

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Южно-Уральский государственный университет»
(национальный исследовательский университет)
Факультет лингвистики
Кафедра лингвистики и межкультурной коммуникации

РАБОТА ПРОВЕРЕНА
Рецензент, доцент

_____ /Н.С. Скрипичникова/

ДОПУСТИТЬ К ЗАЩИТЕ
Заведующий кафедрой,
д.филол.н., доцент

_____ /Т.Н. Хомутова/

**ЛИНГВИСТИЧЕСКИЙ АНАЛИЗ ДОКУМЕНТА-ОБРАЗЦА,
ОРИЕНТИРОВАННЫЙ НА SEO-ОПТИМИЗАЦИЮ ТЕКСТА**

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
ЮУрГУ – 450501.2017.912.ВКР

Руководитель, к.филол.н., доцент
_____ /О.И. Бабина/
« ____ » _____ 2017 г.

Автор
студент группы ЛМ-533
_____ /В.С. Андреева/
« ____ » _____ 2017 г.

Нормоконтролер,
к.филол.н., доцент
_____ /О.И. Бабина/
« ____ » _____ 2017 г.

Работа защищена с оценкой

« ____ » _____ 2017 г.

Челябинск
2017

ОГЛАВЛЕНИЕ

Введение.....	3
Глава 1 SEO-оптимизация текста и лингвистический аспект.....	6
1.1 Поиск информации и ПС.....	6
1.1.1 Современные ИПС.....	7
1.1.2 Запрос к поисковой системе.....	9
1.1.3 Релевантность. Пертинентность. Ранжирование.....	11
1.2 SEO-оптимизация.....	14
1.2.1 SEO-копирайтинг.....	15
1.2.2 Создание оптимизированных текстов.....	18
1.2.3 SEO-анализ.....	20
1.2.4 Оптимизация текста под ПС.....	22
Выводы по главе 1.....	24
Глава 2 Лингвистический анализ документа-образца, ориентированный на оптимизацию текста.....	25
2.1 Лингвистический анализ текстов и SEO-оптимизация.....	25
2.2 Рекомендации по оптимизации текстов под ИПС.....	49
Выводы по главе 2.....	51
Заключение.....	54
Библиографический список.....	56
Список литературы, использованной для анализа.....	60

ВВЕДЕНИЕ

В наше время информационные технологии продолжают стремительно развиваться, требуя совершенствования работающих систем. Среди таких систем можно выделить информационно-поисковые системы, которые делают возможным быстрый поиск текстов на просторах Интернета. Поскольку поток информации постоянно увеличивается, то возникает проблема ее обработки: большая часть веб-сайтов не входит в базу данных, в силу неинформативности контента. Систему научили по определенным параметрам отбирать ценный материал для базы данных.

Создателям контента необходимо придерживаться определенных правил при создании текстов, чтобы система не просто внесла информацию в свою базу данных, но и смогла выдать в первых рядах при запросе. Так появилось направление SEO-копирайтинг, которое занимается созданием текстов с учетом требований поисковых систем. В данном направлении используется SEO-анализ текстов. Такой анализ направлен на выявление лингвистических компонентов текста, которые оказывают влияние на отбор поисковыми системами. SEO-анализ, осуществляемый специальными программами, дает возможность увидеть насколько текст является «приемлемым». Однако стоит отметить, что информации о том, как следует писать оптимизированные тексты мало, так большинство рекомендаций, которые можно найти в книгах по продвижению сайтов, не имеют четкой теории и неясно, как работают на практике.

Существует вполне конкретная проблема – стремительное совершенствование поисковых систем и их механизма по отбору информации. SEO-анализ также необходимо совершенствовать, чтобы успешно создавать тексты, которые не будут противоречить политике поисковых систем. Предполагается, что расширение лингвистического анализа текстов для SEO-оптимизации позволит более точно определять недостатки и недоработки, а также даст возможность их исправить.

Актуальность исследования объясняется необходимостью совершенствовать SEO-оптимизацию текстов. В настоящее время лингвистический аспект оптимизации не имеет соответствующего строго научного обоснования ни в теории, ни в практике. Данное направление может представлять научный интерес для лингвистики, так как связано с текстом и его компонентами на разных уровнях.

Объект нашего исследования – единицы текстов на лексическом, грамматическом и стилистическом уровнях.

Предметом исследования является лингвистический анализ на трех уровнях (лексика, грамматика, стиль), ориентированный на SEO-оптимизацию текстов.

Материалом исследования послужили тексты, полученные по запросам «Альтернативные источники энергии» к поисковым системам Google и Яндекс.

Цель работы – провести лингвистический анализ текстов на компоненты, которые должны пройти SEO-оптимизацию.

Для достижения поставленной цели необходимо было решить следующий ряд **задач**:

- 1) Дать определение понятиям «SEO-анализ» и «SEO-оптимизация».
- 2) Провести стандартный SEO-анализ текстов, полученных при запросе к поисковым системам Google и Яндекс.
- 3) Осуществить SEO-оптимизацию текстов в виде трансформаций на трех уровнях и составить рекомендации.

Методы исследования выбраны с учетом цели и задач исследования: сравнительно-сопоставительный, описательный, трансформационный методы и метод семантического анализа.

Научная новизна дипломной работы заключается в том, что предпринимается попытка усовершенствовать SEO-оптимизацию текстов с помощью лингвистических приемов.

Теоретическая значимость – исследование поможет расширить понятие SEO-оптимизации текстов в лингвистическом аспекте.

Практическая значимость – результаты исследования могут войти в пособия по лексикологии и стилистике.

Дипломная работа содержит: введение, две главы, заключение и библиографический список.

Во введении представлена информация об актуальности, объекте, предмете, цели и задачах исследования, а также новизна исследования, теоретическая значимость и практическая.

В первой главе раскрываются понятия SEO-анализа и оптимизации текста, описываются явления, связанные с ними, и выявляется проблематика исследуемого предмета.

Во второй главе анализируются тексты, полученные в результате запроса к наиболее популярным поисковым системам Google и Яндекс, выявляется соответствие полученных текстов требованиям к ним, рассматриваются лингвистические компоненты, которые необходимо изменить (провести трансформации) в качестве SEO-оптимизации.

В заключении подводится итог результатов исследования.

ГЛАВА 1 SEO-ОПТИМИЗАЦИЯ ТЕКСТА И ЛИНГВИСТИЧЕСКИЙ АСПЕКТ

1.1 Поиск информации и ПС

В современном мире работа с информацией большей частью связана с использованием информационных технологий (ИТ), которые в свою очередь облегчают данный процесс для человека. Так основные операции, связанные с информацией (получение, хранение, обработка и передача), в автоматизированном режиме выполняются значительно быстрее. И поскольку растет информационная потребность, то и возможности, позволяющие ее удовлетворить, должны расширяться. Сейчас любой человек имеет доступ к глобальной информационной среде Интернет, которая позволяет в любой момент и за короткий срок найти все необходимое. Однако любой вид ИТ требует, чтобы механизмы работы с информацией постоянно совершенствовались, а для этого надо найти решения текущих проблем и недостатков систем.

Развитие систем обеспечения выполнения операций с информацией затрагивает самые разные направления, в том числе из разных областей. Лингвистика и ее методы играют немалую роль в развитии ИТ, ведь существует необходимость решить проблемы обработки и поиска информации на естественном языке.

Поиск информации (Information Retrieval) – действия, методы и процедуры, позволяющие осуществить отбор определенной информации из массива данных [Беляева 2007, с. 72].

Наиболее распространенным средством для нахождения необходимой информации в Интернете являются информационно-поисковые системы (ИПС или ПС). Такие системы представляют собой программно-аппаратный комплекс, который дает возможность поиска разного рода информации, в том числе и на естественном языке.

Использование ИПС в среде Интернет позволяет быстро найти необходимый информационный ресурс, так как поисковые системы

упорядочивают их по степени информационной потребности пользователя, а также исключают источники с некачественным содержанием.

1.1.1 Современные ИПС

Каждая информационно-поисковая система имеет свою структуру, которая принципиально отличается от других. Нередко механизмы, которые совершенствуют работу системы, держат в тайне. Все же у всех ИПС есть некоторые общие элементы. Так практически каждая поисковая система состоит из трех основ:

1) Поисковый робот («веб-паук» или «краулер») – это программа, отвечающая за обработку страниц в Интернете и внесение информации о них в базу данных поисковика.

2) Индекс системы – это база данных на поисковом сервере. Поиск запрошенной пользователем информации осуществляется в пределах такой базы данных, которая пополняется поисковым роботом (содержит ссылки и сжатые копии веб-страниц).

3) Программа для обработки запроса пользователя, нахождения в индексе документов (соответствующих запросу) и выдачи результатов поиска.

Для того чтобы увидеть разницу в работе поисковых систем необходимо рассмотреть их классификацию. Оценка работы ИПС может базироваться на разных показателях и проводиться согласно разным направлениям, однако в данном случае уместно привести классификацию, связанную с основной функцией системы (поиск по запросу).

Таким образом, поисковые системы можно разделить на следующие виды [17, с. 34-35]:

1) Направленность поиска:

- Глобальные поисковые системы – такие системы индексируют (передают в базу данных) информационный материал по всему миру и на всех языках. Например, к таким относятся: Google, Baidu, Yahoo!, Bing.

- Региональные поисковые системы – это поисковые системы, которые специализируются на региональных (национальных) информационных ресурсах. Такие поисковые системы индексируют информационный материал выборочно на одном или нескольких языках, а также ориентированы на поиск сайтов в пределах конкретной страны (например, Яндекс и Mail.ru).
- Локальные поисковые системы – это поисковые системы, которые специализируются на локальных информационных ресурсах на одном языке и в пределах одной или нескольких локальных сетей. Такие поисковые системы содержат предусмотренные фильтры, которые дают возможность пользователям блокировать нежелательные сайты. Например: Halalgoogling.

2) Способ хранения информации:

- Поисковые каталоги. В ИПС информация хранится в виде тем или категорий. К поисковым каталогам относится, например, известный сайт *wikipedia.org*, в котором можно найти справочную информацию.
- Рейтинговые системы. Поисковые системы, которые организуют выдачу результатов по количеству просмотров посетителей.
- Поисковые указатели. Такие поисковые системы основываются на поиске информационного материала по ключевым словам. Результаты поиска зависят от набора слов, заданных в запросе. Данный вид представляет большую часть поисковых систем.

В мировом пространстве наиболее популярной поисковой системой является Google, тогда как в России чаще используется Яндекс. В зависимости от того, какую информацию пользователь хочет найти, он делает выбор в пользу той или иной поисковой системы. Выбор напрямую зависит от результатов, которые пользователь получает при запросе.

1.1.2 Запрос к поисковой системе

Разница в работе современных информационно-поисковых систем видна по результатам любого поискового запроса. Результаты зависят от того, что искать и как формулировать запрос.

Поисковый запрос – исходная информация для осуществления поиска с помощью поисковой системы [Калитин 2011, с. 30]. Запрос зависит как от устройства поисковой системы, так и от типа информации для поиска. В основном поисковый запрос задается в виде набора слов или фразы, иногда используются расширенные возможности языка запросов поисковой системы.

Любой запрос пользователя можно связать с определенной целью, которая в свою очередь отражает желание получить в результате необходимую на данный момент информацию. При этом запрос может прямо или косвенно выражать цель пользователя.

