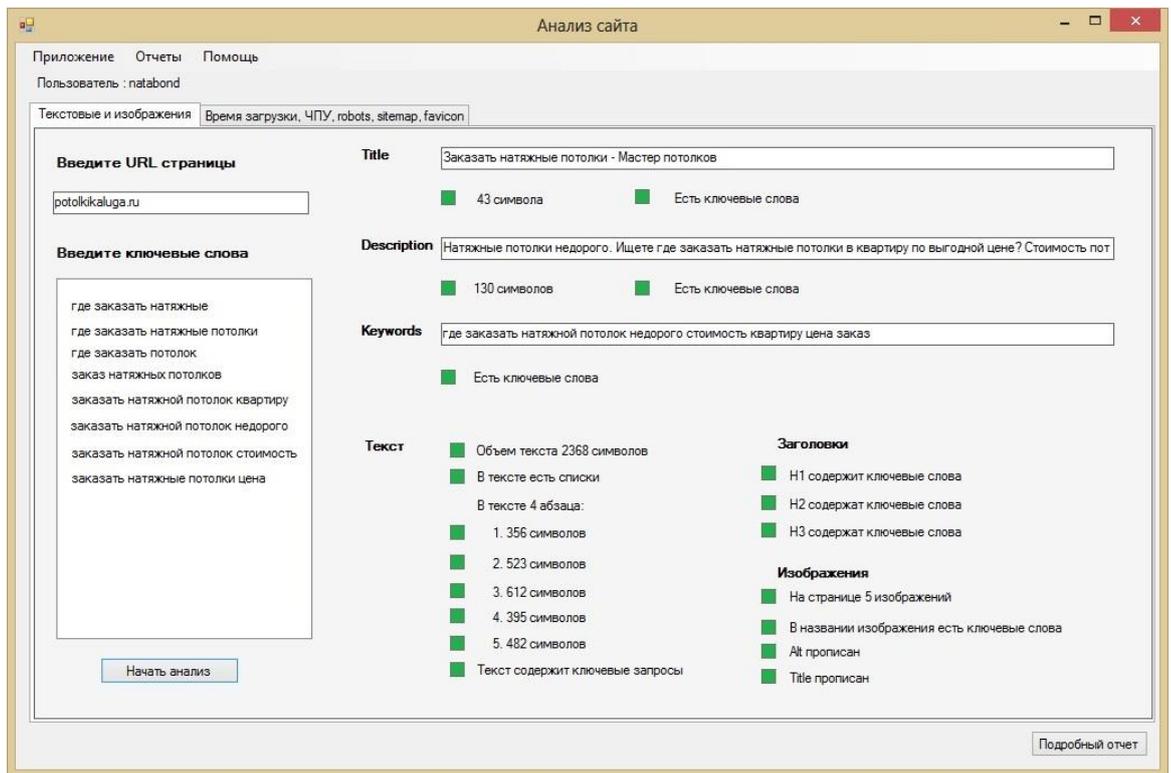


Рисунок 3.4 - Результат анализа страницы <http://potolkikaluga.ru/> после оптимизации, вкладка «Текстовые и изображения»

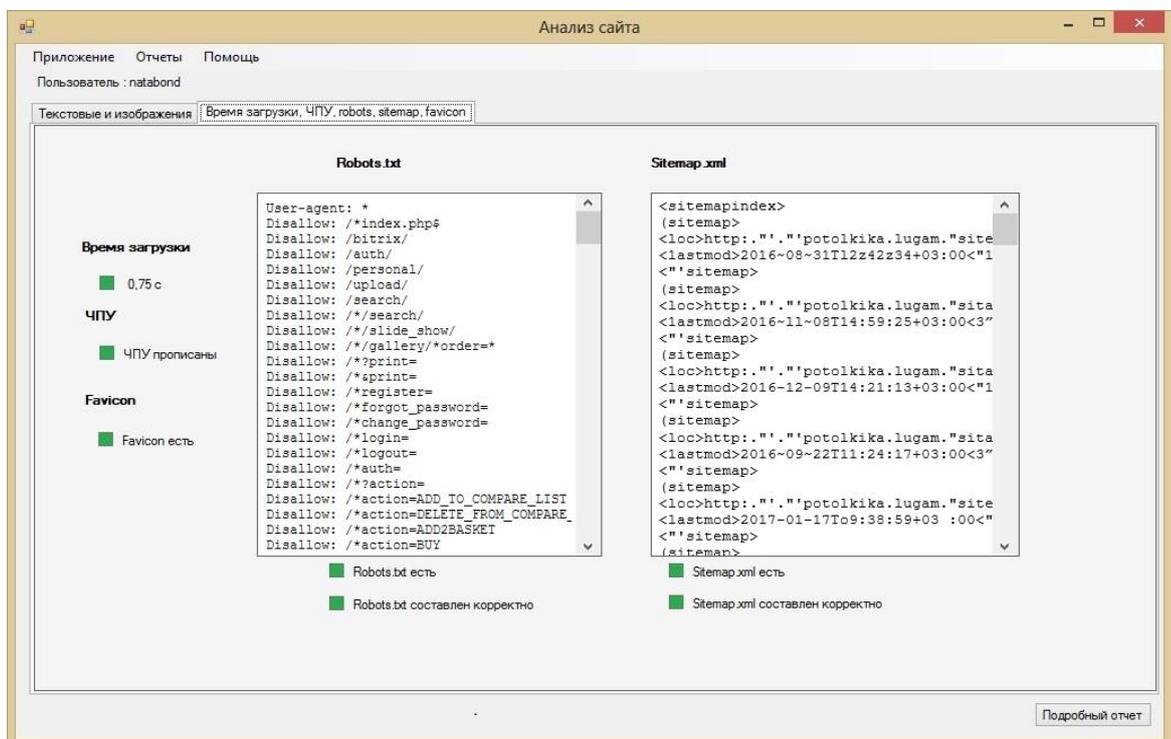


Рисунок 3.5 - Результат анализа страницы <http://potolkikaluga.ru/> после оптимизации, вкладка «Время загрузки, человеко-понятные URL, robots, sitemap, favicon»

Далее было проведено сравнение двух отчетов: до проведения работ по оптимизации и после. На рисунке 3.6 видно, что после редактирования страницы сайта, были ликвидированы ее недостатки.

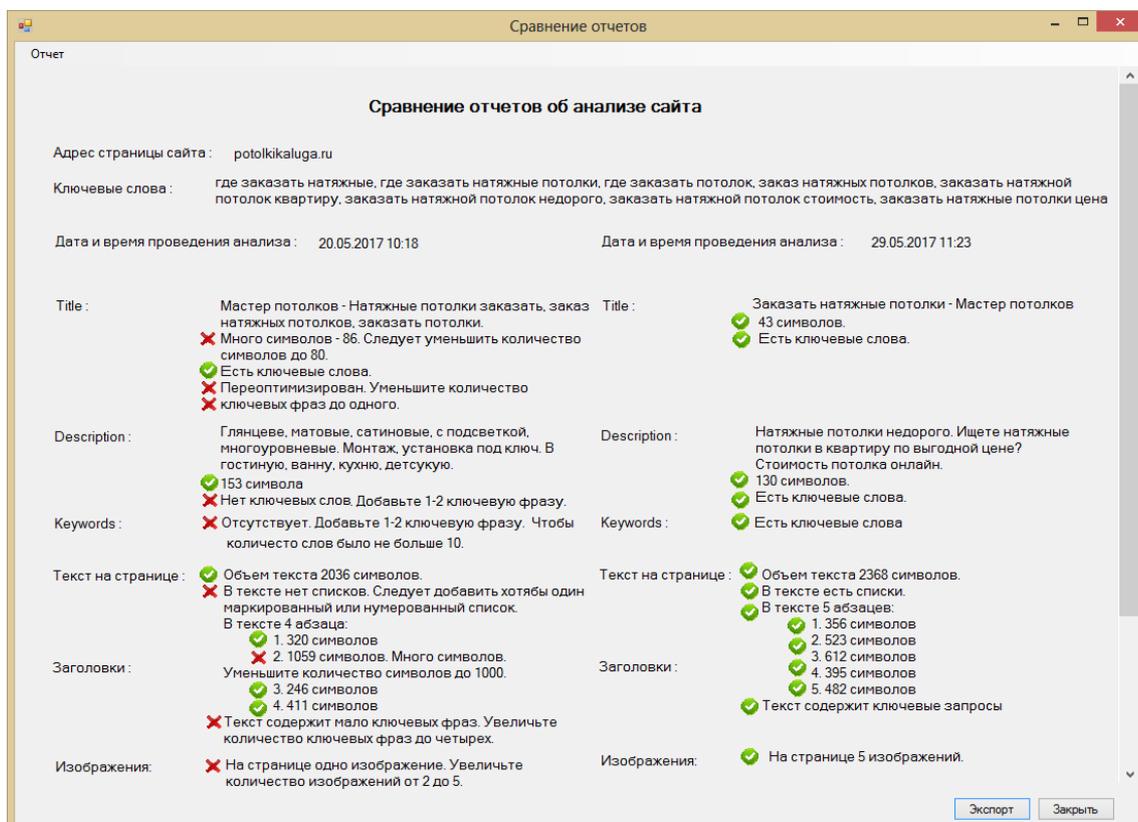


Рисунок 3.6 – Сравнение подробных отчетов

Так же во время внедрения системы были определены позиции страницы сайта в поисковых системах до и после проведения работ по оптимизации, согласно полученным рекомендациям.

В результате внедрения системы было выявлено, что при учете всех рекомендаций, сформированных на основе данных страницы сайта и эталонных значений, оптимизированная страница сайта поднялась в Топ-10 поисковых систем Google и Яндекс по определенным запросам (Таблица 3.1).

Таблица 3.1 – Позиции страницы сайта в поисковых системах

Ключевой запрос	До оптимизации		После оптимизации	
	Google	Яндекс	Google	Яндекс
Где заказать натяжные	23	35	6	5
Где заказать натяжные потолки	28	19	3	2.
Где заказать потолок	32	41	7	4
Заказ натяжных потолков	18	21	1	2
Заказать потолок квартиру	45	34	4	3
Заказать потолок недорого	19	25	2	1
Заказать потолок стоимость	27	16	1	3
Заказать потолок цена	28	22	1	2

## 4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Целью данного раздела является анализ перспективности проведения научно-исследовательских работ путем оценки коммерческой ценности и технико-экономического обоснования разработки, а также количественного и качественного доказательства целесообразности её внедрения.

### 4.1 Организация и планирование работ

При организации процесса реализации конкретного проекта необходимо оптимально планировать занятость каждого из его участников и сроки проведения отдельных работ.

На данном этапе составляется полный перечень проводимых работ, и определяются их исполнители и оптимальная продолжительность. Результатом планирования работ является сетевой, либо линейный график реализации проекта. Наиболее удобным, простым и наглядным способом для этих целей является использование линейного графика. Для его построения составим перечень работ и соответствие работ своим исполнителям, продолжительность выполнения этих работ и сведем их в таблицу 4.1.

Таблица 4.1 – Перечень работ и продолжительность их выполнения

№	Этапы работы	Исполнители	Загрузка исполнителей
1	Постановка целей и задач, получение исходных данных	НР	НР – 100%
2	Составление и утверждение ТЗ	НР, С	НР – 100% С – 30%
3	Подбор и изучение материалов по тематике	НР, С	НР – 20% С – 100%
4	Разработка календарного плана	НР, С	НР – 100% С – 30%

5	Обсуждение литературы	НР, С	НР – 30% С – 100%
6	Изучение аналогов	НР, С	НР – 20% С – 100%
7	Разработка информационной системы	С	НР – 5% С – 100%
8	Внедрение системы	С	С – 100%
9	Оформление расчетно-пояснительной записки	С	С – 100%
10	Оформление графического материала	С	С – 100%
11	Подведение итогов	НР, С	НР – 60% С – 100%

Примечания к таблице 5.1:

НР – Научный руководитель

С – Студент

#### 4.1.1 Продолжительность этапов работ

Расчет продолжительности этапов работ может выполняться двумя методами: технико-экономическим, опытно-статистическим.

В данном случае используется опытно-статистический метод, который может быть реализован двумя способами: аналоговый и вероятностный.

Для расчета ожидаемого значения продолжительности работ  $t_{ож}$  применяется две оценки:  $t_{min}$  и  $t_{max}$  (метод двух оценок).

$$t_{ож} = \frac{3 \cdot t_{min} + 2 \cdot t_{max}}{5}, \quad (4.1)$$

где  $t_{min}$  – минимальная трудоемкость работ, чел/дн.;

$t_{max}$  – максимальная трудоемкость работ, чел/дн.

Для выполнения перечисленных в таблице 4.1 работ требуются специалисты: студент, научный руководитель.

Для построения линейного графика необходимо рассчитать длительность этапов в рабочих днях, а затем перевести в календарные дни. Расчет продолжительности выполнения каждого этапа в рабочих днях выполняется по формуле:

$$T_{РД} = \frac{t_{ож}}{K_{ВН}} \cdot K_{Д} \quad (4.2)$$

где  $t_{ож}$  – трудоемкость работы, чел/дн.;

$K_{ВН}$  – коэффициент выполнения работ ( $K_{ВН} = 1$ );

$K_{Д}$  – коэффициент, учитывающий дополнительное время на компенсации и согласование работ ( $K_{Д} = 1,2$ ).

Расчет продолжительности этапа в календарных днях ведется по формуле:

$$T_{КД} = T_{РД} \cdot T_{К}, \quad (4.3)$$

где  $T_{РД}$  – продолжительность выполнения этапа в рабочих днях;

$T_{КД}$  – продолжительность выполнения этапа в календарных днях;

$T_{К}$  – коэффициент календарности.

Коэффициент календарности рассчитывается по формуле:

$$T_{К} = \frac{T_{КАЛ}}{T_{КАЛ} - T_{ВД} - T_{ПД}} \quad (4.4)$$

где  $T_{КАЛ}$  – календарные дни ( $T_{КАЛ} = 365$ );

$T_{ВД}$  – выходные дни ( $T_{ВД} = 104$ );

$T_{ПД}$  – праздничные дни ( $T_{ПД} = 14$ ).

$$T_{К} = \frac{365}{365 - 104 - 14} = 1,48$$

В таблице 4.2 приведена длительность этапов работ и число исполнителей, занятых на каждом этапе.