Согласно наиболее распространенной классификации, все поисковые запросы делятся на [2, с. 66-68]:

1) Навигационные запросы – запросы, которые отражают цель пользователя в нахождении информации или сайта, связанного с конкретным известным лицом или компанией. При этом считается, что результат в виде ссылки на официальный сайт будет в полной мере соответствовать ожидаемому результату. Так, например, на запросы «администрация челябинска» или «бмв» поисковая система (в данном случае Google) выдаст среди первых результатов ссылки на официальные сайты *cheladmin.ru* и *bmw.ru*.

2) Информационные запросы – запросы, которые осуществляет пользователь с целью найти какие-либо конкретные сведения или информацию, однако источник при этом не принципиально важен. К информационным запросам можно отнести, например, «реформы Петра I» или номер телефона определенного медицинского центра. В первом случае результатов в виде научных статей или онлайн книги по истории будет

достаточно. Во втором – сайт со справочным материалом (необязательно сайт самого медицинского центра).

3) Транзакционные запросы – запросы, которые делает пользователь с целью совершить какое-либо действие (или транзакцию) в Интернете. К таким запросам можно отнести следующие примеры: «книгу купить онлайн», «доставка еды на дом», «путевки в Египет дешево» и т.д.

4) Нечеткие запросы – запросы, которые не отражают цели пользователя. К данному виду запросов относятся все запросы, представляющие слова и выражения, не уточняющие ожидаемый результат пользователя. Например, «билеты» или «большая квартира» представляют собой именно нечеткие запросы, поскольку остается неясна цель пользователя (речь может идти о заказе билетов онлайн, также не совсем понятно на какой транспорт должны быть билеты; большая квартира может быть связана с покупкой, а может и с ремонтом).

Существуют и другие классификации, которые рассматривают запросы не только через призму цели пользователя. Например, с учетом частотности (частотность – количество запросов, которые пользователи совершают за какой-либо определенный период) запросы делятся на: высокочастотные, среднечастотные, низкочастотные и редкие запросы. Или, например, запросы, которые определяются «геозависимостью»: геозависимые (привязка к региону) и геонезависимые (не зависят от месторасположения).

Запросы могут быть связаны с определенным временем года, тогда их можно разделить на: сезонные запросы – такие запросы, которые наблюдаются в какое-то определенное время года; и всесезонные – запросы, которые не зависят от времени года.

Нетрудно заметить, что каждая классификация отражает отдельный фактор, влияющий на результаты поиска информации.

Современные поисковые системы используют механизм ориентировки на запрос пользователя: система обрабатывает фразу или набор слов и выдает максимально соответствующие им результаты.

1.1.3 Релевантность. Пертинентность. Ранжирование

Говоря о поиске информации, невозможно не уделить внимание трем основным понятиям, которые характеризуют результаты процесса.

Первым таким понятием является релевантность.

Релевантность (от англ. relevant – уместный, относящийся к делу) – соответствие запроса пользователя результатам выдачи поиска [Терехов 2010, с. 7].

Релевантность страниц сайта является важным компонентом поисковой выдачи, поскольку задача состоит в том, чтобы отделить полезный и качественный информационный материал от материала, в котором нет информативности и уникальности по теме запроса.

Поисковая выдача формируется главным образом в соответствии с релевантностью. Большинство поисковых систем сначала находят наиболее релевантные страницы для каждого случая (запроса), а затем распределяют их при выдаче. Определение релевантности – это целый алгоритм, и у каждой поисковой системы он свой. К общим критериям такого алгоритма (для любой поисковой системы) можно отнести [2, с. 240]:

1) Внутренние критерии:

- Частота (или плотность) ключевых слов (слов, соответствующих тематике запроса и текста на веб-сайте) в тексте.
- Расположение ключевых слов в тексте.
- Степень идентичности поискового запроса фрагменту текста на веб-странице.
- Словоформы и синонимы ключевых слов в тексте.
- Обновление веб-страниц (требуется регулярное обновление или пополнение информационного материала).
- Общий объем информации (например, количество страниц или разделов).

2) Внешние критерии:

- Популярность ресурса.

- Количество обращений к веб-странице по запросу.
- Посещаемость ресурса.
- Срок существования сайта.

Другим основным понятием является пертинентность. Данное понятие представляет собой субъективную оценку (в отличие от релевантности) пользователем соответствия качества полученной информации сделанному запросу.

Пертинентность (от англ. *pertinent* – соответствующий) – соответствие результатов поиска информационной потребности пользователя [Терехов 2010, с. 7]. Пользователь при получении результатов поиска оценивает информацию, которая может соответствовать или не соответствовать его информационной потребности.

Важно помнить, что оба понятия характеризуют не только информационный результат, но и работу, проделанную ИПС. Поисковым системам приходится не только хранить расширяющуюся индексную базу, но и постоянно обновлять ее, поскольку информация должна соответствовать тому, что появляется в Интернете. Необходимо как индексировать новые тексты, так и обновлять ранее зафиксированную информацию.

Одной из важных и масштабных проблем, с которыми сталкивается любая поисковая система, является проблема удовлетворение информационных потребностей пользователя. Трудность заключается в понимании, что ожидает пользователь получить в результатах поисковой выдачи. Если поисковые системы смогут «понять» намерения пользователя, то проще будет сформировать максимально подходящую для данного запроса поисковую выдачу, которая удовлетворит запрос пользователя. Когда пользователь увидит «понимание» со стороны поисковой системы, то сможет использовать ее постоянно, зная, что получит точный ответ на любые запросы. Решить проблемы такого «понимания» ИПС крайне

сложно, однако оценка результата поиска с точки зрения релевантности и пертинентности все же необходима.

Если проблему релевантности удастся решить, то благодаря возможному объективному восприятию данного понятия. Как уже было сказано, пертинентность можно воспринимать только субъективно, поэтому достижение точного совпадения по результатам поиска пока не представляется возможным. К сожалению, поисковые системы не могут настраиваться на информационные нужды одного пользователя.

Ранжирование – последнее понятие из трех основных.

Ранжирование – процесс выбора страниц из базы поисковой системы, соответствующих запросу пользователя, и упорядочение их по степени убывания соответствию запроса или релевантности [Ашманов 2007, с. 258].

Из определения следует, что ранжирование тесно связано с понятием релевантности.

Существует ряд параметров текста, которые влияют на ранжирование [2, с. 208]:

1) Тематика и содержание. Информация должна быть ценной и понятной пользователю. Система не включает в результат информационные ресурсы, которые содержат несвязный текст или просто набор слов.

2) Наличие и плотность ключевых слов. Текст, в котором будут найдены ключевые слова, будет релевантен поисковому запросу пользователя. При этом учитывается частота повторов ключевых слов и место их расположения в тексте. Слишком низкая или слишком высокая плотность приведет к тому, что поисковая система не выдаст ссылку на текст в результатах.

3) Уникальность. Тексты должны отличаться уникальностью, то есть они должны быть специально созданными для каждого отдельного сайта.

4) Структурированность. Текст должен быть разбитым на части, содержать подзаголовки, таблицы или списки.

Постепенно мы переходим от принципов поиска информации к текстовому содержанию. Современный подход к представлению информации в среде Интернет представляет собой понимание работы поисковой системы и умение «конструировать» материал таким образом, что сайт, где будет размещен материал, окажется релевантным поиску по теме материала.

Текстовый контент (содержимое) представляет собой отдельный элемент, который необходимо исследовать более детально.

1.2 SEO-оптимизация

В поисках информации пользователь обращается в информационную среду Интернет. Здесь существуют свои механизмы поиска и структурирования данных. В такой среде меняются принципы функционирования информации, и меняется естественный язык. Когда пользователь ищет электронный вариант обычного текста, он использует приемы поиска, которые существенно отличаются от тех, что существуют в обычной жизни. Информационная среда постепенно формирует новое направление, которое изучает специфический язык данной среды, а также новые способы представления привычной информации, новые формы «общения».

«Общение» между человеком и информационной средой основывается на законах, которые не похожи на те, что действуют в отношении общения человека с человеком. Пользователь при работе с поисковыми системами вначале прибегал к использованию естественного языка, однако теперь он «выстраивает общение», применяя при запросе короткие фразы или ключевые слова, которые система сможет распознать. Естественный язык с трудом воспринимается системами Яндекс и Google, поэтому он заменен языком ключевых слов. И такой язык отличается от естественного (например, русского языка) и лексикой, и правилами

согласования компонентов. Язык общения с системой является компромиссом между естественным языком и языком поисковых систем.

Компромисс в таком случае приводит к появлению правил формирования текстов и представление текстовой информации. Принципиальное отношение пользователя к контенту приводит к созданию поискового продвижения или оптимизации информации под язык поисковых запросов. С помощью анализа поисковых запросов пользователей комплектуется «полезный контент» для информационных ресурсов. Оптимизированные тексты создаются в соответствии с поисковыми запросами. Язык компромисса постепенно распространяется в среде Интернет, а SEO-оптимизация текстов становится масштабным явлением.

Оптимизация текста – это приведение его в такое соответствие требованиям поисковых систем, при котором сайт, содержащий оптимизированный текст, будет высоко ранжироваться поисковыми роботами [Крохина 2012, с. 156].

Оптимизированные тексты имеют несколько преимуществ. Во-первых, такие тексты пишут согласно определенным «правилам», поэтому они будут составлены аккуратно и продуманно. Во-вторых, такие тексты должны облегчать выбор материала для базы данных. В-третьих, создание оптимизированных текстов может повысить уровень уникальности контента в сети Интернет.

1.2.1 SEO-копирайтинг

Создание оптимизированных текстов относится к SEO-копирайтингу. Развитие данного направления связывают именно с необходимостью решения проблемы соотношения поисковых запросов и текстового контента веб-страниц. Однако SEO-копирайтинг занимается и проблемой написания оригинальных и интересных текстов в Интернете [2].

Контент (от англ. content – содержание, содержимое) – любого вида информация, которая содержится на страницах веб-ресурса [16, с. 87].

К контенту можно отнести тексты, аудио- и видеофайлы, таблицы, картинки и т.д.

Согласно способу представления, контент может быть: авторским (представлен на своем сайте самим владельцем) и пользовательским (посетители добавляют свой контент на какой-либо сайт). Согласно содержанию различают: тематический контент (в случае, если текст соответствует теме сайта) и нетематический (в случае, если текст на тему из другой области).

Контент является одним из главных факторов, влияющих на ранжирование в поисковой выдаче. Контент должен соответствовать определенным требованиям [16]:

1) Уникальность. Текст должен быть оригинальным и неповторимым, поскольку веб-страницы с неуникальным контентом могут не попасть в базу данных поисковых систем.

2) Оптимизирование под поисковые запросы. Ключевые слова в текстах дают возможность поисковым роботам точно определять тематику и направленность сайта, чтобы впоследствии выводить их как наиболее релевантные результаты соответствующим запросам пользователей.

3) Информативность и актуальность. Качественные интересные тексты, аудио- и видеоматериалы способствуют привлечению пользователей, которые будут обращаться к информационному ресурсу вновь.

4) Легкость восприятия. Отдельные элементы контента должны быть логично и удобно распределены в тексте.