Таблица 4.2 - Трудозатраты на выполнение проекта

Этап	Исполнители	Продолжительность работ, дни			Трудоемкость работ по исполнителям чел.- дн.			
		$t_{min}$	$t_{max}$	$t_{ож}$	$T_{РД}$		$T_{КД}$	
					НР	С	НР	С
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1. Постановка задачи	НР	2,00	4,00	2,80	3,36	–	4,97	–
2. Разработка и утверждение технического задания (ТЗ)	НР, И	4,00	7,00	5,20	6,24	1,87	9,24	2,77
3. Подбор и изучение материалов по тематике	НР, И	24,00	36,00	28,80	6,91	34,56	10,23	51,15
4. Разработка календарного плана	НР, И	1,00	3,00	1,80	2,16	0,65	4,97	1,49
5. Обсуждение литературы	НР, И	4,00	6,00	4,80	1,73	5,76	2,56	8,52
6. Изучение аналогов	НР, И	2,00	3,00	2,40	0,58	2,88	0,85	4,26
7. Разработка информационной системы	НР, И	28,00	36,00	31,20	1,87	37,44	2,77	55,41
8. Внедрение системы	И	2,00	4,00	2,80	–	3,36	–	4,97
9. Оформление расчетно-пояснительной записки	И	20,00	28,00	23,20	–	27,84	–	41,20
10. Оформление графического материала	И	3,00	5,00	3,80	–	4,56	–	6,75
11. Подведение итогов	НР, И	2,00	5,00	3,20	2,30	3,84	3,41	5,68
<b>Итого:</b>				<b>110,00</b>	<b>26,35</b>	<b>123,12</b>	<b>39,00</b>	<b>182,22</b>

Таблица 4.3 - Линейный график работ

Этап	НР	И	Ноябрь	Декабрь				Январь			Февраль			Март			Апрель			Май		
			10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	
1	4,97	-	■																			
2	9,24	2,77	■	■	■																	
3	10,23	51,15		■	■	■	■	■	■	■												
4	4,97	1,49							■	■												
5	2,56	8,52							■	■	■											
6	0,85	4,26								■	■											
7	2,77	55,41								■	■	■	■	■	■							
8	-	4,97														■						
9	-	41,20														■	■	■	■	■	■	
10	-	6,75																		■	■	
11	3,41	5,68																		■	■	

Примечания к таблице 4.3:

■ - Научный руководитель

■ - Студент

#### 4.1.2 Расчет накопления готовности проекта

Цель данного пункта – оценка текущих состояний (результатов) работы над проектом. Величина накопления готовности работы показывает, на сколько процентов по окончании текущего (*i*-го) этапа выполнен общий объем работ по проекту в целом.

Введем обозначения:

- $TP_{\text{общ}}$  – общая трудоемкость проекта;
- $TP_i$  ( $TP_k$ ) – трудоемкость *i*-го (*k*-го) этапа проекта,  $i = \overline{1, I}$ ;
- $TP_i^H$  – накопленная трудоемкость *i*-го этапа проекта по его завершению;
- $TP_{ij}$  ( $TP_{kj}$ ) – трудоемкость работ, выполняемых *j*-м участником на *i*-м этапе, здесь индекс  $j = \overline{1, m}$  – индекс исполнителя, в нашем примере  $m = 2$ .

Степень готовности определяется формулой (5.5)

$$CG_i = \frac{TP_i^H}{TP_{\text{общ.}}} = \frac{\sum_{k=1}^i TP_k}{TP_{\text{общ.}}} = \frac{\sum_{k=1}^i \sum_{j=1}^m TP_{km}}{\sum_{k=1}^I \sum_{j=1}^m TP_{km}}. \quad (4.5)$$

Применительно к таблице (4.2) величины  $TP_{ij}$  ( $TP_{kj}$ ) находятся в столбцах (6,  $j = 1$ ) и (7,  $j = 2$ ).  $TP_{\text{общ.}}$  равна сумме чисел из итоговых клеток этих столбцов. Пример расчета  $TP_i$  (%) и  $CG_i$  (%) на основе этих данных содержится в таблице 4.4.

Таблица 4.4 – Нарастание технической готовности работы и удельный вес каждого этапа

Этап	$t_{\text{ож}}$	$t_{\text{нi}}$	$TP_i, \%$	$CG_i, \%$
1. Постановка задачи	2,80	2,80	2,55	2,55
2. Разработка и утверждение технического задания (ТЗ)	5,20	8,00	4,73	7,27
3. Подбор и изучение материалов по тематике	28,80	36,80	26,18	33,45

4. Разработка календарного плана	1,80	38,60	1,64	35,09
5. Обсуждение литературы	4,80	43,40	4,36	39,45
6. Изучение аналогов	2,40	45,80	2,18	41,64
7. Разработка информационной системы	31,20	77,00	28,36	70,00
8. Внедрение системы	2,80	79,80	2,55	72,55
9. Оформление расчетно-пояснительной записки	23,20	103,00	21,09	93,64
10. Оформление графического материала	3,80	106,80	3,45	97,09
11. Подведение итогов	3,20	110,00	2,91	100,00

#### **4.2 Расчет сметы затрат на выполнение проекта**

В состав затрат на создание проекта включается величина всех расходов, необходимых для реализации комплекса работ, составляющих содержание данной разработки. Расчет сметной стоимости ее выполнения производится по следующим статьям затрат:

- материалы и покупные изделия;
- заработная плата;
- социальный налог;
- расходы на электроэнергию (без освещения);
- амортизационные отчисления;
- командировочные расходы;
- оплата услуг связи;
- арендная плата за пользование имуществом;
- прочие услуги (сторонних организаций);
- прочие (накладные расходы) расходы.

#### 4.2.1 Расчет затрат на материалы

К данной статье расходов относится стоимость материалов, покупных изделий, полуфабрикатов и других материальных ценностей, расходуемых непосредственно в процессе выполнения работ.

Данная работа выполнялась преимущественно на компьютерном оборудовании и не потребовала существенных затрат на материалы. Мелкие расходы (канцелярия и др.) могут быть отнесены к статье прочих расходов (см. п. 4.2.6). Следовательно, можно принять  $C_M = 0$ .

#### 4.2.2 Расчет заработной платы

Данная статья расходов включает заработную плату научного руководителя и студента, а также премии, входящие в фонд заработной платы. Расчет основной заработной платы выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя.

Среднедневная заработная плата рассчитывается по формуле:

$$ЗП_{дн-г} = MO/20,58 \quad (4.6)$$

Расчеты затрат на заработную плату приведены в таблице 4.5. При расчете учитывалось, что в году 247 рабочий день и, следовательно, в месяце 20,58 рабочих дня. Затраты времени на выполнение работы по каждому исполнителю брались из таблицы 4.2. Также был принят во внимание коэффициент, учитывающий коэффициент по премиям  $K_{ПР} = 1,1$ , коэффициент дополнительной заработной платы (учитывающий оплату отпуска, простой и пр.)  $K_{ДОП} = 1,188$  и районный коэффициент  $K_{РК} = 1,3$ .

Таблица 4.5 — Затраты на основную заработную плату

Исполнитель	Оклад, руб./мес	Среднедневная ставка, руб./день	Затраты времени, дни	Коэффи циент	Фонд з/платы, руб.
НР	32000	1554,9	26,35	1,592	65 226,81
С	15000	728,36	123,12	1,592	142 763,69
<b>Итого:</b>					<b>207 990,5</b>

Таким образом, затраты на заработную плату составили СЗП = 207 990,5 руб.

#### 4.2.3 Расчет отчислений от заработной платы

Затраты по этой статье составляют отчисления по единому социальному налогу (ЕСН).

Отчисления по заработной плате определяются по следующей формуле:

$$C_{\text{соц.}} = C_{\text{зп}} * 0,3, \quad (4.7)$$

где  $K_{\text{соц}}$  – коэффициент, учитывающий размер отчислений из заработной платы. Данный коэффициент составляет 30% от затрат на заработную плату и включает в себя:

- отчисления в пенсионный фонд;
- на социальное страхование;
- на медицинское страхование.

Итак, отчисления из заработной платы составили:

$$C_{\text{соц}} = 0,3 \cdot 207\,990,5 = 62\,397,15 \text{ руб.}$$

#### 4.2.4 Расчет затрат на электроэнергию

Данный вид расходов включает в себя затраты на электроэнергию при работе оборудования, а также затраты на электроэнергию, потраченную на

освещение. Затраты на электроэнергию при работе оборудования для технологических целей рассчитываются по формуле:

$$C_{\text{эл.об.}} = P_{\text{об.}} \cdot t_{\text{об.}} \cdot \text{ЦЭ} \quad (4.8)$$

где  $P_{\text{об}}$  – мощность, потребляемая оборудованием, кВт;

ЦЭ – тарифная цена за 1 кВт·час.

$t_{\text{об}}$  – время работы оборудования, час.

Для ТПУ ЦЭ = 5,782 руб./кВт·час (с НДС).

Время работы оборудования вычисляется на основе данных для  $T_{\text{рд}}$  таблицы 4.2 для студента из расчета, что продолжительность рабочего дня равна 8 часов.

Мощность, потребляемая оборудованием, определяется по формуле:

$$t_{\text{об}} = T_{\text{рд}} * K_t, \quad (4.9)$$

где  $K_t \leq 1$  – коэффициент использования оборудования по времени, равный отношению времени его работы в процессе выполнения проекта к  $T_{\text{рд}}$ , определяется исполнителем самостоятельно. В ряде случаев возможно определение  $t_{\text{об}}$  путем прямого учета, особенно при ограниченном использовании соответствующего оборудования.

Мощность, потребляемая оборудованием, определяется по формуле:

$$P_{\text{об}} = P_{\text{ном.}} * K_c \quad (4.10)$$

где  $P_{\text{ном.}}$  – номинальная мощность оборудования, кВт;

$K_c \leq 1$  – коэффициент загрузки, зависящий от средней степени использования номинальной мощности. Для технологического оборудования малой мощности  $K_c = 1$ .

Пример расчета затраты на электроэнергию для технологических целей приведен в таблице 4.7.

Таблица 4.3 – Затраты на электроэнергию технологическую

Наименование оборудования	Время работы оборудования $t_{об}$ , час	Потребляемая мощность $P_{об}$ , кВт	Затраты Эоб, руб.
Персональный компьютер	588	0,3	1019,95
Струйный принтер	25	0,1	14,45
<b>Итого:</b>			<b>1034,4</b>

#### 4.2.6 Расчет расходов, учитываемых непосредственно на основе платежных (расчетных) документов (кроме суточных)

Сюда относятся:

- командировочные расходы, в т.ч. расходы по оплате суточных, транспортные расходы, компенсация стоимости жилья;
- арендная плата за пользование имуществом;
- оплата услуг связи;
- услуги сторонних организаций.

Норма оплаты суточных – 100 руб./день.

Во время проведения работ по НИР таких расходов не было, таким образом  $C_{ни} = 0$  руб.

#### 4.2.5 Расчет амортизационных расходов

В статье «Амортизационные отчисления» от используемого оборудования рассчитывается амортизация за время выполнения работы для оборудования, которое имеется в наличии.

Амортизационные отчисления рассчитываются на время использования ЭВМ по формуле:

$$C_{AM} = \frac{N_A * C_{OB} * t_{рф} * n}{F_D}, \quad (4.11)$$

где  $N_A$  – годовая норма амортизации,  $N_A = 25\%$ ;

$C_{OB}$  – цена оборудования,  $C_{OB} = 57\,000$  руб.;

$F_D$  – действительный годовой фонд рабочего времени,  $F_D = 2416$  часов;

$t_{BT}$  – время работы вычислительной техники во время проведения НИР,  
 $t_{BT} = 588$  часа;

$n$  – число задействованных ПЭВМ,  $n = 1$ .

Итак, затраты на амортизационные отчисления составили:

$$C_{AM} (ПК) = (0,25 * 57000 * 588 * 1) / 1976 = 4240,38 \text{ руб.}$$

Для принтера  $C_{OB} = 12000$  руб., его  $F_D = 500$  час.,  $N_A = 50\%$ , тогда его амортизация составит:

$$C_{AM} (ПК) = (0,25 * 12000 * 25 * 1) / 500 = 150 \text{ руб.}$$

Итого начислено амортизации 3 664,84 руб.

#### 4.2.6 Расчет прочих расходов

В статье «Прочие расходы» отражены расходы на выполнение проекта, которые не учтены в предыдущих статьях, их следует принять равными 10% от суммы всех предыдущих расходов, т.е.

$$C_{проч.} = (C_{мат} + C_{зн} + C_{соц} + C_{эл.об.} + C_{ам} + C_{нп}) \cdot 0,1 \quad (4.12)$$

Подставим значения и вычислим:

$$C_{проч.} = (0 + 207990,5 + 62397,15 + 1034,4 + 4240,38 + 0) \cdot 0,1 = 28\,465,73 \text{ руб.}$$

#### 4.2.7 Расчет общей себестоимости разработки

Проведя расчет по всем статьям сметы затрат на разработку, можно определить общую себестоимость проекта «Разработка системы анализа факторов ранжирования web-ресурсов».

Таблица 4.4 – Смета затрат на разработку проекта

Статья затрат	Условное обозначение	Сумма, руб.
Материалы и покупные изделия	$C_{\text{мат}}$	0
Основная заработная плата	$C_{\text{зп}}$	207990,5
Отчисления в социальные фонды	$C_{\text{соц}}$	62397,15
Расходы на электроэнергию	$C_{\text{эл.}}$	1034,4
Амортизационные отчисления	$C_{\text{ам}}$	4240,38
Непосредственно учитываемые расходы	$C_{\text{нр}}$	0
Прочие расходы	$C_{\text{проч}}$	28 465,73
<b>Итого:</b>		<b>304 128,16</b>

Таким образом, затраты на разработку составили  $C = 304\ 128,16$  руб.

#### 4.2.8 Расчет прибыли

Примем прибыль в размере 5 - 20 % от полной себестоимости проекта. В нашем случае она составляет 60519,12 руб. (20 %) от расходов на разработку проекта.

#### 4.2.9 Расчет НДС

НДС составляет 18% от суммы затрат на разработку и прибыли. В нашем случае это  $(302\ 595,58 + 60519,12) * 0,18 = 65360,65$  руб.