SEO-копирайтинг (от англ. search engine optimization copywriting) – это создание уникальных и «активных» текстов для сайтов на основе семантического ядра [2, с. 282].

Уникальность и «активность» текстов достаточно просто определить: текст не должен быть копией текста другого автора; текст должен привлекать читателей информативностью или побуждением к чему-либо.

Уникальность контента является обязательным условием эффективного продвижения. При индексации проверяется содержание сайта на оригинальность, если будут обнаружены веб-страницы с контентом-копией, то информационный ресурс не появится в результатах выдачи поисковых систем.

Согласно способам создания уникальных текстов направления можно разделить на следующие группы: легальные (селфрайтинг, копирайтинг), полулегальные (рерайтинг, перевод с иностранных языков) и нелегальные (сканрайтинг, размножение статей).

Внимание стоит уделить и такому аспекту как семантическое ядро.

Семантическое ядро (СЯ) – это набор слов, их морфологических форм и словосочетаний, которые наиболее часто встречаются в тексте или при запросах [27, с. 4].

Поисковые запросы для составления семантического ядра подбираются проведением анализа тематики или направления контента веб-страницы, статистики сайта и статистики поисковой системы. Ключевые слова, которые определяются в семантическое ядро, должны относиться к тематике контента.

Семантическое ядро составляется следующим образом: производится подбор с помощью специализированных сервисов (например, Яндекс.Директ или Google.AdWords) возможных ключевых слов и словосочетаний, отбрасываются лишние слова, затем в соответствии с классификацией запросов определяются оставшиеся ключевые слова и словосочетания, которые распределяются по тексту [27].

Процесс подготовки семантического ядра заключается в выборе запросов, по которым текст будет попадать в результаты выдачи поисковой системы.

Существуют разные методики по составлению СЯ. Однако у них есть некоторые общие положения такие, как:

- 1) Составление первичного списка слов и словосочетаний.

На начальном этапе составления СЯ необходимо составить первичный список, который будет содержать основные слова (связанные с конкретной темой). Для составления первичного списка ключевых фраз можно использовать следующее [2, с. 92]:

- Сокращения, варианты ошибочного написания, сленг, использование перевода, синонимов или транслитерации. Например, Mercedes – мерседес.
- Варианты названия составных частей товара или услуг. Например, оргтехника – факс, копир, телефон.
- Слова качества, привлекательности или надежности товаров или услуг. Например, недорогой дизайн, бесшумный пылесос.

2) Расширение списка с использованием ассоциированных запросов.

Собрать все ассоциированные запросы помогут поисковые системы. Благодаря ассоциированным запросам можно расширить первичный список ключевых фраз. Для расширения списка с помощью ассоциированных запросов, потребуется одно или два главных слова.

3) Обработка и получение списка поисковых запросов.

Из полученного списка запросов следует убрать наименее подходящие по теме запросы. В семантическое ядро должны входить только те запросы, по которым можно найти контент по теме.

1.2.2 Создание оптимизированных текстов

При создании уникального текста для сайта необходимо помнить следующее правило [27, с. 19]: необходимо работать не просто над текстом, а ориентироваться на удовлетворение информационной потребности пользователя, при этом соблюдая «золотую середину» между маркетинговыми задачами текста и техническими требованиями к нему.

Поскольку мы уже определили требования к текстовому контенту при размещении его в Интернете, то сейчас рассмотрим лингвистический аспект создания текстов.

Известно, что поисковые системы находят разницу между естественным текстом, написанным для людей, и текстом, который сформирован только для соответствия требованиям ПС. Крайне важно учитывать следующие несколько пунктов, чтобы текст был естественным [27]:

1) Структура.

Данный параметр подразумевает логику повествования. Текст следует писать так, чтобы он был понятен пользователям. Также структура должна соответствовать теме (обычно выражается в поисковых запросах), задачи (удовлетворение информационной потребности пользователя) и условию (пространство размещения текста) коммуникации.

Модель структуры текста выглядит следующим образом: заголовок, введение, основная часть и заключение.

2) Абзац и предложения.

При составлении предложений следует учитывать: расположение поисковых запросов, их согласование с другими словами в предложении, количество слов в предложении и типы предложений.

В пределах предложения поисковый запрос развивает его основную мысль, поэтому занимает сильные позиции – начало или конец предложения. Также стоит помнить о том, что поисковый запрос выступает в роли члена предложения, поэтому должен быть согласован с другими словами (слова употребляют в общепринятых значениях и соблюдают их лексическую сочетаемость). Слов в предложении должно быть ровно столько, чтобы выражалась одна главная мысль. При этом тип предложения также играет свою роль – простые и нераспространенные предложения легче воспринимаются читателем, чем сложные (с двумя и более грамматическими основами), распространенные (с большим количеством второстепенных членов), осложненные (рядами однородных членов, причастными и деепричастными оборотами, вводными и вставными конструкциями).

В начале каждого абзаца сообщается наиболее важная информация. Стоит исключать повторы одной мысли, ведь каждый абзац развивает

не более одной микротемы. Объемы абзацев в одном тексте должны быть примерно одинаковы.

3) Грамотность.

Одной из главных задач при создании оптимизированных текстов является его грамотное написание. Грамотность не только вызывает доверие пользователей, но и повышает авторитетность для поисковых систем.

4) Лексика.

Одинаковые слова (поискового запроса) следует равномерно распределять по тексту, чтобы избежать множественных повторов. Также необходимо использовать синонимы и местоимения.

Создание оптимизированного текста подразумевает, что учитываться будут все параметры. Только так можно достичь высокого качества текста, который будет положительно воспринят как пользователями, так и поисковыми системами.

1.2.3 SEO-анализ

Как было сказано ранее, для поисковых систем важнейшим критерием является уникальность текста. Поэтому не стоит использовать тексты чужого авторства в оригинальном виде, иначе контент не будет хорошо ранжироваться поисковыми системами. SEO-оптимизация текста учитывает следующие основные факторы [16]:

- Уникальность (90% и более).
- Объем текста (от 300 слов).
- Плотность ключевых слов и заспамленность.

Основой SEO-оптимизации принято считать SEO-анализ.

Семантический анализ текста (или SEO-анализ) – это оценка качества текста с точки зрения его соответствия целям оптимизации [Ашманов 2007, с. 93]. Семантический анализ призван показать соотношение общего количества слов в тексте и значимых слов, составляющих семантическое ядро.

Такой анализ является именно показателем качества текста, который составляется и размещается на страницах сайта с целью его продвижения. Написание текста для оптимизации требует наличия в нем ряда показателей, которые не должны превышать соответствующие нормы. Другими словами, анализ дает возможность понять, какой у текста уровень заспамленности, какие слова являются подавляющими и т.д.

Таким образом, SEO-копирайтер может сделать определенные выводы и по результатам семантического анализа произвести необходимую корректировку текста, сделав его более привлекательным и для поисковой системы, и для человека.

Основными показателями семантического анализа являются [33]:

- Заспамленность (иногда «тошнота») ключевых слов – частота употребления в тексте слов, составляющих семантическое ядро. Запросы, по которым продвигается страница, по частоте приведения в тексте должны преобладать над всеми остальными словами.

- Стоп-слова – количество в тексте слов, не несущих смысловой нагрузки (предлоги, союзы, местоимения и др.). При индексации такие слова не учитываются поисковиками, что имеет в виду при вычислении процента плотности в тексте ключевых слов.

Семантический анализ проводится с помощью специальных программ, которые сами учитывают все необходимые показатели. Среди таких программ можно выделить Advego и Textus Pro. Такие программы имеют установки определения разных показателей, которые влияют на качество текста. Качество текста учитывается не только для того, чтобы контент попал в базу данных, но и для того, чтобы текст был интересен пользователю.

1) Advego.

Данная программа анализирует следующее:

- Количество символов в тексте (с пробелами и без пробелов).
- Общее количество слов в тексте (с предлогами).
- Количество уникальных слов, встречающихся только один раз в тексте.

- Количество значимых для текста слов.
- Количество грамматических ошибок.
- Заспамленность текста.
- Плотность ключевых слов в тексте.

Стоит отметить, что процесс анализа текста данным сервисом не всегда понятен. В некоторых случаях полученные проценты показателей могут быть спорными. Некоторые предпочитают доверять своему профессионализму и проверяют качество текстов самостоятельно.

2) Textus Pro.

Программа анализирует следующие показатели:

- Количество знаков (с пробелами и без).
- Количество слов в тексте (со стоп-словами и без).
- Заспамленность текста (норма – 4-5, без стоп-слов – 3,5).
- Повторы всех слов.
- Плотность распределения слов в тексте (норма 2,5% для одного слова).

В практической части мы проведем SEO-анализ с помощью программы Textus Pro. Результаты данного анализа станут основой для оптимизации текстов.

1.2.4 Оптимизация текста под ПС

Оптимизированные тексты необязательно должны создаваться с нуля. Практически любой текст можно оптимизировать таким образом, что поисковые системы смогут воспринимать его, как если бы он создавался специально для публикации в сети Интернет.

Мы уже говорили о том, как создаются оптимизированные тексты, а также о том, что необходимо при этом учитывать. Для оптимизации уже готового текста стоит опираться на три основных лингвистических уровня:

- лексический,
- грамматический,

- стилистический.

Все три уровня помогут обеспечить успешную оптимизацию под поисковые системы, поскольку на каждом из данных уровней будут учтены особенности текста, влияющие на восприятие содержимого. Оптимизация таких особенностей подразумевает их преобразование для лучшего восприятия пользователя и для повышения качества (показатели, которые учитываются при семантическом анализе). Здесь уместно говорить о лингвистических трансформациях, которые дадут возможность сделать текст «приемлемым».

Лингвистические трансформации – это такие преобразования слов, словосочетаний, конструкций и предложений, которые не искажают главного смысла текста.

Именно трансформации наиболее точно помогут повысить качество текста на разных его уровнях за счет внесения изменений без ущерба для общей целостности текста и его главной идеи. Так оптимизированный текст нуждается в осторожной «модификации».

Три уровня текста включают следующие виды трансформаций:

1) На лексическом уровне применяются следующие приемы трансформаций: замена, опущение, генерализация, конкретизация и добавление.

2) К грамматическим трансформациям относятся такие приемы, как перестановка, замена, опущение, добавление и перестройка предложения.

3) Стилистические трансформации включают: различные замены, компенсации, описания и синонимические замены.

Данные лингвистические приемы помогут структурно работать над оптимизацией, вносить только необходимые изменения в текст, помогут избежать переписывания, и тем самым избежать лишней траты времени.

На практике нам предстоит рассмотреть, как данные приемы трансформаций повлияют на основные показатели качества текста при SEO-анализе.

Выводы по главе 1

К содержанию текстового контента необходимо относиться серьезно, поскольку соблюдение правил создания и оптимизации дают возможность поисковым системам отбирать интересные для пользователя информационные ресурсы.

Направление SEO-копирайтинг занимается созданием и оптимизацией текстов с учетом требований поисковых систем. В данном направлении используется SEO-анализ текстов, который выявляет показатели качества текста. Среди таких показателей выделяют следующие: общее количество слов со всеми служебными словами (и количество слов без них), уровень заспамленности текста, плотность распределения ключевых слов и стоп-слов. Для семантического анализа текста используют специальные программы (например, Advego или Textus Pro).

Так SEO-анализ является основой не только для создания уникальных текстов, но и для оптимизации уже готовых текстов, которые не соответствуют требованиям поисковых систем.

Чтобы сделать оптимизацию текстов под поисковые системы необходимо применять лингвистические трансформации на трех основных уровнях текста, которые позволят достичь главной цели – повысить качество для пользователей и поисковых систем. При этом не будет нарушена целостность и смысл текста.

Поскольку информации о том, как следует проводить оптимизацию текстов недостаточно, то в следующей главе мы проверим на практике подход к решению проблемы удачной оптимизации путем использования приемов лингвистических трансформаций.

ГЛАВА 2 ЛИНГВИСТИЧЕСКИЙ АНАЛИЗ ДОКУМЕНТА-ОБРАЗЦА, ОРИЕНТИРОВАННЫЙ НА ОПТИМИЗАЦИЮ ТЕКСТА

В предыдущей главе мы в теории обосновали понятия «SEO-оптимизация», рассмотрели современные требования к оптимизированным текстам и определили лингвистические уровни, на которых может быть осуществлена оптимизация текста. В данной главе мы переходим к практической части исследования. Нам предстоит с помощью программы TEXTUS PRO провести стандартный SEO-анализ текстов (одной тематики), полученных в результате запросов к поисковым системам Google и Яндекс. Мы проведем лингвистический анализ всех текстов на трех основных уровнях и выделим особенности, которые следует оптимизировать. Затем на основе лексических, грамматических и стилистических особенностей будут обобщены рекомендации по работе с оптимизацией текстов.

2.1 Лингвистический анализ текстов и оптимизация

Практическая часть исследования начинается с определения тематики текстов, которые предстоит анализировать. Тема – альтернативные источники энергии. При этом запрос к поисковым системам будет совпадать с нашей темой – «Альтернативные источники энергии».

Мы выбрали наиболее популярные в нашей стране поисковые системы: Google и Яндекс. Среди полученных по запросам результатов в системах мы отбираем текстовый контент по трем первым ссылкам на первой и третьей страницах. Так мы сможем увидеть некоторую разницу в качестве начального варианта текста и разницу после проведения оптимизации.

В результате поиска в системе Google были получены следующие статьи: «Альтернативная энергетика», «Альтернативные источники энергии», «8 необычных источников альтернативной энергии для дома, офиса

и отдыха», «Есть ли будущее у альтернативных источников энергии?», «Альтернативная энергия – будущее человечества» и «Альтернативные источники энергии».

Статьи, которые были найдены по ссылкам в системе Яндекс: «Альтернативная энергетика», «Альтернативные источники энергии», «8 необычных источников альтернативной энергии для дома, офиса и отдыха», «Альтернативные источники энергии», «Альтернативные источники энергии» и «10 источников энергии будущего».

Каждый из данных текстов статей мы анализируем отдельно. Однако первый и третий тексты с первой страницы в Google совпадают с первым и третьим текстами с первой странице в Яндекс, поэтому два идентичных текста будут считаться за один (потому как взяты с одного и того же информационного ресурса). SEO-анализ будет проведен с помощью программы TEXTUS PRO.

Итак, анализ текстов, полученных в результате поиска в системе Google:

1) Статья «Альтернативная энергетика».

Источником данной статьи является известный информационный ресурс: https://ru.wikipedia.org/wiki/Альтернативная_энергетика.

Вероятно, что популярность ресурса, который предоставляет справочную информацию по разным темам, влияет на выдачу, поскольку и в Google, и в Яндекс ссылка на статью оказалась первой.

Тема статьи несколько шире, чем было обозначено в запросе. В ней дано определение понятию альтернативной энергетике, а также рассмотрены направления, которые относятся к данной области. Информация об альтернативных источниках энергии, в данном случае, приведена в качестве пункта в разделе «Направления альтернативной энергетике», и с учетом того, что здесь есть краткое описание и классификации, стоит считать статью информативной, хотя и не совсем точно отвечающей нашему запросу.

Согласно SEO-анализу:

- Общее количество слов составляет 1041 (без учета стоп-слов – 817).
- Уровень заспамленности текста равен 6,71 (без учета стоп-слов – 5,66).

Отсюда следует, что на уровень заспамленности в большей мере должны повлиять лексические трансформации.

- Среди часто встречаемых слов следует выделить следующие: «энергия» с разными морфологическими формами (например, «энергии» с 32 повторами и 3,07% плотности) повторяется 43 раза и имеет плотность 4,13%; «источник» повторяется 21 раз при плотности 2,02% (форма «источников» имеет 11 повторов и 1,06% плотности); «электростанция» – 12 повторов и 1,15% плотности; «электроэнергия» – 12 повторов и 1,15% плотности.

- Частота стоп-слов, которая требует оптимизации: «в» (45 повторов при 4,32% плотности), «и» (32 повтора при 3,07% плотности).

В тексте наблюдается повышенный уровень заспамленности за счет повторов ключевых слов и стоп-слов.

Рассмотрим данный текст на трех уровнях, где можно провести трансформации, которые улучшат оптимизацию.

Лексический уровень.

1. Основным направлением альтернативной энергетики является поиск и использование **альтернативных (нетрадиционных) источников энергии**.

Основным направлением альтернативной энергетики является поиск и использование **нетрадиционных способов получения ресурса**.

Чтобы избежать повторов ключевых слов «альтернативный», «источник» и «энергия» можно опустить первое слово (в предложении уже используется его синоним), а в отношении двух других – применить конкретизацию и генерализацию.

2. Альтернативный источник энергии является возобновляемым ресурсом, он заменяет собой традиционные **источники энергии**, функционирующие на нефти, добываемом природном газе и угле, которые при сгорании

выделяют в атмосферу углекислый газ, способствующий росту парникового эффекта и глобальному потеплению.

Альтернативный источник энергии является возобновляемым ресурсом, он заменяет собой традиционные **источники**, функционирующие на нефти, добываемом природном газе и угле, которые при сгорании выделяют в атмосферу углекислый газ, способствующий росту парникового эффекта и глобальному потеплению.

Слово «энергии» здесь можно опустить, поскольку смысл после опущения будет также понятен.

3. Грозная энергетика – это способ **использования энергии путём** поимки и перенаправления энергии молний в электросеть.

Грозная энергетика – это **способ** поимки и перенаправления энергии молний в электросеть.

Опущение словосочетания «использования энергии путём» избавит предложение от «лишних» слов.

Грамматический уровень.

1. Предполагалось, что **эта энергия окажется** значительно дешевле **энергии**, полученной с помощью современных источников, окупаться такая установка будет за 4-7 лет.

Предполагалось, что **так будет** значительно дешевле **использования** современных источников, окупаться такая установка будет за 4-7 лет.

После перестройки предложения с причастным оборотом, который был опущен, стало возможным устранить повторяющееся слово «энергия».

2. Синтез более тяжёлых атомных ядер из более лёгких с целью получения энергии, который носит управляемый характер. **До сих пор не применяется.**

Синтез более тяжёлых атомных ядер из более лёгких с целью получения энергии, который носит управляемый характер, **до сих пор не применяется.**

Объединение двух предложений в одно обусловлено необходимостью правильно донести смысл. Начальный вариант является не совсем корректным.

3. Используется **как для** нагрева воды **для отопления, так и для** производства электроэнергии.

Используется для нагрева воды (**отопления**) и производства электроэнергии.

Конструкция «как для..., так и для...» часто встречается в тексте. В данном предложении конструкция приводит к повторению стоп-слова «для», поэтому применимо опущение.

Стилистический уровень.

1. Приливные электростанции (ПЭС) **пока имеются лишь** в нескольких странах – Франции, Великобритании, Канаде, России, Индии, Китае.

Приливные электростанции (ПЭС) **есть только** в нескольких странах – Франции, Великобритании, Канаде, России, Индии, Китае.

Словосочетание требует замены, поскольку нарушает стилистическое «звучание» текста.

В результате применения трансформаций на всех трех уровнях (использование опущений, замен, синонимов и перестроек предложений) наблюдались существенные изменения в показателях SEO-анализа:

- Общий уровень заспамленности – 5,39 (5,20 без стоп-слов).
- Частотность ключевых слов – «энергия» (27 повторов, 2,75% плотности), «источник» (16 повторов, 1,63% плотности).
- Частотность стоп-слов – «в» (29 повторов, 2,96% плотности), «и» (26 повторов, 2,65% плотности).

2) Статья «Альтернативные источники энергии».

Источник статьи: <http://elektrik.info/main/news/614-alternativnye-istochniki-energii.html>.

Сайт «Электрик Инфо» предлагает статьи, обучающие материалы, обзоры и информацию для начинающих электриков и домашних мастеров.

Данная статья написана доступным языком без использования специализированных терминов. Тема статьи напрямую отражает цель запроса. В ней рассмотрены некоторые виды альтернативных источников энергии и дано краткое описание их преимуществ и недостатков.

SEO-анализ текста показал, что:

- Общее количество слов – 506 (без учета стоп-слов – 378).
- Уровень заспамленности текста с учетом стоп-слов равен 5,48 (5,48 без них).
- В тексте часто встречаются следующие слова: «энергия» с морфологическими формами «энергии» и «энергию» повторяется 30 раз с плотностью в 5,93%, «источник» («источники», «источников» и «источника») повторяется 19 раз при 3,75% плотности.
- Частота стоп-слов в тексте: «в» с 16 повторами и 3,16% плотности, «и» с 11 повторами и 2,17% плотности.

В данном случае уровень заспамленности связан с ключевыми словами, поэтому им будет уделено больше внимания при использовании трансформаций.

Лексический уровень.

1. На вопрос «что делать в преддверии энергетического кризиса?» уже давно найден ответ: надо искать другие источники **энергии** – альтернативные, нетрадиционные, возобновляемые.

На вопрос «что делать в преддверии энергетического кризиса?» уже давно найден ответ: надо искать другие источники **добычи ресурса** – альтернативные, нетрадиционные, возобновляемые.

Приемы конкретизации и генерализации помогают избежать лексического повтора слова «энергии» в тексте.

2. Все мы привыкли использовать в качестве источников **энергии** органическое топливо – уголь, газ, нефть.

Все мы привыкли использовать в качестве **ее** источников органическое топливо – уголь, газ, нефть.

Чтобы избежать постоянного повтора слова «энергия» в тексте, существительное заменяется местоимением.

3. Всевозможные гелиоустановки используют солнечное излучение как альтернативный источник **энергии**.

Всевозможные гелиоустановки используют солнечное излучение как **альтернативный источник**.

В предложении возможно опущение слова «энергии», потому как из контекста уже понятно, о чем идет речь.

Грамматический уровень.

1. Кроме того, ветрогенераторы, **производя энергию**, не загрязняют атмосферу вредными выбросами.

Кроме того, ветрогенераторы **не загрязняют атмосферу вредными выбросами**.

В данном случае прием опущения конструкции (деепричастный оборот) приемлем, поскольку оборот не представляет важной для понимания информации.

2. **Энергию** этих геотермальных источников **и предлагают использовать** в качестве **альтернативного источника сторонники геотермальной энергетики**.

Сторонники геотермальной энергетики предлагают использовать геотермальные источники в качестве **альтернативы**.

Приемы перестановки, опущения и замены связаны с необходимостью исключения повторяющегося в тексте слова (в данном случае синоним не подходит).