#### **4.2.10 Цена разработки НИР**

Цена равна сумме полной себестоимости, прибыли и НДС, в нашем случае:

$$C_{\text{НИР(КР)}} = 302\,595,58 + 60\,519,12 + 65\,360,65 = 428\,475,346 \text{ руб.}$$

#### **4.3 Оценка эффективности проекта**

В рамках работы не представляется возможным оценить экономическую эффективность проекта, потому что для этого необходимо проводить специальное исследование. Невозможность количественной оценки обусловлена тем, что информация о прототипах и аналогах системы анализа web-ресурсов практически недоступна, в связи с чем в рамках данной работы невозможно ее систематизировать.

Кроме того, результатом данной научно-исследовательской работы является программный продукт, разработанный для закрытого внутреннего пользования, поэтому информация по экономическому эффекту является закрытой.

##### **4.3.1 Оценка научно-технического уровня НИР**

Научно-технический уровень характеризует, в какой мере выполнены работы и обеспечивается научно-технический прогресс в данной области. Для оценки научной ценности, технической значимости и эффективности, планируемых и выполняемых НИР, используется метод бальных оценок. Бальная оценка заключается в том, что каждому фактору по принятой шкале присваивается определенное количество баллов. Обобщенную оценку проводят по сумме баллов по всем показателям или рассчитывают по формуле. На этой основе делается вывод о целесообразности НИР.

Сущность метода заключается в том, что на основе оценок признаков работы определяется коэффициент ее научно-технического уровня по формуле:

$$K_{НТУ} = \sum_{i=1}^3 R_i \cdot n_i, \quad (4.13)$$

где  $K_{НТУ}$  – коэффициент научно-технического уровня;

$R_i$  – весовой коэффициент  $i$ -го признака научно-технического эффекта;

$n_i$  – количественная оценка  $i$ -го признака научно-технического эффекта, в баллах.

Таблица 4.9 — Весовые коэффициенты признаков НТУ

Признаки научно-технического эффекта НИР	Характеристика признака НИР	Ri
Уровень новизны	Систематизируются и обобщаются сведения, определяются пути дальнейших исследований	0,4
Теоретический уровень	Разработка способа (алгоритм, программа мероприятий, устройство, вещество и т.п.)	0,1
Возможность реализации	Время реализации в течение первых лет	0,5

Таблица 4.10– Баллы для оценки уровня новизны

Уровень новизны	Характеристика уровня новизны – $n_1$	Баллы
Принципиально новая	Новое направление в науке и технике, новые факты и закономерности, новая теория, вещество, способ	8 – 10
Новая	По-новому объясняются те же факты, закономерности, новые понятия дополняют ранее полученные результаты	5 – 7
Относительно новая	Систематизируются, обобщаются имеющиеся сведения, новые связи между известными факторами	2 – 4
Не обладает новизной	Результат, который ранее был известен	0

Таблица 4.11 – Баллы значимости теоретических уровней

Теоретический уровень полученных результатов – $n_2$	Баллы
Установка закона, разработка новой теории	10
Глубокая разработка проблемы, многоспектральный анализ взаимодействия между факторами с наличием объяснений	8
Разработка способа (алгоритм, программа и т. д.)	6
Элементарный анализ связей между фактами (наличие гипотезы, объяснения версии, практических рекомендаций)	2
Описание отдельных элементарных факторов, изложение наблюдений, опыта, результатов измерений	0,5

Таблица 4.12 – Возможность реализации результатов по времени

Время реализации – $n_3$	Баллы
В течение первых лет	10
От 5 до 10 лет	4
Свыше 10 лет	2

Так как все частные признаки научно-технического уровня оцениваются по 10-балльной шкале, а сумма весов  $R_i$  равна единице, то величина интегрального показателя также принадлежит интервалу  $[0, 10]$ . В таблице 4.13 указано соответствие качественных уровней НИР значениям показателя, рассчитываемого по формуле (4.13).

Таблица 4.13 - Соответствие качественных уровней НИР значениям показателя

Уровень НТЭ	Показатель НТЭ
Низкий	1-4
Средний	4-7
Высокий	8-10

Для используемого в пособии примера частные оценки уровня  $n_i$  и их краткое обоснование даны в таблице 4.14.

Таблица 4.5 – Оценки научно-технического уровня НИР

<b>Значимость</b>	<b>Фактор НТУ</b>	<b>Уровень фактора</b>	<b>Выбранный балл</b>	<b>Обоснование выбранного балла</b>
0,4	Уровень новизны	Относительно новая	4	Облегчит процесс анализа сайтов
0,1	Теоретический уровень	Глубокая разработка проблемы, многоспектральный анализ взаимодействия между факторами с наличием объяснений	8	Разработка системы анализа основных факторов сайта, влияющих на поисковое продвижение
0,5	Возможность реализации	В течение первых лет	10	Время реализации в течение полу года

Отсюда интегральный показатель научно-технического уровня для нашего проекта составляет:  $I_{нту} = 0,4*4 + 0,1*8 + 0,5*10 = 1,6 + 0,8 + 5 = 7,8$

Таким образом, исходя из данных таблицы 4.12, данный проект имеет средний уровень научно-технического эффекта.

## **5 Социальная ответственность**

Целью выпускной квалификационной работы является разработка системы анализа факторов ранжирования web-ресурсов. Разрабатываемая система позволяет автоматизировать процесс анализа web-ресурса при поисковом продвижении. В данном разделе представлены и рассмотрены основные факторы, оказывающие влияние на работников предприятия согласно ГОСТ 12.0.003- 74 [21]. Основная часть выпускной квалификационной работы студента выполнялась с использованием персонального компьютера, поэтому основной целью данного раздела является обнаружение и анализ вредных и опасных факторов, влияющих на работу разработчика-программиста, а также методов защиты от них. Основными вредными факторами являются освещение, микроклимат помещения, а также рабочая поза, так как работа ведётся в закрытом помещении с использованием персонального компьютера. К физическим опасным факторам, воздействующим на программиста, относится опасность поражения электрическим током. Поэтому требуется изучение и создание оптимальных условий труда, а также следует учесть организацию пожарной безопасности на предприятии.

Работа над проектом осуществлялась в 418 аудитории, расположенной на кафедре Программной инженерии в Кибернетическом центре Томского политехнического университета с использованием персонального компьютера.

### **5.1 Производственная безопасность**

Работы по разработке системы анализа факторов ранжирования относятся к категории работ, связанных с опасными и вредными производственными факторами.

Вредные производственные факторы – факторы, воздействие которых на человека могут привести к его заболеванию и снижению работоспособности.

Опасные производственные факторы – факторы, воздействие которых могут привести к травме, отравлению, внезапному резкому ухудшению здоровья, или смерти человека.

### **5.3.1 Вредные факторы при разработке и эксплуатации проектируемого решения**

#### **5.1.1.1 Микроклимат**

Микроклимат помещения – это комплекс метеорологических условий в нем. Микроклимат является одним из наиболее важных факторов, влияющих на работоспособность и здоровье человека [22]. Основные факторы, характеризующие микроклимат помещения, устанавливаются в соответствии с СанПиН 2.2.4.548 – 96 [23]. К ним относятся:

- температура воздуха;
- скорость движения воздуха;
- влажность;
- интенсивность теплового облучения.

Согласно вышеуказанному документу [23], работа разработчика-программиста относится к категории работ 1б, так как основная часть работы происходит с использованием ПЭВМ.

При допустимых значениях микроклимата работник может ощущать небольшой дискомфорт и понижение работоспособности, при этом ухудшение состояния здоровья возникать не будет. При оптимальных значениях наблюдается высокий уровень работоспособности и обеспечивается нормальное состояние организма работника.

Допустимые и оптимальные значения показателей микроклимата холодного и тёплого периода года для категории работ 1б представлены в таблице 5.1.

Таблица 5.1 – Оптимальные и допустимые значения показателей микроклимата (СанПин 2.2.4.548-96)

Тип величины	Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Оптимальный	Холодный	21-23	20-24	40-60	0,1
	Тёплый	22-24	21-25		
Допустимый	Холодный	19-24	18-25	15-75	0,1-0,2
	Тёплый	20-28	19-29		0,1-0,3

Микроклимат комнаты поддерживается на оптимальном уровне системой водяного центрального отопления, естественной вентиляцией, а также искусственным кондиционированием в теплое время года и дополнительным прогревом в холодное время года. В помещениях также должно проводиться систематическое проветривание после каждого часа работы на ПЭВМ.

### 5.1.1.2 Освещение

Рабочее (общее) освещение – это основное освещение, которое обеспечивает нормальные условия для нахождения человека в помещении [24].

Правильное освещение является основным фактором, соблюдение которого требуется для снижения развития профессиональных болезней при работе с ПЭВМ и увеличения работоспособности. Отсутствие хорошего освещения может привести к профессиональным заболеваниям, а также ухудшению концентрации работников. Работа разработчика-программиста в основном проводится за дисплеем персонального компьютера. В результате недостаточной освещённости рабочего места или при избыточном освещении помещения у работника может возникнуть переутомление, а также ухудшится зрение.

Для комфортной работы необходимо создать систему освещения рабочего места, отвечающую нормам, указанным в соответствии со СНиП 23-05-95 [24]. Согласно вышеуказанному документу [24], освещённость рабочего места должна быть равна 200 лк, так как работа программиста за компьютером относится к 3 категории зрительной работы.

Рассчитаем освещённость рассматриваемой аудитории, размеры которой указаны в пункте «Характеристика рабочего места». Необходимо рассчитать индекс помещения по формуле (5.1):

$$i = \frac{S}{h \cdot (A + B)} \quad (5.1)$$

где  $i$  – индекс помещения;

$S = 30 \text{ м}^2$ , площадь помещения;

$h = 2,5 \text{ м}$ , высота помещения;

$A = 6 \text{ м}$ , длина помещения;

$B = 5 \text{ м}$ , ширина помещения.

В результате вычислений индекс помещения  $i = 1,1$ . Коэффициент использования светового потока с люминесцентными лампами равен 43%. Рассчитаем освещённость аудитории, исходя из того, что в ней установлено 7 светильников по 2 лампы в каждом:

$$E_{\text{факт}} = \frac{N \cdot n \cdot \Phi_{\text{ст}} \cdot \eta}{S \cdot K_3 \cdot Z} \quad (5.2)$$

где  $E_{\text{факт}}$  – фактическая освещённость;

$N = 7$ , число светильников в помещении;

$n = 2$ , число ламп в светильнике;

$\Phi_{\text{ст}} = 1650 \text{ лм}$ , величина стандартного светового потока (для люминесцентных ламп мощности 36Вт);

$\eta = 43\%$ , коэффициент использования светового потока;

$S = 30 \text{ м}^2$ , площадь помещения;

$K_3 = 1,5$ , коэффициент запаса для помещений с малым выделением пыли;

$Z = 1,1$ , коэффициент неравномерности освещения люминесцентных ламп.

В результате фактическая освещённость  $E_{\text{факт}} = 200$  лк равна нормативной освещённости.

$$\Delta E = \frac{(E_{\text{факт}} - E_{\text{н}})}{E_{\text{н}}} \cdot 100\% \quad (5.3)$$

где  $\Delta E$  – показатель разности между фактической и нормативной освещённостью;

$E_{\text{факт}} = 200$  лк, фактическая освещённость;

$E_{\text{н}} = 200$  лк, нормативная освещённость.

В результате показатель разности  $\Delta E = 0\%$ . Из этого следует, что фактическое значение освещённости соответствует нормативному, что говорит о хорошей системе освещения в аудитории.

### **5.3.2 Опасные факторы при разработке и эксплуатации проектируемого решения**

Первым опасным фактором является опасность поражения электрическим током. Исходя из анализа состояния помещения, данную лабораторию по степени опасности поражения электрическим током можно отнести к классу помещений без повышенной опасности.

В помещении подавляющая часть электрической проводки является скрытой. Поражение электрическим током возможно только при возникновении оголенных участков на кабеле, а также нарушении изоляции распределительных устройств, однако кабель имеет двойную изоляцию, поэтому опасность поражения значительно снижается. Не исключается также опасность поражения и от токоведущих частей компьютера в случае их пробоя и нарушении изоляции. Кроме того, компьютер является основным источником статического электричества. Местами скопления статических зарядов, как

правило, служит поверхность экрана монитора. Частое поражение статическим электричеством может привести к возникновению кожных заболеваний и нервно-психическим расстройствам.

Возникновение пожара является вторым опасным производственным фактором, т.к. пожар на предприятии наносит большой материальный ущерб, а также часто сопровождается травмами и несчастными случаями. Регулирование пожаробезопасности производится СНиП 21-01-97.

В помещениях с ПЭВМ повышен риск возникновения пожара из-за присутствия множества факторов: наличие большого количества электронных схем, устройств электропитания, устройств кондиционирования воздуха; возможные неисправности электрооборудования, освещения, или неправильная их эксплуатация может послужить причиной пожара.

Возможные виды источников воспламенения:

- Искра при разряде статического электричества;
- Искры от электрооборудования;
- Искры от удара и трения;
- Открытое пламя [25].

## **5.2 Экологическая безопасность**

### **5.3.1 Источники загрязнения**

В данном разделе рассматривается воздействие на окружающую среду деятельности по разработке проекта, а также самого продукта в результате его реализации на производстве.

Работа с ПЭВМ не несет в себе потенциальной экологической опасности, поэтому помещение, в котором проходила разработка, относятся к пятому классу с размером санитарно-защитных зон равным 50 метрам [26].

Непосредственно программный продукт, разработанный в ходе выполнения магистерской диссертации, не наносит вреда окружающей среде ни

на стадиях его разработки, ни на стадиях эксплуатации. Однако, средства, необходимые для его разработки и эксплуатации могут наносить вред окружающей среде.

При разработке любых автоматизированных систем возникает необходимость утилизировать производственные отходы, в качестве которых в данном случае выступают бумажные отходы (макулатура) и неисправные детали персональных компьютеров, плат, контроллеров.

Под защитой окружающей среды подразумевается утилизация отходов оргтехники, которые содержат различные долго разлагающиеся пластмассы, люминесцентные лампы, а также вредные вещества и их соединения: свинец, ртуть, мышьяк и т.д. Утилизировать их вместе с бытовыми отходами нельзя.

### **5.3.2 Мероприятия по предотвращению загрязнения**

Утилизация электрооборудования производится специальными организациями, имеющими лицензию на такую деятельность [27]. В настоящее время в Томской области утилизацией занимаются две компании: городской полигон и ООО НПП «Экотом». Сбор опасных отходов в специальный контейнер осуществляется по заявке физических и юридических лиц, поэтому в случае необходимости нужно оформить соответствующую заявку в вышеуказанные компании.

Бумажные отходы должны передаваться в соответствующие организации для дальнейшей переработки во вторичные бумажные изделия. Еще одним из способов снижения бумажных отходов является хранение данных на электронных носителях. Данное ПО позволяет хранить данные на жестком диске в электронном виде.

## **5.3 Безопасность в чрезвычайных ситуациях**

### **5.3.1 Кибертерроризм**

Под кибертерроризмом понимают совокупность противоправных действий, связанных с угрозами безопасности личности, обществу и государству, деструктивными действиями в отношении материальных объектов, искажением объективной информации или другими действиями с целью получения преимущества при решении политических, экономических или социальных задач.

Арсенал компьютерных террористов – различные вирусы, логические бомбы – команды, встроенные заранее в программу и срабатывающие в нужный момент. Современные террористы используют Интернет в основном как средство пропаганды, передачи информации, а не как новое оружие. Однако можно предполагать, что компьютерный терроризм сегодня уже представляет реальную угрозу обществу. В настоящее время существует весьма мало систем, которые можно назвать надежно защищенным [28].

Разрабатываемая система позволяет сохранять результаты анализа web-ресурсов на сервере, которые могут быть подвержены кибератакам. Для предотвращения и нейтрализации последствий кибератак необходимо принять следующие меры:

- обеспечить разграничение доступа в помещения (или на объекты), где находятся информационные ресурсы;
- установить системы защиты от несанкционированного доступа к ресурсам компьютеров и локальных сетей;
- защитить (межсетевыми экранами) входы в сеть извне, в том числе и при наличии соединения с Internet;
- организовать постоянный мониторинг и аудит сети;
- обеспечение антивирусной защиты;

Кроме того, существуют различные организационные меры, такие как, например, психологическая проверка сотрудников, запрет на вынос дискет или распечаток и, наоборот, на внос видео или фотоаппаратуры и т.д. Конечно, все принимаемые меры по обеспечению безопасности позволят значительно снизить риск и тем самым сохранить информацию и здоровье ее владельца.

### **5.3.2 Пожарная безопасность**

Пожарная безопасность представляет собой единый комплекс организационных, технических, режимных и эксплуатационных мероприятий по предупреждению пожаров и взрывов. Так как пожар является одним из наиболее вероятных чрезвычайных ситуаций, то помещения должны быть обеспечены всем необходимым для быстрой локализации и/или уничтожения очага [29].

По взрыво- и пожароопасности все помещения, согласно техническому регламенту НПБ 105-95 [29], делятся на 5 категорий, в зависимости от применяемых на производстве веществ и их количества. Учебные аудитории Кибернетического центра относятся к категории Г, так как в качестве горючего материала выступает бумага.

Возможными источниками и причинами пожара в помещении являются:

- короткое замыкание в электропроводке вследствие неисправности самой проводки или электrorаспределительных щитов;
- неисправная аппаратура;
- нарушение правил пожарной безопасности работниками.

Каждый работник при обнаружении пожара, признаков горения, задымления, запаха гари, повышения температуры и т.п. должен [30]:

- незамедлительно сообщить об это в пожарную охрану, поставить в известность службу охраны;

- принять по возможности меры по эвакуации людей и материальных ценностей в соответствии с планом эвакуации;
- по возможности отключить электроэнергию и приступить к тушению пожара первичными средствами пожаротушения, не подвергая свою жизнь опасности.

Для предотвращения возникновения пожароопасных ситуаций необходимо для работников проводить следующие мероприятия [25]:

- проведение противопожарных инструктажей для работников предприятия;
- определение действий работников предприятия в случае возникновения пожара;
- разработка плана эвакуации работников предприятия;
- размещение планов эвакуации, плакатов с инструкциями на каждом этаже;
- проведение инструктажей по работе за оборудованием;
- плановый осмотр и своевременный ремонт оборудования;
- проверка средств пожаротушения и средств оказания первой помощи пострадавшим.

Учебные аудитории Кибернетического центра оснащены ручными углекислотными огнетушителями ОУ-2, а также аптечками первой помощи согласно требованиям [31]. Так же во всех аудиториях установлены датчики дыма, которые срабатывают в случае возникновения пожара и задымления. Также на всех этажах должны быть вывешены планы эвакуации.

## **5.4 Правовые и организационные вопросы обеспечения безопасности**

### **5.4.1 Организационные мероприятия при компоновке рабочей зоны**

Учебная аудитория, в которой выполнялась работа, располагается на четвертом этаже Кибернетического центра Научно-исследовательского Томского политехнического университета и имеет следующие параметры: длина – 6 м, ширина – 5 м, высота – 2,5 м. Для естественного освещения аудитория оснащена двумя окнами, имеющими одинаковые размеры: ширина – 1,7 м, высота – 1,5 м. Дверь изготовлена из металла и имеет серый цвет. На подвесном потолке установлена прессованная плитка из полистирола. На стенах поклеены обои белого цвета. Напольным покрытием является линолеум. Площадь аудитории равна 30 м<sup>2</sup>, объём – 75 м<sup>3</sup>.

Согласно документу СанПиН 2.2.2/2.4.1340-03 [32], в котором указаны нормы к помещениям для работы с ПЭВМ, норма площади рабочего места с персональным компьютером составляет 4,5 м<sup>2</sup>. В рассматриваемой аудитории установлено шесть рабочих мест с персональными компьютерами и жидкокристаллическими экранами. Соответственно, на одного человека приходится 5 м<sup>2</sup>, что соответствует вышеуказанным требованиям.

Главными элементами рабочего места являются стол и кресло. Основным рабочим положением является положение сидя.

Рабочая поза сидя вызывает минимальное утомление сотрудника. Рациональная планировка рабочего места предусматривает четкий порядок и постоянство размещения предметов, средств труда и документации. То, что требуется для выполнения работ чаще, расположено в зоне легкой досягаемости рабочего пространства.

Оптимальная зона моторного поля рабочего места, ограничена дугами, описываемыми предплечьями при движении в локтевых суставах с опорой в точке локтя и с относительно неподвижным плечом (рисунок 5.1).

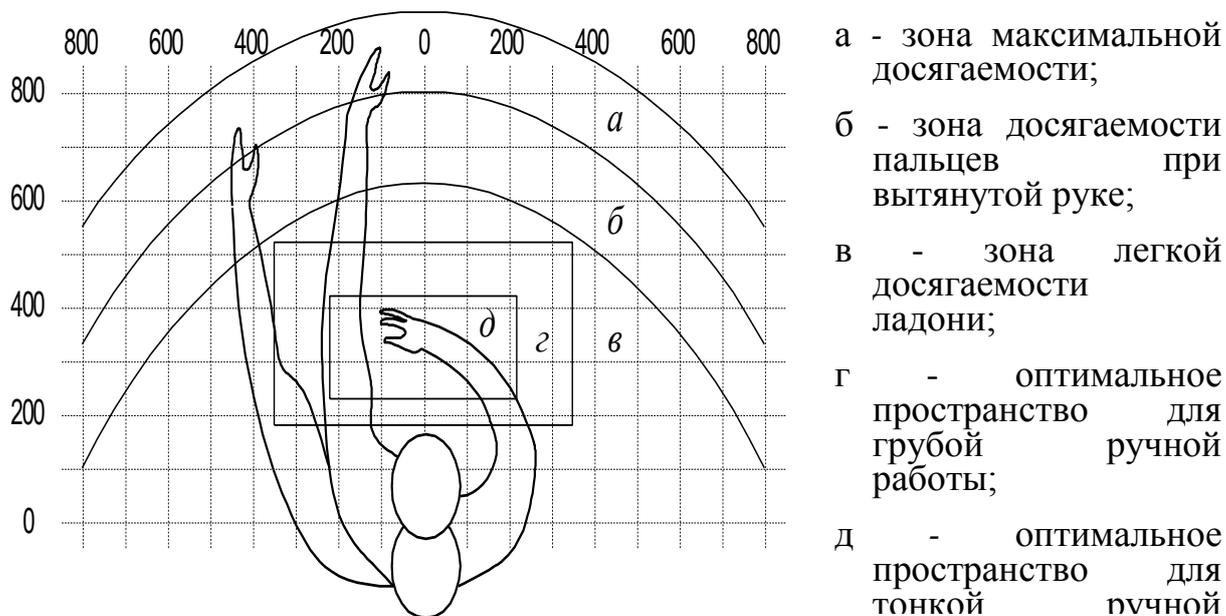


Рисунок 5.1. Зоны досягаемости рук в горизонтальной плоскости

Оптимальное размещение предметов труда и документации в зонах досягаемости:

- Дисплей размещается в зоне а (в центре);
- системный блок размещается в предусмотренной нише стола;
- клавиатура - в зоне г/д;
- «мышь» - в зоне в справа;
- сканер в зоне а/б (слева);
- принтер находится в зоне а (справа);
- документация: необходимая при работе - в зоне легкой

досягаемости ладони – в, а в выдвижных ящиках стола - остальная литература.

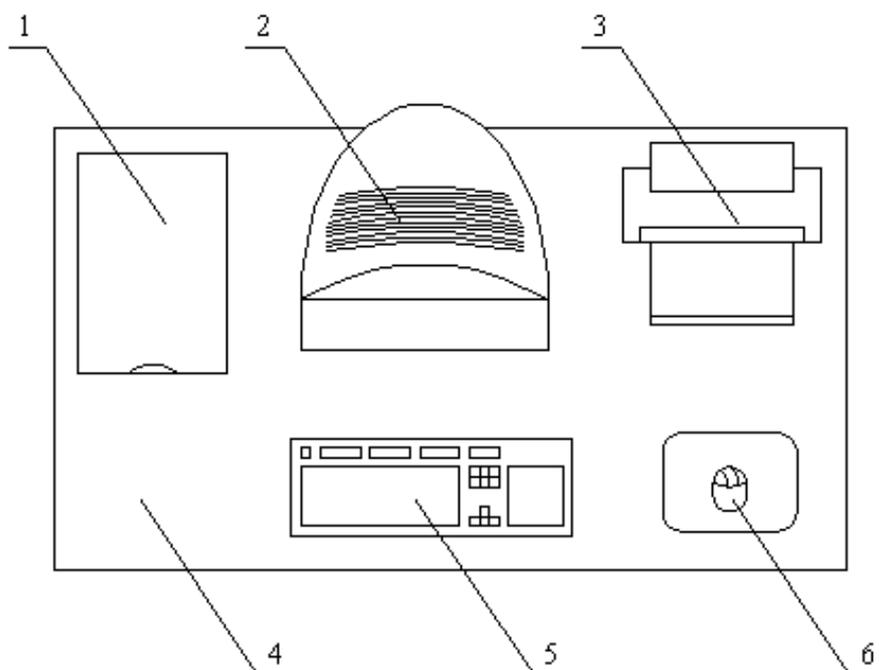


Рисунок 5.2. Размещение основных и периферийных составляющих

На рисунке 5.2 показан пример размещения основных и периферийных составляющих ПК на рабочем столе сотрудника.

1 – сканер, 2 – монитор, 3 – принтер, 4 – поверхность рабочего стола,  
5 – клавиатура, 6 – манипулятор типа «мышь».

Для комфортной работы стол должен удовлетворять следующим условиям [33]:

- высота стола должна быть выбрана с учетом возможности сидеть свободно, в удобной позе, при необходимости опираясь на подлокотники;
- нижняя часть стола должна быть сконструирована так, чтобы сотрудник мог удобно сидеть, не был вынужден поджимать ноги;
- поверхность стола должна обладать свойствами, исключающими появление бликов в поле зрения сотрудника;
- конструкция стола должна предусматривать наличие выдвижных ящиков (не менее 3 для хранения документации, листингов, канцелярских принадлежностей);

- высота рабочей поверхности рекомендуется в пределах 680-760мм.

Высота поверхности, на которую устанавливается клавиатура, должна быть около 650мм.

Большое значение придается характеристикам рабочего кресла. Так, рекомендуемая высота сиденья над уровнем пола находится в пределах 420-550 мм. Поверхность сиденья мягкая, передний край закругленный, а угол наклона спинки - регулируемый.

Необходимо предусматривать при проектировании возможность различного размещения документов: сбоку от видеотерминала, между монитором и клавиатурой и т.п. Кроме того, в случаях, когда видеотерминал имеет низкое качество изображения, например, заметны мелькания, расстояние от глаз до экрана делают больше (около 700 мм), чем расстояние от глаза до документа (300-450 мм). При высоком качестве изображения на видеотерминале расстояние от глаз пользователя до экрана, документа и клавиатуры может быть равным.

Положение экрана определяется:

- расстоянием считывания (0,6...0,7м);
- углом считывания, направлением взгляда на  $20^\circ$  ниже горизонтали

к центру экрана, причем экран перпендикулярен этому направлению.

Должна также предусматриваться возможность регулирования экрана:

- по высоте +3 см;
- по наклону от  $-10^\circ$  до  $+20^\circ$  относительно вертикали;
- в левом и правом направлениях.

Таким образом, конструкция оборудования и организация рабочего места должны обеспечивать оптимальное положение работающего, соответствующее ГОСТ 12.2.033-78 ССБТ [34]. Требования к рабочему месту при выполнении работ сидя указаны в таблице 5.2.

Таблица 5.2 - Рекомендуемые параметры рабочего пространства РМ операторов  
ПЭВМ

№	Параметры	Величина, в мм
1.	Ширина пространства	1050
2.	Глубина пространства	400
3.	Высота рабочей поверхности стола над полом (регулируется)	750 (650.. .850)
4.	Расстояние между высотой рабочей поверхности и креслом оператора	280
5.	Пространство для ног: • высота • ширина • высота кресла • глубина	600 500 400-500 (регулируется) 400
6.	Угол отклонения спинки кресла к поверхности сидения	100- 115 градусов
7.	Ширина сидения кресла	550
8.	Глубина сидения кресла	450
9.	Ширина подлокотников	50
10.	Длина подлокотников	200
11.	Высота над поверхностью сидения	250
12.	Радиус изгиба кресла	350
13.	Подставка для ног	400x300x150
14.	Угол наклона	0-20 градусов

Основное значение придается правильной рабочей позе пользователя. При неудобной рабочей позе могут появиться боли в мышцах, суставах и сухожилиях. Требования к рабочей позе пользователя видеотерминала следующие:

- голова не должна быть наклонена более чем на 20°;
- плечи должны быть расслаблены;

- локти - под углом 80°...100°;
- предплечья и кисти рук - в горизонтальном положении.