3. Поэтому **для** отработанную воду необходимо закачивать обратно в подземный водоносный горизонт.

Поэтому отработанную воду необходимо закачивать обратно в подземный водоносный горизонт.

Использование приема опущения связано с ошибкой использования слова в предложении.

Стилистический уровень.

1. Как видим, альтернатива традиционным **источникам энергии** – существует.

Как видим, альтернатива традиционным **способам добычи ценного ресурса** – существует.

Во избежание использования фразы «источники энергии», применяется трансформация в виде замены, а затем и добавления слова «ценный», чтобы замена не нарушала атмосферу статьи.

Результат оптимизации текста с помощью трансформаций (чаще лексические опущения и замены):

- Заспамленность текста составляет 4,00 (по сравнению с 5,48).
- Частотность ключевых слов: «энергии» (16 повторов, плотность 3,27%), «источник» (16 повторов, плотность 3,27%).
- Частотность стоп-слов в тексте: «в» (16 повторов и 3,27% плотности), «и» (10 повторов и 2,04% плотности).

3) Статья «8 необычных источников альтернативной энергии для дома, офиса и отдыха».

Информационным источником данной статьи является: <http://www.novate.ru/blogs/280415/31040/>. Данный сайт представляет собой электронный журнал, в котором можно найти статьи по разным темам.

Данная статья рассчитана на широкий круг читателей. Целью статьи является составление интересной информации о современных альтернативных источниках энергии, которые могут быть доступны всем.

SEO-анализ выявил, что в данном тексте:

- Общее количество слов – 1033 (без учета стоп-слов составляет 712).
- Уровень заспамленности текста - 6,78 (4,12% без стоп-слов).

Уровень заспамленности существенно снижается, когда исключаются стоп-слова, поэтому оптимизация будет направлена именно на них.

- В тексте статьи часто встречаются слова: «энергия», «энергией», «энергии» и «энергию» – 17 повторов и 1,65% плотности; «power» – 9 повторов и 0,87% плотности.

- Среди частых стоп-слов встречаются: «в» с 46 повторами и 4,45% плотности, «на» – 21 повтор и 2,03% плотности.

Лексический уровень.

1. Солнечные панели **в окнах**.

Солнечные панели **для окон**.

Трансформационный прием замены помогает изменить предлог, чтобы плотность «в» стала ниже.

Грамматический уровень.

1. Компания Uprise создала необычную ветряную турбину высокой мощности, которую можно использовать **как в быту, так и в промышленных масштабах**.

Компания Uprise создала необычную ветряную турбину высокой мощности, которую можно использовать **в быту и промышленности**.

Применяется трансформация в виде перестройки конструкции «как в..., так и в...», поскольку есть необходимость исключить повтор предлога «в» в тексте.

2. **В сложенном состоянии с турбиной Uprise можно ездить по дорогам общего пользования**.

Турбину Uprise можно сложить, так она не мешает на дорогах общего пользования.

Перестройка предложения и добавление конструкции с союзом «так», потому как первый вариант «в сложенном состоянии» можно отнести к человеку (не только к турбине).

3. Полученная энергия **в данном случае** передается по шнуру, **соединяющем** воздушного змея с базовой станцией.

Полученная энергия передается по шнуру, **соединяющему** воздушного змея с базовой станцией.

«В данном случае» можно опустить без ущерба для информативности предложения. Исправлено нарушение норм согласования в «по шнуру, соединяющем», возможно, это просто опечатка.

Стилистический уровень.

1. Но с недавних пор стало возможным размещать **эти элементы** прямо в окнах, что позволяет использовать **такие** батареи даже владельцам обычных квартир в многоэтажных домах.

Но с недавних пор стало возможным размещать **такие панели** прямо на окнах, что позволяет использовать **батареи** даже владельцам обычных квартир в многоэтажных домах.

Фраза «эти элементы» стоит особняком в данном предложении, поэтому заменяется на более уместную «такие панели». Тогда повторное слово «такие (батареи)» опускается.

2. **Эти элементы** пропускают 99 процентов проходящего через них света, но имеют при этом коэффициент полезного действия в 7%.

Они пропускают 99 процентов проходящего через них света, но имеют при этом коэффициент полезного действия в 7%.

Фразе «эти элементы» можно заменить местоимением, чтобы не допустить повтора слова «элементы».

В результате оптимизации на лексическом, грамматическом и стилистическом уровнях было отмечено следующее:

- Общий уровень заспамленности удалось снизить до 5,48 (4,12% без стоп-слов).

- Часто встречающиеся слова не превышают допустимой нормы, поэтому остаются с прежними показателями: «энергия», «энергией», «энергии» и «энергию» – 17 повторов и 1,65% плотности; «power» – 9 повторов и 0,87% плотности.

- Была необходимость изменить частотность стоп-слов: «в» (30 повторов и 3,02% плотности), «на» (19 повторов и 1,91% плотности).

4) Статья «Есть ли будущее у альтернативных источников энергии?».

Источником статьи послужил новостной ресурс – <https://pronedra.ru/alternative/2017/03/22/buduschee-alternativnoy-energetiki/>.

Статьи относятся к разным рубрикам (например, политика, энергетика и т.д.).

Тема статьи связана с применением альтернативных источников энергии в России, в ней также приведено сравнение с положением в других странах, много статистики и цифр.

Согласно SEO-анализу в тексте статьи:

- Общее количество слов насчитывает 1647, если не учитывать стоп-слова количество слов составит 1186.

- Уровень заспамленности текста равен 8,83% (4,69% без стоп-слов).

- Часто встречаемые слова в тексте: «Россия» и «российский» повторяются 25 раз при 1,52% плотности, «альтернативный» – 22 повтора и 1,34% плотности, «источник» – 22 повтора и 1,34% плотности.

- Частота стоп-слов: «в» (78 повторов при 4,74% плотности), «и» (39 повторов при 2,37% плотности), «на» (36 повторов и 2,19% плотности).

Высокий уровень повторов стоп-слов «в» и «и» в тексте, необходимо применить лингвистические приемы в виде трансформаций.

Лексический уровень.

1. Альтернативные источники энергии представляют интерес, в первую очередь, из-за выгоды их использования – ресурсы ВИЭ **в основном** неисчерпаемы – при достаточно низком риске причинения вреда окружающей природе.

Альтернативные источники энергии представляют интерес, в первую очередь, из-за выгоды их использования – ресурсы ВИЭ **преимущественно** неисчерпаемы – при достаточно низком риске причинения вреда окружающей природе.

Применение трансформации в виде замены фразы «в основном» связано с целью исключения использования предлога «в».

2. Однако вопрос перехода на альтернативные источники в России стоит рассматривать не **в административной, а в экономической плоскости.**

Однако вопрос перехода на альтернативные источники в России стоит рассматривать не **с административной, а с экономической стороны.**

Лексическая замена связана с заменой предлогов (по причине высокой частотности данных стоп-слов).

Грамматический уровень.

1. **В РФ существуют целевые показатели для развития ВИЭ:** до 2020 года **в стране** должны быть построены солнечные электростанции с суммарной мощностью 1,6 ГВт.

Целевые показатели для развития ВИЭ страны выглядят следующим образом: до 2020 года должны быть построены солнечные электростанции с суммарной мощностью 1,6 ГВт.

Перестановка и добавление в начале предложения дают возможность не использовать повторяющиеся предлоги и слова.

2. Геотермальные станции могут быть и мощнее, примером тому является самая крупная в мире ГеоЭС **в Кении** с мощностью 140 МВт.

Геотермальные станции могут быть и мощнее, примером тому является самая крупная в мире ГеоЭС (**Кения**) с мощностью 140 МВт.

Исключить повторяющийся предлог «в» (у которого высокая плотность в тексте) можно, заключив в скобки слово в качестве пояснения.

3. **В случае с АЭС остается,** пусть и минимальная, но угроза, что в окружающую среду могут попасть радионуклиды.

АЭС представляет, пусть и минимальную, но угрозу, что в окружающую среду могут попасть радионуклиды.

Пассивный залог перестраиваем в активный, чтобы не оставлять «лишних» слов, влияющих на уровень заспамленности.

Результаты оптимизации текста:

- Уровень заспамленности текста удалось снизить до 7,07%.

- Часто встречаемые слова в тексте: «Россия» с 23 повторами при 1,43% плотности.

- Частота стоп-слов: «в» – 50 повторов при 3,12% плотности, «и» – 34 повтора при 2,12% плотности, «на» – 34 повтора при 2,12% плотности.

5) Статья «Альтернативная энергия – будущее человечества».

Источник статьи: <http://madenergy.ru/stati/alternativnye-istochniki-ehnergii-kak-bolshoj-shag-v-razvitie-i-ispolzovanie-ehlektroehnergii.html>.

Сайт MadEnergy.RU является информационным, научно-популярным и профессиональным Интернет-ресурсом. Здесь представлены новости, пресс-релизы и обзоры прессы, которые постоянно обновляются.

Статья кратко, но подробно рассказывает о том, что такое альтернативные источники энергии, какую роль они играют для общества и какие основные виды источников существуют.

SEO-анализ текста данной статьи:

- Общее количество слов – 712, количество слов без учета стоп-слов – 540.

- Уровень заспамленности текста составляет 5,00 (без стоп-слов – 4,69).

- В тексте частотность ключевых слов выглядит следующим образом: «энергия» 26 раз повторяется при 3,65% плотности, «альтернативный» – 18 повторов и 2,53% плотности, «источник» – 16 повторов и 2,25% плотности.

- Частотность стоп-слов: «в» (25 повторов, 3,51% плотности), «на» (12 повторов, 1,69% плотности).

Лексический уровень.

1. При этом одним из направлений развития данной отрасли является использование любого вида энергетики, который представляет интерес с экономической точки зрения, в силу низкой стоимости за единицу получаемой энергии и с экологической точки зрения, поскольку альтернативные виды **энергии**, как правило, отличаются своей безопасностью и не наносят вред окружающей среде.

При этом одним из направлений развития данной отрасли является использование любого вида энергетики, который представляет интерес с экономической точки зрения, в силу низкой стоимости за единицу получаемой энергии и с экологической точки зрения, поскольку ее альтернативные виды, как правило, отличаются своей безопасностью и не наносят вред окружающей среде.

Повтор слова «энергия» удастся избежать с помощью замены на местоимение.

Грамматический уровень.

1. В настоящее время для получения необходимого **количество** электрической энергии применяются **не возобновляемые** источники, которые представляют собой природные ископаемые, такие как уголь, нефть или природный газ.

В настоящее время для получения необходимого **количества** электрической энергии применяются **невозобновляемые** источники, которые представляют собой природные ископаемые, такие как уголь, нефть или природный газ.

Исправлена ошибка в согласовании и словообразовании в данном предложении.

2. Альтернативная **энергетика** – это новая отрасль **энергетики**, представляющая собой общность перспективных направлений, ставящие своей задачей поиск новых способов получения, передачи и применения энергии, **источником** которой являются альтернативные **источники энергетики**.

Альтернативная энергетика – это новая отрасль, представляющая собой общность перспективных направлений, ставящие своей задачей поиск новых **альтернативных** способов получения, передачи и применения энергии.

В пределах данного предложения были осуществлены: перестановка слова и опущение определительного придаточного предложения.

3. При этом одним из направлений **развития** данной отрасли является использование любого вида энергетики, который представляет интерес с экономической **точки зрения, в силу** низкой стоимости **за единицу получаемой энергии** и с экологической точки зрения, **поскольку** альтернативные виды энергии, **как правило, отличаются своей безопасностью** и не наносят вред окружающей среде.

При этом одним из направлений **развития** данной отрасли является использование любого вида энергетики, который представляет интерес с экономической (**низкая стоимость**) и с экологической точки зрения (**альтернативные виды энергии не наносят вред окружающей среде**).

Есть возможность упростить данное предложение за счет опущения повторяющихся словосочетаний, вводных фраз и конструкций.

В результате оптимизации:

- Уровень заспамленности текста составил 4,71 (без стоп-слов – 4,32).
- Частотность ключевых слов: «энергия» – 21 повтор при 3,00% плотности, «альтернативный» – 15 повторов и 2,01% плотности, «источник» – 14 повторов и 2,00% плотности.
- Частотность стоп-слов: «в» (20 повторов, 2,87% плотности), «на» (11 повторов, 1,55% плотности).

б) Статья «Альтернативные источники энергии».

Статья взята с сайта <http://www.mywindenergy.ru>. Данный сайт принадлежит компании «ОМТех». Компания занимается продажей ветрогенераторов и солнечных батарей. На сайте можно найти не только информацию о продаваемой продукции, но и статьи, посвященные тематике альтернативных источников энергии.

Данная статья является релевантной запросу. В ней дана информация о том, что такое альтернативные источники энергии и какие существуют виды в настоящий момент. Цель статьи – привлечь покупателей.

Согласно SEO-анализу в данном тексте:

- Общее количество слов – 862; количество слов без учета стоп-слов – 580.
- Уровень заспамленности текста равен 6,16 (без учета стоп-слов – 5,83).
- Часто встречаемые слова: «энергия», у которых 45 повторов и 5,22% плотности, «ветра» – 11 повторов и 1,28% плотности.
- Частота стоп-слов: «в» (38 повторов, 4,41% плотности), «и» (28 повторов, 3,25% плотности), «на» (16 повторов, 1,86% плотности).

Лексический уровень.

1. Однако, выход есть и он уже очевиден – это переход на альтернативные источники **энергии**.

Однако выход есть и он уже очевиден – это переход на альтернативные источники **ресурса**.

Генерализация слова «энергии».

2. И не только те, что не имеют своих источников **энергии и** вынуждены постоянно закупать **ресурсы** у других стран, но и те, кто думает о будущем и понимает, что ресурсы не безграничны и сохранив хоть часть их сегодня дадут ему возможность на существование потом.

И не только те, что не имеют своих **источников, вынуждены** постоянно закупать у других стран, но и те, кто думает о будущем и понимает, что ресурсы не безграничны и сохранив хоть часть их сегодня дадут ему возможность на существование потом.

Опущение повторов слов «энергии» и «ресурсы».

Грамматический уровень.

1. И не только те, что не имеют своих источников энергии и вынуждены постоянно закупать ресурсы у других стран, но и те, кто думает о будущем и понимает, что ресурсы не безграничны и **сохранив хоть часть их сегодня** дадут ему возможность на существование потом.

И не только те, что не имеют своих источников, вынуждены постоянно закупать у других стран, но и те, кто думает о будущем и понимает,

что ресурсы не безграничны и, **если сохранить хоть часть их сегодня**, они дадут возможность на существование потом.

Ошибка в употреблении деепричастного оборота исправлена приемом перестройки части предложения в придаточное условия.

Стилистический уровень.

1. На данный **момент времени** общее потребление энергии в мире составляет приблизительно 1050 млрд. кВт/ч в год.

На данный **момент** общее потребление энергии в мире составляет приблизительно 1050 млрд. кВт/ч в год.

Слово «времени» в данном случае лучше опустить.

2. Россия от общего мирового **потребления** энергии **потребляет** около 5%.

На долю России от общего мирового **потребления** энергии **приходится** около 5%.

Тавтологию в данном предложении следует исключить приемами замены и добавления.

После проведения оптимизации:

- Уровень заспамленности текста был снижен до 5,80 (без учета стоп-слов – 5,43).

- Частотность слов: «энергия» (34 повтора при 3,62% плотности), «ветра» (11 повторов и 1,28% плотности).

- Частотность стоп-слов: «в» (27 повторов, 3,05% плотности), «и» (21 повтор, 2,78% плотности), «на» (16 повторов, 1,86% плотности).

Переходим к лингвистическому анализу текстов, полученных при запросе к поисковой системе Яндекс:

1) Статья «Альтернативные источники энергии».

Источником статьи является информационный ресурс: <http://infoelektrik.ru/alternativnoe-elektrichestvo/alternativnoe-elektrichestvo.html>.

Сайт публикует статьи по теме «Электричество».

Данная статья освещает тему альтернативных вариантов получения энергии, описаны основные виды альтернативных источников.

Согласно SEO-анализу в данном тексте:

- Общее количество слов – 1582, количество слов без учета стоп-слов – 1217.
- Уровень заспамленности текста равен 7,87 с учетом стоп-слов и 6,08 без них.
- В тексте часто встречаются слова: «энергия» с 47 повторами и 2,97% плотности, «альтернативный» – 38 повторов и 2,40% плотности, «источник» – 34 повтора и 2,15% плотности.
- Частота стоп-слов: «и» (62 повтора, 3,92% плотности), «в» (46 повторов, 2,91% плотности), «для» (33 повтора, 2,09% плотности).

Лексический уровень.

1. Человек издревле использовал **энергию ветра** как движитель для кораблей, что позволяло развиваться торговле.

Человек издревле использовал **ветер** как движитель для кораблей, что позволяло развиваться торговле.

Использование приема генерализации, чтобы не повторять слово «энергию».

2. **Энергия** перепада воды приводила в действие мельничные жернова.

Механизм перепада воды приводил в действие мельничные жернова.

Замена с целью устранения часто встречающегося слова «энергия».

Грамматический уровень.

1. **В качестве материалов используется** дерево, металл, пластик и т.д.

К используемым материалам относятся: дерево, металл, пластик и т.д.

«В качестве...» встречается несколько раз в одном абзаце, повтор исключается заменой.

2. Биологический газ **в качестве источника** энергии получают в основном двумя способами – это пиролиз и анаэробное (без доступа кислорода) разложение органических веществ.

Биологический газ, **как источник энергии**, получают в основном двумя способами – это пиролиз и анаэробное (без доступа кислорода) разложение органических веществ.

«В качестве...» встречается несколько раз в одном абзаце, повтор исключается заменой.

«... в качестве» повторяется в пределах одного абзаца более одного раза и требует замены.

Стилистический уровень.

1. Другим **колоссальным** источником альтернативной энергии может служить атмосферное электричество.

Другим **мощным** источником альтернативной энергии может служить атмосферное электричество.

Синонимическая замена объясняется лексической несочетаемостью.

После оптимизации текста:

- Уровень заспамленности составил 6,31 с учетом стоп-слов и 5,98 без них.

- Часто встречаются слова: «энергия» с 38 повторами и 2,16% плотности, «альтернативный» – 32 повтора и 2,07% плотности, «источник» – 24 повтора и 1,55% плотности.

- Частотность стоп-слов: «и» (49 повторов, 2,95% плотности), «в» (40 повторов, 2,59% плотности), «для» (35 повторов, 2,27% плотности).

2) Статья «Альтернативные источники энергии».

Источником статьи является информационный ресурс новостей науки и техники: <http://s-and-t.ru/technology/item/70-alternativnye-istochniki-energii/>.

В статье рассказывается о некоторых видах альтернативных источников энергии, а также приведены примеры использования в разных странах.

SEO-анализ выявил, что в данном тексте:

- Общее количество слов – 848 (количество слов без учета стоп-слов – 594).

- Уровень заспамленности текста – 6,48% (3,16% без учета стоп-слов).

- Часто встречаются следующие слова: «энергия» (10 повторов и 1,18% плотности), «солнца» (6 повторов и 0,71% плотности).

- Частотность стоп-слов: «в» (42 повтора, 4,95% плотности), «и» (28 повторов, 3,30% плотности), «на» (13 повторов, 1,53% плотности).

Лексический уровень.

1. Архимед поджег флот неприятелей, используя зеркала **в роли линз**.

Архимед поджег флот неприятелей, используя зеркала **как линзы**.

Произведена замена, чтобы исключить предлог «в».

Грамматический уровень.

1. **В наши дни** разработки в области ветрогенерации набирают оборотов.

Сейчас разработки в области ветрогенерации набирают **обороты**.

Исправлена ошибка в согласовании.

Стилистический уровень.

1. Так, что рано или поздно **количество полезных ископаемых подойдет к концу** и настанет, так называемый, энергетический кризис.

Так, что рано или поздно **полезные ископаемые закончатся** и настанет, так называемый, энергетический кризис.

Применены приемы упущения и замены с целью упрощения предложения.

Оптимизация текста дала следующие результаты:

- Уровень заспамленности текста – 4,90% (3,16% без учета стоп-слов).

- Частотность слов: «энергия» (10 повторов и 1,18% плотности), «солнца» (6 повторов и 0,71% плотности).

- Частотность стоп-слов: «в» (24 повтора, 2,94% плотности), «и» (23 повтора, 2,82% плотности).

3) Статья «Альтернативные источники энергии».

Источником статьи является следующий информационный ресурс:
http://alternativa.dviger.com/projects/work/c_409.html.

Сайт является молодежным порталом «Движер». Различные блоки данного портала предоставляют возможность проявить себя пользователям во многих сферах, путем размещения информации на сайте.

Текст представляет собой проект по теме альтернативных источников энергии. В тексте приведена информация о выгодных способах добычи человеком энергии. Проект был разработан, с целью выявить наиболее перспективные способы добычи электроэнергии с максимальной пользой и минимальными затратами. Даны описание, виды, история и статистика.

SEO-анализ показал, что в данном тексте:

- Общее количество слов – 4404, количество слов без учета стоп-слов – 3298.
- Уровень заспамленности текста равен 14,21% с учетом стоп-слов и 7,21% без них.
- Часто встречаемые слова: «энергия», у которых 128 повторов и 2,91% плотности, «электроэнергии» – 56 повторов и 1,27% плотности.
- Частота стоп-слов: «в» (202 повтора, 4,59% плотности), «и» (134 повтора, 3,04% плотности), «на» (71 повтор, 1,61% плотности).

Лексический уровень.

1. Сложность в принятии проекта ВЭС к реализации и бюрократические проволочки не только могут **в несколько раз** увеличить стоимость проекта, но и просто не дать возможности его довести до конца.

Сложность в принятии проекта ВЭС к реализации и бюрократические проволочки не только могут **существенно** увеличить стоимость проекта, но и просто не дать возможности его довести до конца.

Замена помогает сократить количество слов, вместо трех слов можно использовать одно.

Грамматический уровень.

1. Альтернативный источник энергии — способ, устройство или сооружение, позволяющее получать электрическую энергию (или другой

требуемый вид **энергии**) и **заменяющий** собой традиционные источники **энергии**, функционирующие на нефти, добываемом природном газе и угле.