В целях преодоления указанных недостатков даются общие рекомендации: лучше передвижная клавиатура; должны быть предусмотрены специальные приспособления для регулирования высоты стола, клавиатуры и экрана, а также подставка для рук [33].

#### **5.4.2 Правовые нормы трудового законодательства для рабочей зоны оператора ПЭВМ**

В соответствии с государственными стандартами и правовыми нормами обеспечения безопасности предусмотрена рациональная организация труда в течение смены, которая предусматривает:

- длительность рабочей смены не более 8 часов;
- установление двух регламентируемых перерывов (не менее 20 минут после 1-2 часов работы, не менее 30 минут после 2 часов работы);
- обеденный перерыв не менее 40 минут.

Обязательно предусмотрен предварительный медосмотр при приеме на работу и периодические медосмотры.

Каждый сотрудник должен пройти инструктаж по технике безопасности перед приемом на работу и в дальнейшем, должен быть пройден инструктаж по электробезопасности и охране труда.

Предприятие обеспечивает рабочий персонал всеми необходимыми средствами индивидуальной защиты.

Оплата труда, социальные пособия, дополнительные выплаты устанавливаются в соответствии со степенью вредности и опасности выполняемых обязанностей. В случае задержки заработной платы более чем на 15 дней, работник имеет право приостановить работу, письменно уведомив

работодателя. Однако, в некоторых случаях возможно удержание заработной платы [35].

### **5.4.3 Предотвращение нервно-эмоционального напряжения**

При работе с ПЭВМ основным фактором, влияющим на нервную систему программиста или пользователя, является огромное количество информации, которое он должен воспринимать. Это является сложной задачей, которая очень сильно влияет на сознание и психофизическое состояние из-за монотонности работы. Поэтому меры, позволяющие снизить воздействие этого вредного производственного фактора, которые регулируются СанПиН 2.2.2/2.4.1340-03, являются важными в работе оператора ПЭВМ. Они позволяют увеличить производительность труда и предотвратить появление профессиональных болезней.

Организация работы с ПЭВМ осуществляется в зависимости от вида и категории трудовой деятельности. Виды трудовой деятельности разделяются на 3 группы: группа А – работа по считыванию информации с экрана с предварительным запросом; группа Б – работа по вводу информации; группа В – творческая работа в режиме диалога с ПЭВМ. Работа программиста-разработчика рассматриваемой в данной работе системы относится к группам А и Б. Для снижения воздействия рассматриваемого вредного фактора предусмотрены регламентированные перерывы для каждой группы работ (Таблица 5.3).

Таблица 5.3 – Суммарное время регламентированных перерывов в зависимости от продолжительности работы, вида категории трудовой деятельности с ПЭВМ [12]

Категория работы с ПЭВМ	Уровень нагрузки за рабочую смену при видах работ с ПЭВМ		Суммарное время регламентированных перерывов, мин.	
	группа А, количество знаков	группа Б, количество знаков	при 8-часовой смене	при 12-часовой смене
I	до 20 000	до 15 000	50	80
II	до 40 000	до 30 000	70	110
III	до 60 000	до 40 000	90	140

#### 5.4.4 Влияние использования программного продукта на рабочий процесс

Разрабатываемая система предназначена для автоматизации процесса анализа сайтов при продвижении в поисковых системах. Процесс анализа сайтов при ручном выполнении занимает от одного до двух часов на одну страницу сайта, с помощью разработанной системы этот процесс будет занимать около пяти минут.

В месяц один работник анализирует 5 – 8 сайтов с количеством страниц от одной до 30. Таким образом среднее количество сайтов в месяц  $\frac{5+8}{2} = 6.5$ , а количество страниц  $\frac{1+30}{2} = 15.5$ .

Для подсчета времени на проведение анализа в месяц используем следующую формулу:

$$T = N * M * t \quad (5.4)$$

где N – среднее количество сайтов;

M – среднее количество страниц сайтов;

t – время проведения анализа.

При ручном проведении анализа получается:

$$6,5 * 15,5 * 1,5 = 151,125 \text{ часа.}$$

При проведении анализа с использованием системы:

$$6,5 * 15,5 * 0,083 = 8,4 \text{ часа.}$$

В результате с помощью программы время проведения анализа в среднем уменьшится на 142,725 часа.

Таким образом с использованием системы анализ будет происходить намного быстрее, следовательно, работник сможет выполнять остальную работу по продвижению более качественно, либо выполнять больший объем работы.

## **Заключение**

В данной магистерской диссертации была поставлена цель разработать систему анализа факторов ранжирования, предназначенную для автоматического выполнения анализа страницы сайта при поисковом продвижении. Также были сформулированы следующие задачи:

1. Выбрать факторы ранжирования.
2. Определить эталонные значения.
3. Рассмотреть существующие системы.
4. Определить функциональные требования.
5. Разработать архитектуру системы.
6. Разработать пользовательский интерфейс.
7. Провести внедрение системы.

В качестве факторов ранжирования были выбраны Title, Description, Keywords, текст на странице сайта, заголовки h1-h6, изображения, скорость загрузки страницы, favicon, человеко-понятные URL, файл robots.txt и карта сайта. После изучения рекомендаций по оптимизации этих факторов были выделены их эталонные значения, которые использовались для анализа. Так же был проведен обзор аналогичных систем, где были выявлены их недостатки, которые были учтены в требованиях к системе.

Информационная система была спроектирована и реализована согласно поставленным функциональным требованиям. Архитектура системы представлена в виде диаграммы потоков данных, а также были декомпозированы некоторые основные процессы. Разработана база данных для хранения информации об учетной записи пользователя и обработанных данных с сайта, и интуитивно-понятный интерфейс.

В результате работы была разработана система для анализа факторов ранжирования web-ресурсов, которая выполняет экспресс-анализ сайта и предоставляет подробный отчет об ошибках оптимизации на странице сайта с

рекомендациями по их устранению. Разработанная система соответствует всем сформулированным требованиям, таким образом имеет некоторые преимущества перед аналогами.

## Список публикаций

1. Бондарчук Н. В. Система для анализа внутренних параметров Web-ресурса / Н. В. Бондарчук; науч. рук. А. О. Савельев // Молодежь и современные информационные технологии: сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых, г. Томск, 7-11 ноября 2016 г. : в 2 т. — Томск : Изд-во ТПУ, 2016. — Т. 2. — [С. 106-107].

## Список использованных источников:

1. 130 факторов ранжирования: «ПОЛНЫЙ СПИСОК!». [Электронный ресурс]. – Режим доступа: <http://convertmonster.ru/blog/seo-blog/osnovnye-principy-ranzhirovaniya-sajta-v-poiskovyh-sistemah/>
2. И. Ашманов, А. Иванов. Оптимизация и продвижение сайта в поисковых системах: - М. Издательство: Питер, 2013 – 464 с.
3. И. Севостьянов. Поисковая оптимизация. Практическое руководство по продвижению сайта в Интернете. – М. Издательство «Питер», 2010 – 240 с.
4. Л. Гроховский, М. Сливинский, А. Чекушин, С. Ставский. SEO: руководство по внутренним факторам. – М.: Центр исследований и образования «ТопЭксперт.РФ», 2011. – 133 с.
5. Как заставить поисковики полюбить ваш сайт — руководство по SEO для начинающих: Глава 4 [Электронный ресурс]. – Режим доступа: <http://blog.contentmonster.ru/2014/06/seo-guide-4>
6. Тег title — название страницы. [Электронный ресурс]. – Режим доступа: <http://seo-in.ru/tag-title.html>
7. Мета-теги в SEO (Description и Keywords). [Электронный ресурс]. – Режим доступа: <http://seo-in.ru/meta-teg.html>
8. Как Правильно Использовать HTML Мета-тег Keywords. [Электронный ресурс]. – Режим доступа: <http://vsemproblemam.net/seo/kak-pravilno-ispolzovat-html-meta-teg-keywords.html>
9. Релевантность. [Электронный ресурс]. – Режим доступа: <https://www.webeffector.ru/wiki/%D0%A0%D0%B5%D0%BB%D0%B5%D0%B2%D0%B0%D0%BD%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%8C>
10. SEO оптимизация картинок и изображений для сайта. [Электронный ресурс]. – Режим доступа: <https://convertmonster.ru/blog/seo-blog/seo-optimizacija-izobrazhenij-kartinok-na-sajte-dlja-poiskovyh-sistem/>

11. Заголовки (теги h1, h2, h3) в HTML5 и SEO заблуждения. [Электронный ресурс]. – Режим доступа: <http://shpargalkablog.ru/2015/02/h1-seo.html>
12. Что необходимо знать о favicon.ico. [Электронный ресурс]. – Режим доступа: <https://devaka.ru/articles/favicon-in-seo>
13. Robots.txt — инструкция для SEO. [Электронный ресурс]. – Режим доступа: <https://seoprofy.ua/blog/optimizaciya-sajtov/robots-txt>
14. Карта сайта. [Электронный ресурс]. – Режим доступа: [https://www.webeffector.ru/wiki/%D0%9A%D0%B0%D1%80%D1%82%D0%B0\\_%D1%81%D0%B0%D0%B9%D1%82%D0%B0](https://www.webeffector.ru/wiki/%D0%9A%D0%B0%D1%80%D1%82%D0%B0_%D1%81%D0%B0%D0%B9%D1%82%D0%B0)
15. Влияет ли время загрузки страницы сайта на его позиции. [Электронный ресурс]. – Режим доступа: <https://seo-akademiya.com/baza-znaniij/vnutrennyaya-optimizacziya/vremya-zagruzki-sajta/>
16. SEO и скорость загрузки сайтов. [Электронный ресурс]. – Режим доступа: [https://www.darvin-studio.ru/about/articles/state-seo\\_speed.html](https://www.darvin-studio.ru/about/articles/state-seo_speed.html)
17. Что такое ЧПУ URL и в чем его преимущества? [Электронный ресурс]. – Режим доступа: <http://pgdv.ru/samostoyatelno/url-chpu>
18. Что такое уникальность текста и как ее определить. [Электронный ресурс]. – Режим доступа: <http://www.rabota-kopiraiterom.ru/pervyj-vkhod-na-advego/unikalnost-teksta-na-advego.html>
19. Яндекс.Помощь. [Электронный ресурс]. – Режим доступа: <https://yandex.ru/support/webmaster/recommendations/intro.html>
20. Руководство по поисковой оптимизации для начинающих. [Электронный ресурс]. – Режим доступа: <https://static.googleusercontent.com/media/www.google.ru/ru/ru/intl/ru/webmasters/docs/search-engine-optimization-starter-guide-ru.pdf>

21. ГОСТ 12.0.003-74 (1999) ССБТ Опасные и вредные производственные факторы. Классификация. – М.: Издательство стандартов, 1986.
22. ГОСТ 12.1.005-88. Общие санитарно-гигиенические требования к воздуху рабочей зоны. – М.: Стандартинформ, 2006.
23. СанПиН 2.2.4.548 – 96. Гигиенические требования к микроклимату производственных помещений. – М.: Информационно-издательский центр Минздрава России, 1997.
24. СНиП 23-05-95. Естественное и искусственное освещение. – М.: Центр проектной продукции в строительстве, 2011.
25. СНиП 21-01-97. Пожарная безопасность зданий и сооружений. – М.: Центр проектной продукции в строительстве, 2011.
26. СанПиН 2.2.1/2.1.1.1200-03 Санитарно-защитные зоны и санитарная классификация предприятий, сооружений и иных объектов. – М.: Госкомсанэпиднадзор России, 2003.
27. Постановление Правительства Российской Федерации от 25.04.2012 N 390 "О противопожарном режиме".
28. В. А. Мазуров. Кибертерроризм: понятие, проблемы противодействия URL: <http://cyberleninka.ru/article/n/kiberterrorizm-ponyatie-problemy-protivodeystviya>
29. НПБ 105-95. Определение категорий помещений и зданий по взрывопожарной и пожарной опасности. / Шебеко Ю.Н. – М.: ВНИИПО, 1998.
30. ППБ 01-03. Правила пожарной безопасности в Российской Федерации. – М.: ФГУ ВНИИПО МЧС России, 2003.
31. ГОСТ Р 51057-01. Огнетушители переносные. Общие технические требования. Методы испытаний. – М.: Издательство стандартов, 2001.

32. СанПиН 2.2.2/2.4.1340-03. Гигиенические требования к персональным электронно-вычислительным машинам и организации работы. – М.: Информационно-издательский центр Минздрава России, 2003.