Альтернативный источник энергии — способ, устройство или сооружение, позволяющее получать электрическую энергию (или другой требуемый ее вид), **а также** заменяющее собой традиционные источники, функционирующие на нефти, добываемом природном газе и угле.

Замена существительного местоимением – устранение повтора, опущение повтора.

Стилистический уровень.

1. Как и в случае с ветровой энергией, в самое последнее время ситуация начала меняться, **но чуть-чуть**.

Как и в случае с ветровой энергией, в последнее время ситуация начала меняться, **однако медленным темпом**.

Замена необходима в данном случае, поскольку лексика не соответствует общей атмосфере статьи.

2. Действительно, только представьте себе фантастическую **на сегодня** картину – несколько сотен таких установок **хотя бы лишь** в наиболее солнечных регионах Казахстана – от Мангышлака до Талдыкоргана, сколько угля, дорогостоящего мазута и газа можно было бы **экономить!**

Действительно, только представьте себе фантастическую **для нашего времени** картину – несколько сотен таких установок **хотя бы** в наиболее солнечных регионах Казахстана – от Мангышлака до Талдыкоргана, сколько угля, дорогостоящего мазута и газа можно было бы **экономить!**

Замена и опущение, в данном случае, необходимы для того, чтобы статья не нарушала правило грамотного представления текста.

Оптимизация текста дала следующие результаты:

- Уровень зашамленности текста равен 8,59% с учетом стоп-слов и 7,01% без них.

- Часто встречаемые слова: «энергия», у которых 111 повторов и 2,48% плотности, «электроэнергии» – 41 повтор и 0,89% плотности.

- Частота стоп-слов: «в» (151 повтор, 3,49% плотности), «и» (114 повторов, 2,56% плотности), «на» (71 повтор, 1,61% плотности).

4) Статья «10 источников энергии будущего».

Источником данной статьи является: <https://rodovid.me/energy/10-istochnikov-energii-buduschego.html>.

Сайт «Родовид» представляет собой коллективный блог, где каждый пользователь-участник может выкладывать статьи о своих достижениях и открытиях в области ответственного потребления и устойчивого развития.

В статье рассказывается о возможных источниках энергии будущего.

В данном тексте, согласно SEO-анализу:

- Общее количество слов – 1385, тогда как количество слов без учета стоп-слов – 939.

- Уровень заспамленности текста равен 6,86 с учетом стоп-слов, и 4,36 без них.

- Среди часто встречаемых слов можно увидеть: «энергия» (34 повтора и 2,45% плотности), «мы» (7 повторов и 1,28% плотности).

- Часто встречаемые стоп-слова: «в» (47 повторов, 3,39% плотности), «и» (26 повторов, 1,88% плотности), «на» (23 повтора, 1,66% плотности).

Лексический уровень.

1. **Из-за упора** на энергию солнца и ветра, приливную энергетику вытеснили из первых рядов, **но** она **становится** более эффективной.

После перехода на энергию солнца и ветра, приливную энергетику вытеснили из первых рядов, **однако** она **считается** более эффективной.

Приемы замены позволяют сохранить логику повествования.

Грамматический уровень.

1. Каждый час земля получает **столько солнечной энергии, больше, чем** земляне ее **используют** за целый год.

Каждый час земля получает **солнечную энергию в объемах, которые** земляне **не** используют **даже** за целый год.

Перестройка предложения объясняется необходимостью использовать правильную форму изложения информации.

2. Один из способов **использование** этой энергии, создание гигантских солнечных ферм, которые будут собирать часть высокоинтенсивного и бесперебойного солнечного излучения.

Один из способов **использования** этой энергии, создание гигантских солнечных ферм, которые будут собирать часть высокоинтенсивного и бесперебойного солнечного излучения.

Исправление ошибки согласования.

3. **Речь идет о микроэлектронике, но потенциал велик**, при целевой аудитории в миллиард людей.

Речь идет о микроэлектронике. Потенциал направления велик, при целевой аудитории в миллиард людей.

Необходимость разделения предложения на две части, чтобы правильно передать смысл сказанного.

Стилистический уровень.

1. Одна из причин, почему этот проект **находится на стадии идеи** – это его **огромная** стоимость.

Одна из причин, почему этот проект **пока не реализуется** – это его **высокая** стоимость.

Замена связана с необходимостью сохранить стилистическую атмосферу статьи.

Результаты оптимизации текста:

- Уровень заспамленности текста равен 5,67 с учетом стоп-слов, и 5,24 без них.

- Среди часто встречаемых слов можно увидеть: «энергия» (31 повтор и 2,29% плотности), «мы» (7 повторов и 1,28% плотности).

- Часто встречаемые стоп-слова: «в» (33 повтора, 2,42% плотности), «и» (24 повтора, 1,76% плотности), «на» (23 повтора, 1,66% плотности).

Все найденные статьи можно считать релевантными запросу «Альтернативные источники энергии», потому как напрямую относились к данной теме. Изначально по показателям качества текста (согласно SEO-анализу) можно было сделать вывод, что тексты не были оптимизированы, что наталкивает на мысль – посещаемость или популярность информационного ресурса имеет большое значение при ранжировании. Тогда оптимизация текстов должна быть инструментом для сайтов, желающих заинтересовать пользователей и тем самым поднять рейтинг посещения.

Примечательно, что тексты, представленные на третьей странице поисковой выдачи в системах Google и Яндекс, могут представлять интерес для пользователей.

Результаты трансформаций, осуществленных на уровне лексики, грамматики и стилистики, привели к существенному изменению показателей при повторном SEO-анализе. Поэтому можно сказать, что данные лингвистические приемы стоит использовать для оптимизации уже готовых текстов, чтобы не писать их заново.

В анализированных текстах статей были найдены разные особенности, на которые стоило обратить внимание при оптимизации: высокая частотность повторов ключевых слов и их словоформ («альтернативный», «энергия» и «источник»), высокая частотность повторов предлогов («в» и «и»), встречались ошибки словообразования и согласования, а также некоторые другие.

Все ошибки в тексте были исправлены, по отношению к лингвистическим компонентам текста, которые снижали качества для поисковых систем и пользователей, были применены соответствующие приемы трансформаций.

Тексты, полученные в результатах выдачи в системе Google (на третьей странице), содержали меньше грамматических и стилистических ошибок. Источники статей также были посвящены узкой тематике, которая отражалась в запросе. В системе Яндекс источники часто

содержались на порталах, где пользователи могут сами представлять какую-либо информацию. Наблюдалось множество грамматических и стилистических ошибок (например, нарушение согласования компонентов предложения).

В качестве результата сравнения работы двух поисковых систем можно заключить следующее: Google в большей степени учитывает уровень грамотности в тексте, что видно даже по результатам выдачи на третьей странице; Яндекс особое внимание уделяет информативности текста и его соответствие запросу, часто в ущерб качеству на уровне грамматики.

2.2 Рекомендации по оптимизации текстов под ИПС

Оптимизация текстов под поисковые системы, как мы увидели, подразумевает умение работать с текстом на разных его уровнях, при этом по отношению к каждому элементу, при условии нарушения каких-либо условий (лингвистических или требований ИПС), необходимо применять соответствующие меры. Такие меры необходимо принимать с учетом того, что текст должен быть в результате грамотным, логичным и естественным (т.е. ориентированным на пользователя).

Говоря о лексике, следует отметить, что наиболее распространенным явлением являются чрезмерные повторы ключевых и служебных (стоп-слов) слов. Частые повторы снижают качество текста, поэтому рекомендуется использовать лингвистические приемы трансформаций (каждый в отдельном случае):

- замены (в виде синонимов и местоимений),
- генерализацию (гиперонимы),
- конкретизацию (гипонимы),
- добавления,
- опущения.

Каждый из данных приемов снижает частотность ключевых слов и предлогов (наиболее распространенные «в» и «и»), которые являются основной причиной высокого уровня зашамленности текстов.

На уровне грамматики часто встречаются конструкции (причастные и деепричастные обороты), которые в некоторых случаях не несут смысловой нагрузки, и, скорее всего, служат для дополнительного объема текста. Некоторые конструкции (например, «как..., так и...») встречаются слишком часто в пределах даже одного абзаца, чего стоит избегать. В этом плане помогут следующие трансформации:

- перестройка предложения,
- опущения,
- добавления,
- перестановки,
- замены.

Всегда следует помнить о правиле упрощения предложений, поскольку текст должен быть легким для восприятия.

Уровень стилистики также предполагает оптимизацию в форме трансформаций, только при условии необходимости. С учетом того факта, что у каждого автора есть свой стиль выражения мысли, все элементы рассматриваются для установления их соответствия общей атмосфере текста. Если целостность нарушается в силу использования определенных слов, фраз и конструкций, то применяются следующие трансформации:

- замены (в том числе синонимы),
- компенсации,
- описания.

Стоит избегать:

- канцеляризм (формальные фразы и выражения, которые используются в официальных документах),

- повторения речевых штампов (без учета контекста),

- плеоназмы (использование оборотов речи, дублирующие смысл сказанного),
- тавтологию (например, однокоренные слова в одном предложении).

Оптимизация текста в большей степени направлена на пользователя, на его информационную потребность. Так требования поисковых систем учитывают показатели качества текста, чтобы результаты отвечали потребности человека. Информативность и грамотность текста являются обязательным условием при отборе и индексации. Именно поэтому необходимо, чтобы все лингвистические компоненты были на своем месте (согласно их функции, семантики и формы).

Данные рекомендации составлены с учетом того, что оптимизация текста под поисковые системы отличается от редактирования. Текст в первом случае не является окончательным вариантом, который можно «поправить», поскольку требует основательной работы. Работа заключается в существенном изменении слов, словосочетаний и конструкций. После внесения изменений в текст (процесса оптимизации) конечный результат можно подвергнуть редактированию.

Выводы по главе 2

В данной главе мы описали этапы проведения практической части исследования.

Во-первых, был сделан запрос к выбранным поисковым системам (Google и Яндекс) в виде фразы «Альтернативные источники энергии». В самом начале работы мы определили, что тексты будут взяты из информационных ресурсов, представленных по первым трем ссылкам на первой странице поисковой выдачи и по трем первым ссылкам на третьей странице. В результате было получено 12 текстов, двое из которых были исключены по причине совпадения с двумя другими текстами (они найдены на одних и тех же сайтах). Все найденные статьи можно считать релевантными запросу «Альтернативные источники энергии»,

поскольку они напрямую относятся к данной теме. Интересен тот факт, что тексты, представленные на третьей странице поисковой выдачи, могут представлять интерес для пользователей.

Во-вторых, был проведен семантический анализ десяти текстов. Данные анализа о компонентах текстов были использованы в качестве основы для оптимизации. Показатели качества текста (по SEO-анализу) говорят о том, что тексты, скорее всего, не были оптимизированы, поскольку уровень заспамленности во многих текстах был выше нормы (по TEXTUS PRO – не более 5). Анализ выявил и высокую частотность ключевых слов и стоп-слов, а также их плотность расположения в текстах.

В-третьих, тексты были оптимизированы с учетом требований поисковых систем и восприятия пользователей сети Интернет. Оптимизация затронула лексику, грамматику и стилистику. В отношении отдельных слов, словосочетаний и предложений были применены лингвистические трансформации (чаще всего применялись замены и опущения, поскольку компоненты текста не несли дополнительной смысловой нагрузки).