33. Гигиенические требования к ВДТ, ПЭВМ и организации работы. Санитарные правила и нормы 2.2.2.542 – 96. – М., 1996

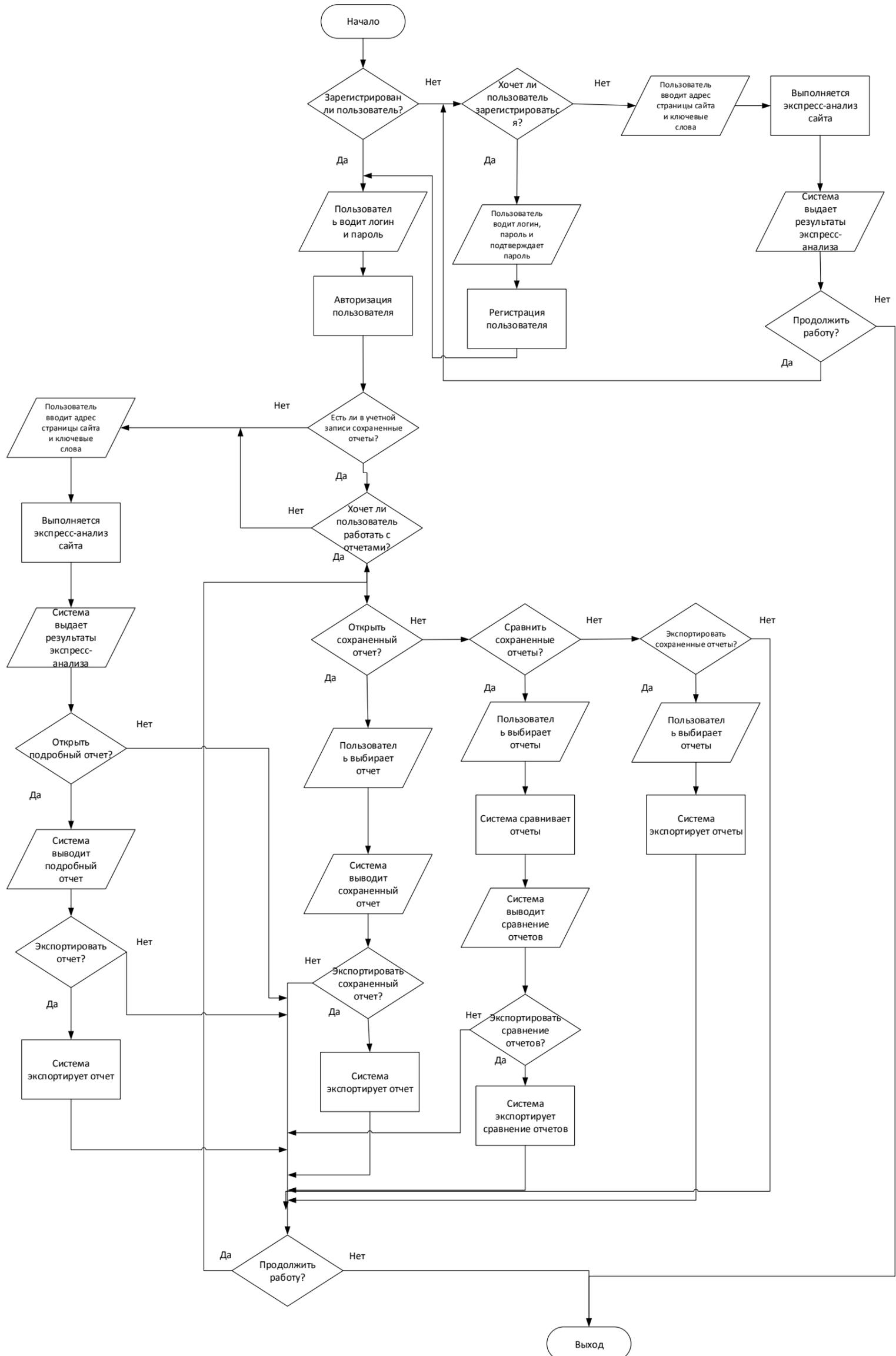
34. ГОСТ 12.2.033-78 ССБТ. Рабочее место при выполнении работ стоя. Общие эргономические требования. Методы испытаний. – М.: Издательство стандартов, 1986.

35. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 03.07.2016). Ст. 137. Ограничение удержаний из заработной платы // Консультант Плюс. 2015. URL: [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_34683/0b97a8e40bb5179ac150db9bb34324c945278f2d/](http://www.consultant.ru/document/cons_doc_LAW_34683/0b97a8e40bb5179ac150db9bb34324c945278f2d/)

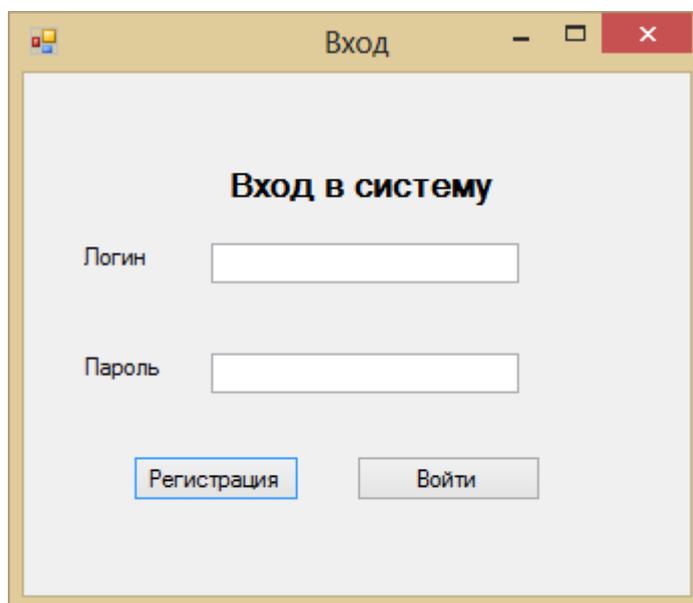
# Приложение А

(обязательное)

## Обобщённый алгоритм работы системы

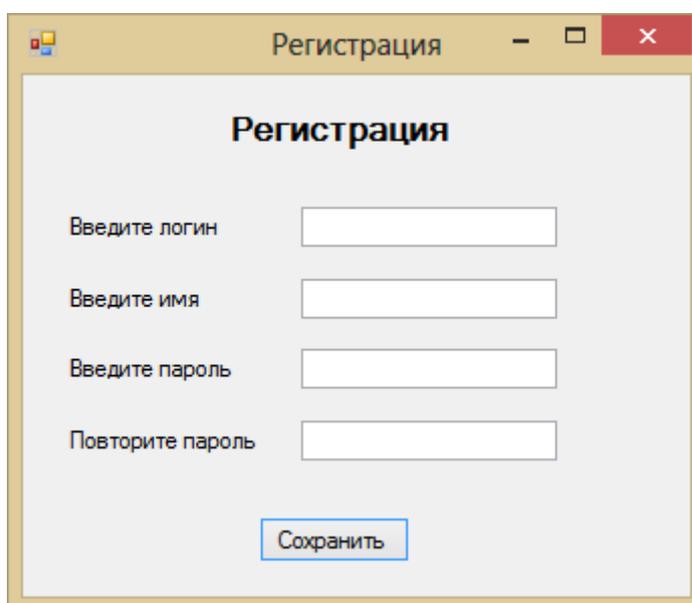


**Приложение Б**  
**(обязательное)**  
**Пользовательский интерфейс**



The screenshot shows a window titled "Вход" (Login). The main heading is "Вход в систему" (Login to the system). Below the heading are two input fields: "Логин" (Login) and "Пароль" (Password). At the bottom, there are two buttons: "Регистрация" (Registration) and "Войти" (Login).

Рисунок Б.1 - Форма входа в систему



The screenshot shows a window titled "Регистрация" (Registration). The main heading is "Регистрация" (Registration). Below the heading are four input fields: "Введите логин" (Enter login), "Введите имя" (Enter name), "Введите пароль" (Enter password), and "Повторите пароль" (Repeat password). At the bottom, there is a "Сохранить" (Save) button.

Рисунок Б.2 - Форма регистрации

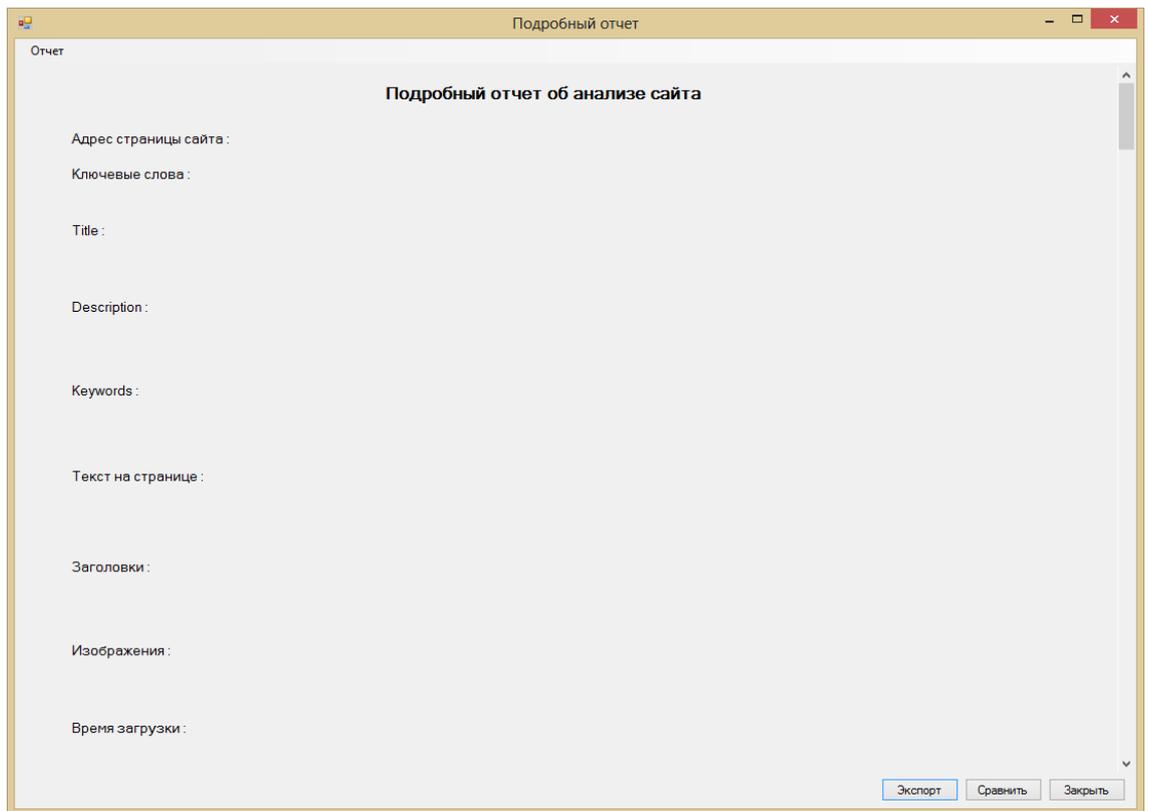


Рисунок Б.3 - Форма подробного отчета

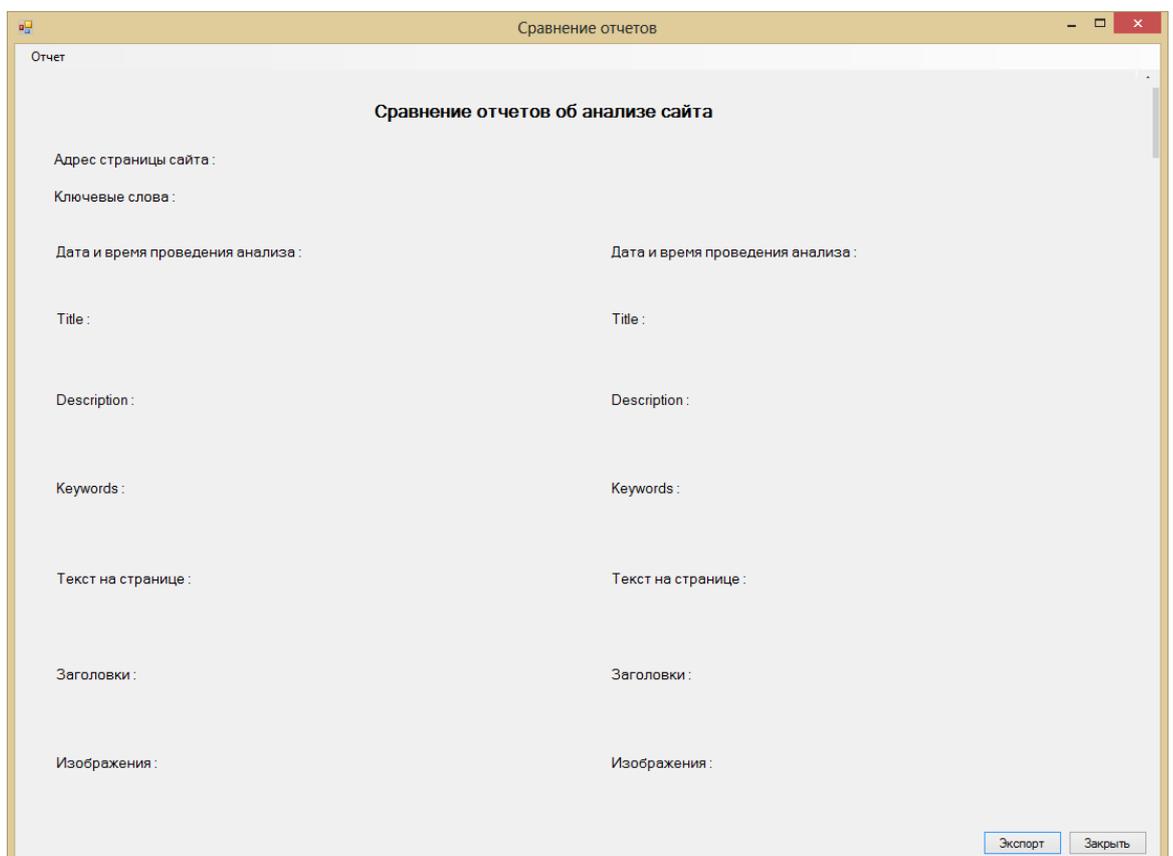


Рисунок Б.4 - Форма сравнения отчетов

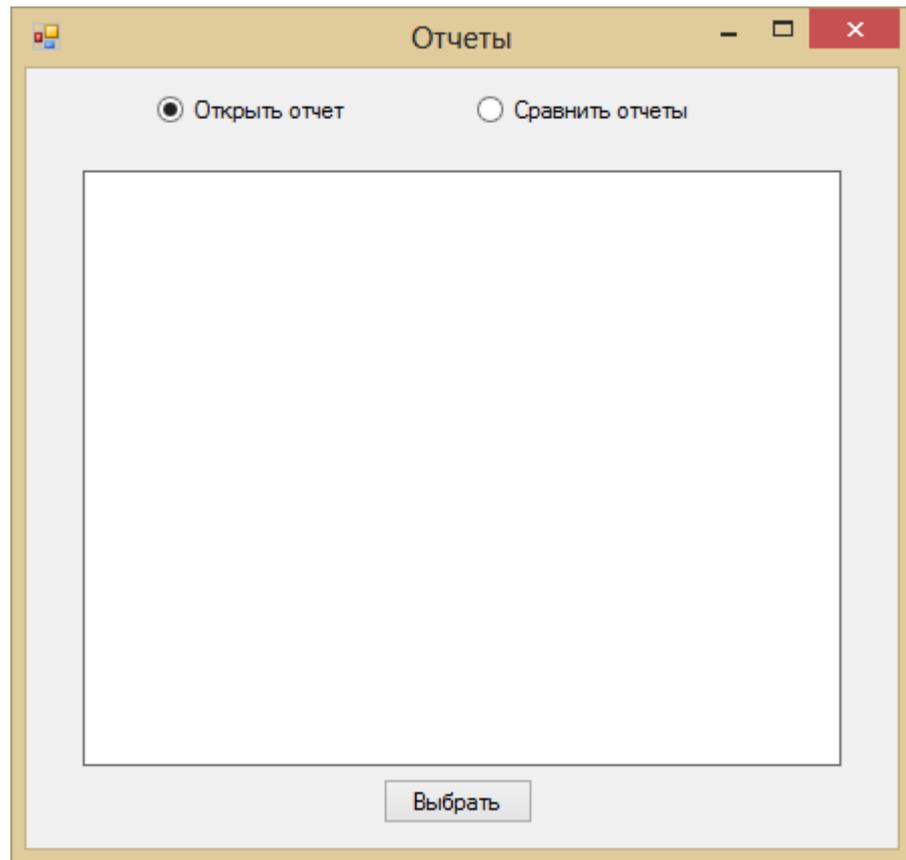


Рисунок Б.5 - Форма "Отчеты"

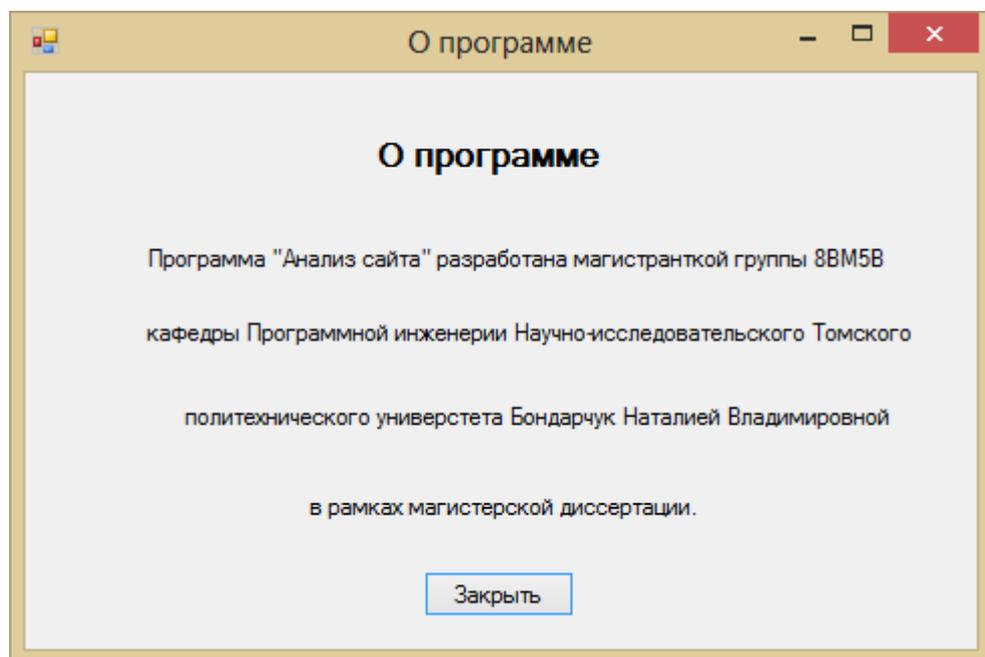


Рисунок Б.6 - Форма "О программе"

## Приложение В

(обязательное)

### Акт внедрения

«УТВЕРЖДАЮ»



Генеральный директор  
ООО «Нева-Томь Консалт»  
Д.В. Бибко

**АКТ**

внедрения системы анализа факторов ранжирования web-ресурсов

Комиссия в лице генерального директора Бибко Дениса Владимировича ООО «Нева-Томь Консалт» рассмотрела научно-исследовательскую работу магистранта кафедры Программной инженерии Института кибернетики Национального исследовательского Томского политехнического университета Бондарчук Наталии Владимировны «Разработка системы анализа факторов ранжирования web-ресурсов», выполненную под руководством ассистента, к.т.н. Савельева Алексея Олеговича.

Разработанная в рамках научно-исследовательской работы Бондарчук Наталией Владимировной система анализа факторов ранжирования web-ресурсов передана в ООО «Нева-Томь Консалт» с целью применения для автоматизации процесса анализа сайтов при поисковом продвижении.

Генеральный директор \_\_\_\_\_ / Бибко Д. В.

## Приложение Г

(обязательное)

### Раздел ВКР, выполненный на иностранном языке

#### OVERVIEW OF THE SUBJECT AREA

Студент

Группа	ФИО	Подпись	Дата
8BM5B	Н.В. Бондарчук		

Консультант кафедры ПИ

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент каф. ПИ	Е. С. Чердынцев	К. Т. Н.		

Консультант – лингвист кафедры ИЯИК

Должность	ФИО	Ученая степень, звание	Подпись	Дата
зав. каф. ИЯ	Т. В. Сидоренко	К. П. Н.		

## Overview of the subject area

### 1.1 Choice of ranking factors affecting the promotion of web resources in search engines

Before creating a system for analyzing the factors of the ranking of web resources, it is necessary to identify which parameters affect the promotion in search engines, such as Yandex and Google. These parameters will be used for analysis.

On the production practice, based on the specialized literature [1-3] on the promotion of sites in the search engines, an analysis of the ranking factors was carried out, which have a greater impact on the search distribution. As a result, ranking factors were tested and those that work more effectively when promoting sites in search engines are highlighted (Figure 1.1).

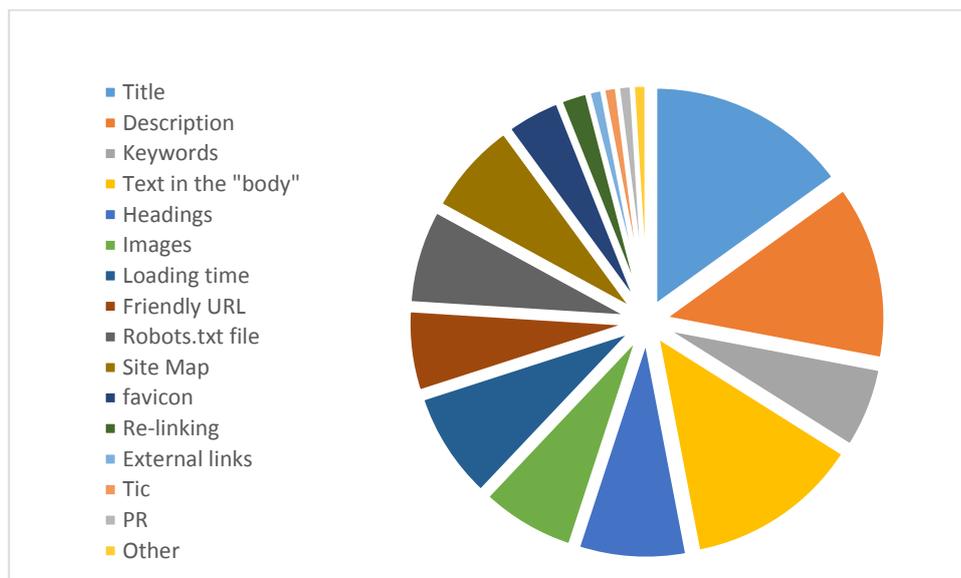


Figure 1.1 - Influence of ranking factors

The main principle of the search engines is to determine the relevance of the content of the page of the site containing, therefore, the factors that influence it were chosen. In addition, other important ranking factors will be considered.

Keywords - a represented word (or a set of words) that describe a site theme or text, and are used to search for exact relevance (relevance) of search results. When search engines index content, they distribute them to databases with different keywords. In other words, instead of storing 25 billion pages in one

database, search engines use millions of smaller databases, each of which contains sites with a keyword. This allows search engines to find relevant pages and produce search results in a fraction of a second [5].

Keywords play a primary role in the Internet search. When a user enters a query, the search engines select pages with the corresponding keywords, the principle of the search engines is shown in Figure 1.2.

The user requests the information he needs, then the request is sent to the cache of the search engine, if there is a response, it is given to the user, if not, the request is sent to the main database, from which the answer is raised and given to the user.

The order of words in the query, punctuation, the use of additional characters - all this helps search engines to find the most relevant answers. To determine the relevance of the page, search engines analyze how the keywords are used in the text, and other components of the page. To optimize a page for a particular keyword, you must use it in the metatags.

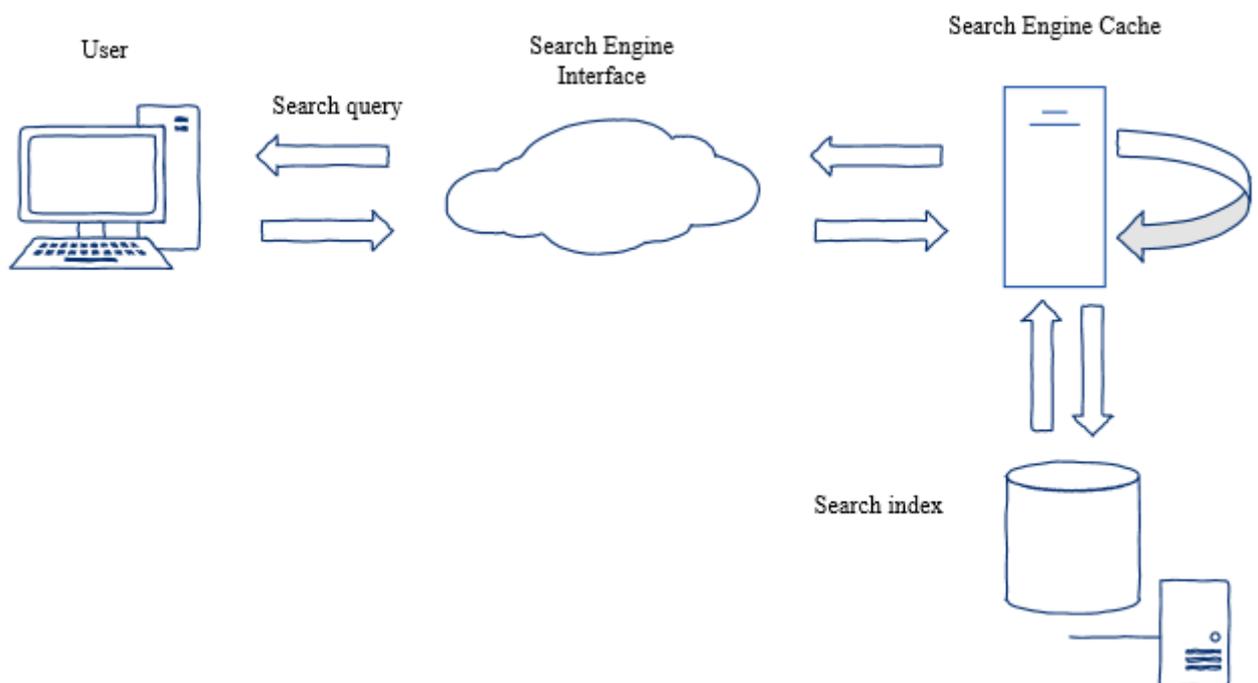


Figure 1.2 - The principle of the search engines

### **1.1.1 Metatags: Title, Description, Keywords**

Title - the most important meta-tag for internal optimization of HTML-code pages. Search engines attach great importance to the text contained in the title, so you should use this factor as effectively as possible.

Example: <title> Page Title </ title>

Physically, the content of the title tag is the name of the page that appears in the topmost field of the browser. Also, the content of the title is displayed in the output of search engines at the request of users (this is the name of the links in the search results). Therefore, the title of the page should reflect its essence as much as possible and necessarily contain key words. For each page, the title content must be unique [6].

Search engines can take into account not all words contained in the title of the page, but only a certain number of the first words (the rest can simply be cut off). For different search engines, these metrics are different, and besides, they can simply change over time as the algorithm is adjusted, so it is strongly recommended that you have keywords and phrases at the beginning of the header.

The Description Metatag is intended to create a short description of the page, which is displayed as an annotation to the link to the page in the SERP. However, at present this rule works mainly in the Google search engine, and that is not always the case. Quite often, search engines as annotations show just a fragment of the text of the page most relevant to the query [7]. However, Google officially recommends using the Description metatag to describe the pages.

Information that determines whether the contents of the Description tag affects the relevance of the page relative to the keywords that it contains are not propagated in any search engine. It can be assumed that this influence is present. In any case, the presence of keywords in the Description is desirable for the above reason.

Keywords is another metatag that plays an important role in SEO. Example of the metatag Keywords in the html document:

<Meta name = "keywords" content = "Keywords, keywords ...">

Meta-tag Keywords search engines are used to determine the relevance, or compliance, of a link. To properly use this metatag, you must use only those words that are contained in the document itself. The use of words that are not on the page is not recommended. A list of relevant keywords for the Keywords metatag can be separated by commas or you can use a space [8]. These words should reflect the most important topics of your site.

The Keywords Metatag is becoming less and less important from the perspective of the search engine. In fact, you can get high positions by not including anything in the Keywords metatag. However, in order to secure itself, it is desirable to include some keywords in this metatag.

### **1.1.2 Text on the site page**

In the ranking mechanism, the most important is to evaluate the relevance of the page content to the query that the user entered. Initially, the relevance of the pages was determined by internal criteria: the density of keywords in the text, the occurrence of keywords in the title, metatags, text elements, and many others. However, after the appearance of doorways optimized for these requirements, it became necessary to develop external parameters that determine the site's compliance with the query [9]. As an example, describing this process, we can consider the following formula (Г.1):

$$R=PR*(T+L) \quad (\Gamma.1)$$

where: R is the relevance value; T - degree of correspondence of internal criteria to specified search systems requirements (text relevance); L - level of reference ranking - the degree to which the texts of incoming links to the document match the search query (referential relevance); PR - indicator of external criteria of the document, not depending on the request (authority of the resource).

This formula is not an exact representation of the work of modern algorithms for ranking search engines, but it gives an idea of the significance and types of criteria from which the relevance of the pages is formed. [10]

To compile index databases and determine relevance, search engines use content, query text, and some elements of html-markup. The text should be interesting and useful to the user, and the number of occurrences and placement of keywords should meet the requirements of the search engine.

Text optimization - improving text content from the perspective of search engines. Actually, when they say about site optimization, they mean the optimization of the texts placed on its pages. All other components, for example, usability or the correctness of the page from the position of HTML-code for search engines, although important, but are secondary. The first thing that gets into the field of view of the search robot and what it works with is textual content.

### **1.1.3 Images on the site page**

Almost every site has pictures, photos, some graphic images or even image galleries. Like texts, images require optimization and organization in addition to other elements of the page. As a result, placement and search optimization of images cannot be made in the rendering of Google in the SERP, but also to the traffic in the rendering of Google and Yandex [10] image services.

### **1.1.4 Headings**

The main two purposes of headings in HTML tags H1-H6 - design and structuring of content on the pages of the resource in order to make it more readable and easily consumed by visitors to the site. It is by the headers of Google and Yandex that determine how the topic of your page coincides with the requests that introduce users. In other words, how relevant this page is. [11]

Ideally, the page should have only one heading of the first level, which as it contains two or three chapters of the second level in which in turn "nested" headings of less rank and so on. The structure of the document should look something like this (Figure 1.3):

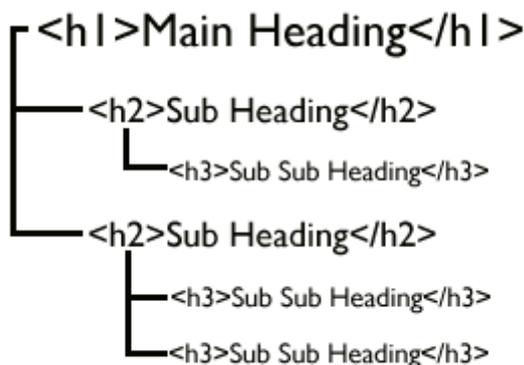


Figure 1.3- Hierarchy of headers

That is, the headers should be, as it were, nested to each other in importance and have a tree-like appearance. The text itself, it is desirable to break into paragraphs using the <p> tag - where the text of each paragraph.

Such text markup is easily "digested" by search engines and other programs working with HTML documents, for example, allows you to snatch one of the paragraphs to describe the site in search results without breaking the head where the logical part of the text begins or ends, or taking into account the header levels automatically compose The content of the document.

Since the text in the headers <h1> - <h6> is much more important for search engines, and simply for the person viewing the page, rather than just the text, they should be filled with keywords.

### 1.1.5 Favicon

Favicon (English Favorite Icon) is an icon in a special format that is displayed in the address bar when you visit a site, for example, this, the browser tab or its bookmarks ("Favorites"). Usually it is stored in the root folder of the site and is called favicon.ico. [12].

The use of this icon has many positive sides. First, this icon is remembered by visitors at times better than the name of the site. Secondly, favicon makes your site unique, highlighting your site among other sites that did not have time to acquire an icon.

Some search engines display this favicon for each site in search results. Thus, this is one of the levers for controlling the clickability of your snippet. The image always attracts the user's attention and, even if your site is not in the first positions in the search results, more visitors will go to you than to the site next to yours without favicon.

### **1.1.6 The Robots.txt file**

The robots.txt file provides important information for search robots that scan the Internet. Before you go through the pages of your site, the search robots check this file. This will allow them to more efficiently crawl the site, since you help the robots immediately start indexing the really important information on your site (provided that you have correctly configured robots.txt). But, both the robots.txt directives and the noindex instruction in the robots metatag are just a recommendation for robots, so they do not guarantee that the closed pages will not be indexed and will not be added to the index. If you need to really close part of the site from indexing, then, for example, you can additionally use the passwords to close the directories. [13]

### **1.1.7 Site map**

Site map - a tool for user navigation on the site. A site map is a complete catalog of resource sections. Site map is placed, as a rule, on the main page of the site so that the user in a short time and without difficulty moved to the desired section of the site. [14]

The site map is necessary for realization of two purposes. Firstly, the site, equipped with a simple and informative map, is popular among users. On the Internet, a very important parameter of the resource is the level of user confidence. The site on which the map is placed, without a doubt, causes such trust. Secondly, then the optimizer can effectively use, presence on a site of a card is highly cited by search engines. Convenient user navigation is, so to speak, one more point in favor of successful website promotion in search engines.

### **1.1.8 Page loading time**

Search engines pay attention to the time of loading a page of the site for one rather simple reason: the slower sites get into the top of the delivery of a particular search engine, the less users will use it. [15]

Accelerating the work of the site leads to an increase in the number of page views the visitor of the resource during one session. Almost eighty percent of users claim that if the page loads for more than four seconds, the likelihood that they will return to this resource is negligible. [16] The average duration of the page loading in the opinion of the Internet community should not exceed two seconds.

It turns out that not to concede to the competitive search system, you need to display on the first page of the issuance only those results that have the optimal time of loading the site and do not upset the visitors with the need to wait long.

### **1.1.9 Friendly URLs**

A friendly URL is the unique address of a web page in the global network, briefly displaying its contents, understandable and friendly. CNC are one of the tools of SEO and positively influence the ranking of pages in the SERP. This is due to the fact that, as a rule, human-to-human URLs contain the keywords of the query, to which the search engines are positive. For example, by typing in Google

a "learning logo for a site", one of the first results you'll see is a link to a page with a friendly URL-key.

A well-organized structure of web addresses with human-understandable URLs greatly helps people to find the pages they are interested in, and more often it is to information from such sites that users leave links. While long URLs with a little sympathetic content scare visitor away. [17]

It is worth noting that it is necessary to make links human-understandable in the first days of the creation of the site. Subsequent modification of the appearance of links is fraught with the removal of pages from the search for a certain period of time - after all, search engines will need to re-index the changed links, delete old ones, and write new ones.

## **1.2 Determination of reference characteristics of parameters**

For each ranking factor, there are recommendations from the search engines. For Yandex, they are located on the Yandex.Help service [19], for the search engine Google - Search Console [20]. Consider all the methods and choose the main recommendations that will be used in the developed system for analyzing web resources.

Recommendations for compiling the title:

1. The title of each page should be unique.
2. Place the key phrase closer to the beginning of the title.
3. If there are several keywords on which the page is advancing, try to combine them into title without compromising readability.
4. You cannot use a simple listing of keywords in the title - this can be the reason for imposing sanctions on search engines.
5. Do not repeat the same word more than 2 times. But the use of "key" twice (in different word forms) is even desirable.

6. It is recommended not to include in the title stop characters / \ | () + =. And if it is necessary, it is better to use them after the key phrase, and not to it.

7. Try to do without stop words (prepositions, pronouns, interjections, introductory verbal constructions, etc.).

8. The length of the title should be a maximum of 80 characters with spaces, but it is better to adhere to the optimal boundaries (40-65 characters).

9. A dot at the end of the title is not set.

10. Do not use a question mark in the title, even if the key phrase requires it, otherwise users can decide that you are asking a question, but do not give an answer.

11. Do not write the headline in capital letters - it can be regarded as spam.

Rules and recommendations for compiling a description:

1. Like the title, description of each page should be unique.

2. Description is written specifically for each Internet article, and not copied from the text of the page.

3. The optimal amount of the web document description is 130-180 characters with spaces. The maximum is 200, although sometimes they occur in the issuance of descriptions and 240 characters each.

4. If there are several queries for the web page, use the highest-frequency ones in the description.

5. The key phrase is recommended to be used 2 times: in exact occurrence and in morphological or diluted.

6. Enter in the description synonyms of words from the key phrase. Look at Yandex, which words it highlights along with those that enter the query. For example, "buy" - "sale", "buy" - "order", etc.

Recommendations for the compilation of keywords:

1. The number of words in the keywords tag is not more than 10.

2. For the correct use of this metatag, only those words that are contained in the document should be used. The use of words that are not on the page is not recommended.

3. The list of keywords should be separated by commas or spaces.

4. Keywords should not be repeated more than 3 times, including different spelling variants.

Directions for texts:

1. Uniqueness of the text is more than 90% The uniqueness of the text is a characteristic by means of which it is possible to determine whether there are partial or complete copies of the given document on the Internet. The less in the article fragments that occur on other resources, the higher the percentage of its uniqueness. [18]

2. Text should be about 2000 characters with spaces

4. Text should contain at least one bulleted or numbered list

5. In the text there should be headings 2-3 on the section: there should be key phrases with exact entry once and diluted entry one time

6. The text should contain two exact occurrences of key and one diluted entry

7. In the text it is necessary to use word forms and root words of keywords

8. Keyword density should not exceed 6%.

Consider the requirements for images:

1. Title and description - a text description and a title allow you to ensure that search engines understand what this picture is about.

2. Size - search engines take care of the load time, so you need to make sure that the image is the appropriate size.

3. The file name is a simple and important detail. The name of your image should be directly related to it and include keywords.

4. Formatting - as a rule, it's better to upload a JPG format photo to the site. This will help search engines find and index them correctly. If the use of JPG format is not possible, adhere to such formats as GIF or PNG.

5. Link - you can refer to your image to an article. Just as you use links to an article, you can do the same by spreading a link to an image.

The headings h1-h6 should be compiled according to the following rules:

1. On the site page there should be only one h1 header and is at the top of the page.

2. In the headings, you must use keywords. The higher the importance of the keyword, the higher the header level.

3. Headings should be informative and not very long at the same time.

4. Point out important keyword phrases as close as possible to the beginning of the title.

5. Try to make the key phrases in sub-headings taking into account morphology and synonyms.

When building a robots.txt file, you should follow some tips:

1. The file name must be written only in capital letters (for example, the names Robots.txt or ROBOTS.TXT will be incorrect).

2. For robots.txt there is a restriction on the weight of the file from Google, this is up to 500 KB, from Yandex to 32 KB.

3. The robots.txt file should be located in the root folder of the site.

Recommendations for creating a site map:

1. The site map should show the hierarchy of the entire site. Try to break the map into convenient sections, points, try to use lists and headings.

2. The site map file must be in the root of the site.

3. The site map should not contain links that are closed for indexing in the robots.txt file.

4. The site map specifies the priorities for indexing the pages of the site from 0.1 for the most unimportant page to 1 for the main page.

For a human-friendly URL, the basic rules are to use keywords, be readable to the site visitor and the crawler.

There are no special recommendations for composing favicon, its main availability and size is 16 by 16 pixels.

The download speed of the site should be 2 seconds, because if the site will be loaded for a long time, visitors will close the site without waiting for a full download.

Having studied the recommendations, we will define the reference values of the ranking factors that we will use for analysis.

One of the important components of the site's page are text parameters (Table 1.1).

Table 1.1 Text parameters

	<b>Number of characters</b>	<b>Entering the query</b>
<b>Title</b>	40-80	1
<b>Description</b>	130-180	1-2
<b>Keywords</b>	to 10 words	2-3
<b>Text</b>	2000-3000	3-4
<b>Paragraph</b>	500-1000	1-2

Text on the site page, in addition to the above characteristics, has a number of important characteristics, such as the presence of bulleted and numbered lists, the presence of headers (H1, H2, H3, etc.), the length of paragraphs (500-1000

characters). Headers should also include a key query, where the H1 header must be one on the page and contain the main query.

On the page of the site there should be 2-5 images, which should have a key query in the title (for example, if the query is "stretch ceilings", then the image can be called natyajnie-potolki.jpg). In the header and in the alternative text of the image, key requests must also be written.

The next factor in the ranking of the site was the time of loading the site. The site should not load more than 2 seconds. As the load time increases, the chance increases that the user will leave the page before it loads. That is, the site will lose a potential customer, which means that it will be inefficient.

For correct indexing by search robots on the site server, there must be a robots.txt file, which indicates the main site mirror, and information about which pages are open for indexing and which are closed. If this file is missing or if it is not filled correctly, the search robots may incorrectly select the site mirror and index all pages of the site. Thus, the search results can be administrative and private pages of the site.

Also for faster indexing of new pages of a site by search robots the site map is necessary, it should not contain pages of a site closed in a file robots.txt. Each page should have a priority of 0.1 to 1.

To increase the CTR (clickability) of the site you need favicon. It allows you to improve the recognition of the site, and therefore increase customer loyalty. Dimensions favicon 16x16 pixels.

The next item is the CNC (human-to-human URL). For search engines to understand what the page of the site in them should be key requests.

### **1.3 Overview of similar systems**

There are a lot of tools for analyzing the promotion of sites in the search engines, they provide a variety of services and programs. These systems can be divided into several groups:

1. Analysis of the text: advego.ru, text.ru, pr- cy.ru/zypfa/
2. Analysis of the site's positions: allpositions.ru/, seogadget.ru/serppos, line.pr- cy.ru/
3. Analysis of the reference mass: linkpad.ru, en.ahrefs.com, rookee.ru
4. Analysis of internal parameters, without recommendations for optimization: seolib.ru, SiteAuditor, analizsaita.com
5. Analysis of internal and external parameters, without recommendations for optimization: xtool.ru, sbup.com
6. Analysis of internal parameters, with recommendations for optimization: WebSite Auditor, iAnalysis

The system being developed belongs to the 6th group, that is, it can analyze the internal parameters and provide information on the necessary amendments to the site.

Let us consider in more detail the systems of interest to us:

- WebSite Auditor allows you to analyze a large number of different ranking factors, such as metatags, images, text, robots.txt. The analysis of the pages takes place taking into account specific keywords. This service does not analyze site maps, download speed, availability of CNC and favicon. The program is paid - the cost of a license in Russia is about \$ 124.

- iAnalysis service is a service analyzing PR and TIC, indexing the site in search engines, links to the site, availability in major catalogs, page size, IP and site URL. The analysis of pages occurs without taking into account the keywords. This service does not analyze the main parameters that affect the promotion of the site in the search engines, but it allows you to save the analysis results and is free.

In the process of studying WebSite Auditor and iAnalysis systems, it was revealed that the analysis is saved to the device on which the program is open, which makes it impossible to open a report from the program on another device. This complicates the process of work, since in many companies several people

work on optimization of sites. For ease of sharing and use on different devices, it is necessary that reports are saved in the user account.

Thus, the weaknesses of similar systems were singled out, which will be taken into account when developing their own system. As a result, the developed system should provide recommendations on all the main factors of the site ranking, be free, provide the ability to create a user account, save reports in the system and compare reports.