Далее были даны рекомендации по проведению оптимизации. Так на трех основных уровнях даны виды трансформаций, которые могут изменить качество текста в лучшую сторону. Также отмечены элементы, которым необходимо уделить особое внимание.

Результаты оптимизации текстов с использованием трансформаций показали существенные изменения при повторном SEO-анализе. Показатели стали ближе к норме (по TEXTUS PRO), что говорит также о повышении их качества, как для пользователей, так и для поисковых систем.

Отсюда следует вывод, что лингвистические приемы необходимо использовать для оптимизации уже готовых текстов, чтобы улучшить их качество.

ЗАКЛЮЧЕНИЕ

Данное исследование было направлено на рассмотрение одного из важнейших лингвистических элементов – текст. Так текстовый контент в Интернете можно увидеть с двух разных сторон: глазами пользователя, целью которого является получение информации в виде текста или документа, и поисковой системы, которая отбирает для пользователя ценные информационные ресурсы. При этом берется во внимание лингвистический аспект, что означает изучение компонентов текста и их влияние на общее восприятие информации. Другими словами проводится анализ на трех основных уровнях: лексика, грамматика и стилистика. Таким образом, можно успешно выявить особенности текста, которые снижают его качество (распределение и использование слов, построение предложений, соблюдение логики).

Современные ПС отбирают информационные ресурсы именно по критериям качества контента, поэтому необходимо умение писать и оптимизировать тексты. Оптимизация осуществляется на основе данных, полученных при SEO-анализе, и с помощью лингвистических приемов трансформаций (например, перестройка предложения или лексическая замена). На практике были получены положительные изменения в оптимизированных текстах.

В данной дипломной работе мы провели лингвистический анализ особенностей текстов, полученных по запросу «Альтернативные источники энергии», и их SEO-оптимизацию на всех трех уровнях. Также обобщили рекомендации для успешного проведения оптимизации текстов под поисковые системы.

С учетом цели и задач исследования мы описали этапы проведения лингвистического анализа и оптимизации, а затем сравнили полученные результаты.

В результате проведенного исследования были сделаны следующие выводы:

1) Даны определения двум основным понятиям в направлении нашего исследования: «SEO-анализ» и «SEO-оптимизация». Два понятия тесно связаны друг с другом.

2) Оптимизация текстов под ПС подразумевает использование в качестве основы данные SEO-анализа и применение лингвистических трансформаций, которые способствуют улучшению качества текста.

3) Обобщение рекомендаций по оптимизации позволяет учитывать большинство параметров работы с текстом.

Результаты исследования помогут расширить понятие SEO-оптимизации текстов в лингвистическом аспекте, что обуславливает теоретическую значимость дипломной работы.

Практическая значимость заключается в том, что результаты исследования могут быть включены в пособия по лексикологии и стилистике.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика [Текст] : учебное пособие / Е. И. Большакова, Э. С. Клышинский, Д. В. Ландэ, А. А. Носков, О. В. Пескова, Е. В. Ягунова. – М.: МИЭМ, 2011. – 272 с.
2. Ашманов, И. С. Продвижение сайта в поисковых системах [Текст] / С. И. Ашманов, А. А. Иванов. – М.: И. Д. Вильямс, 2007. – 304 с.
- Байков, В. Д. Интернет. Поиск информации. Продвижение сайтов [Текст] / В. Д. Байков. – Спб.: БХВ-Петербург, 2000. – 288 с.
3. Баранов, А. Н. Введение в прикладную лингвистику [Текст] : учебное пособие / А. Н. Баранов. – М.: Эдиториал УРСС, 2001. – 360 с.
4. Белоногов, Г. Г. Компьютерная лингвистика и перспективные информационные технологии [Текст] / Г. Г. Белоногов. – М.: Русский мир, 2004. – 248 с.
5. Беляева, Л. Н. Лингвистические автоматы в современных гуманитарных технологиях [Текст] : учебное пособие / Л. Н. Беляева. – СПб.: Книжный дом, 2007. – 192 с.
6. Головкина, С. Х. Лингвистический анализ текста [Текст] / С. Х. Головкина, С. Н. Смольников. – Вологда: Издательский центр ВИРО, 2006. – 124 с.
7. Гроховский, Л. SEO: руководство по внутренним факторам [Текст] / Л. Гроховский. – М.: ТопЭкспертРФ, 2011. – 133 с.
8. Евдокимов, Н. В. Основы контентной оптимизации. Эффективная интернет-коммерция и продвижение сайтов в Интернете [Текст] / Н. В. Евдокимов. – М.: Вильямс, 2010. – 160 с.
9. Евдокимов, Н. В. Раскрутка веб-сайта: практическое руководство по SEO [Текст] / Н. В. Евдокимов, И. В. Лебединский. – М.: Вильямс, 2011. – 288 с.

10. Жуликов, С. Е. Проблема пертинентности современных информационно-поисковых систем [Текст] / С. Е. Жуликов, О. В. Жуликова // Вестник ТГУ. Сер. Естественные и технические науки. – Тамбов: Изд-во ТГУ, 2013. – Т.18.– Вып.1. – С. 224-226.
11. Захаров, В. П. Лингвистические средства информационного поиска в Интернете [Текст] / В. П. Захаров // Библиосфера. – 2005. – №1 – С. 63-71.
12. Зубов, В. А. Информационные технологии в лингвистике [Текст] / А. В. Зубов, И. И. Зубова. – 2-е изд., испр. – М.: Академия, 2012 – 206 с.
13. Иванов, И. И. SEO: Поисковая оптимизация от А до Я. Основы [Текст] / И. И. Иванов. – Интернет-издание, 2011. – 303 с.
14. Иванов, И. И. SEO: Поисковая оптимизация от А до Я. Средний уровень [Текст] / И. И. Иванов. – Интернет-издание, 2011. – 339 с.
15. Иванов, И. И. SEO: Поисковая оптимизация от А до Я. Продвинутый уровень [Текст] / И. И. Иванов. – Интернет-издание, 2011. – 278 с.
16. Кажарнович, В. Ф. SEO на результат. Простые и понятные методы продвижения в Интернете [Текст] / В. Ф. Кажарнович. – СПб.: Питер, 2017. – 320 с.
17. Калитин, С. В. Компьютерные технологии в науке [Текст] : учебное пособие / В. С. Калитин. – Хабаровск: Хабаровская гос. акад. экономики и права, 2011. – 95 с.
18. Кент, П. Поисковая оптимизация для чайников [Текст] / П. Кент. – 4-е издание. – М.: Вильямс, 2011. – 432 с.
19. Колисниченко, Д. Н. Поисковые системы и продвижение сайтов в Интернете [Текст] / Д. Н. Колисниченко. – М.: Вильямс, 2007. – 272 с.
20. Кононова, И. В. SEO-оптимизированный текст как жанровая разновидность компьютерно опосредованного рекламного дискурса [Текст] / И. В. Кононова, Ф. А. Николаев // Вестник ВолГУ. Сер. Языкознание. – Волгоград: ВолГУ, 2016. – Т.15. – Вып. 4. – С. 96-103.
22. Кураков, А. Эффективное продвижение сайтов. Реальное SEO [Текст] / А. Кураков, М. Райцин.: СамИздат, 2011. – 61 с.

23. Левицкая, Т. Р. Теория и практика перевода с английского языка на русский [Текст] / Т. Р. Левицкая, А. М. Фитерман. – М.: Издательство литературы на иностранных языках, 1963. – 125 с.
24. Марчук, Ю. Н. Компьютерная лингвистика [Текст] / Ю. Н. Марчук. – М.: АСТ Восток-Запад, 2007. – 317 с.
25. Первая книга SEO-копирайтера. Как написать текст для поисковых машин и пользователей [Текст] / О. И. Крохина, М. Н. Полосина, А. В. Рубель, О. И. Сахно, Е. В. Селин, М. С. Ханина. – М.: Инфра-Инженерия, 2012. – 216 с.
26. Севостьянов, И. О. Поисковая оптимизация. Практическое руководство по продвижению сайта в Интернете [Текст] / И. О. Севостьянов. – Спб.: Питер, 2010. – 240 с.
27. Семантическое ядро: 6 адовых ошибок – URL: <https://blog.ingate.ru/books/>, свободный. – Загл. с экрана (дата обращения: 10.02.2017).
28. Соснина, Е. П. Введение в прикладную лингвистику [Текст] : учебное пособие / Е. П. Соснина. – 2-е изд., испр. и доп. – Ульяновск : УлГТУ, 2012. – 110 с.
29. Терехов, А. А. Разработка методов и инструментальных средств повышения пертинентности поиска в современных информационных средах [Текст] : дис. ... канд.тех.наук: 05.13.17 / Терехов Алексей Андреевич. – Москва, 2010. – 165 с.
30. Шакин, М. SEO Винчестер Эффективные методы оптимизации коммерческих сайтов [Текст] / М. Шакин.: Интернет-издание, 2014. – 91 с.
31. Щипицина, Л. Ю. Информационные технологии в лингвистике [Текст] : учебное пособие / Л.Ю.Щипицина. – М.: Изд-во ФЛИНТА, 2013. – 128 с.
32. Яковлев, А. А. Раскрутка и продвижение сайтов: основы, секреты, трюки [Текст] / А. А. Яковлев. – Спб.: БХВ-Петербург, 2007. – 336 с.

33. SEO – искусство раскрутки сайтов [Текст] / Э. Энж, С. Спенсер, Р. Фишкин, Д. Стрикчиола. – Пер. с англ., 2-е изд, перераб. и доп. – СПб.: БХВ-Петербург, 2014. – 688 с.

СПИСОК ЛИТЕРАТУРЫ, ИСПОЛЬЗОВАННОЙ ДЛЯ АНАЛИЗА

34. Альтернативные источники энергии [Электронный ресурс]. – URL: <http://elektrik.info/main/news/614-alternativnye-istochniki-energii.html>, свободный. – Загл. с экрана (дата обращения: 7.02.2017).

35. Альтернативные источники энергии [Электронный ресурс]. – URL: <http://infoelektrik.ru/alternativnoe-elektrichestvo/alternativnoe-elektrichestvo.html>, свободный. – Загл. с экрана (дата обращения: 16.02.2017).

36. Альтернативные источники энергии [Электронный ресурс]. – URL: <http://s-and-t.ru/technology/item/70-alternativnye-istochniki-energii/>, свободный. – Загл. с экрана (дата обращения: 16.02.2017).

37. Альтернативные источники энергии [Электронный ресурс]. – URL: http://alternativa.dviger.com/projects/work/c_409.html, свободный. – Загл. с экрана (дата обращения: 16.02.2017).

38. Альтернативная энергия – будущее человечества [Электронный ресурс] – URL: <http://madenergy.ru/stati/alternativnye-istochniki-ehnergii-kak-bolshoj-shag-v-razvitie-i-ispolzovanie-ehlektroehnergii.html>, свободный. – Загл. с экрана (дата обращения: 7.02.2017).

39. Альтернативная энергетика [Электронный ресурс] – URL: https://ru.wikipedia.org/wiki/Альтернативная_энергетика, свободный. – Загл. с экрана (дата обращения: 7.02.2017).

40. Есть ли будущее у альтернативных источников энергии? [Электронный ресурс] – URL: <https://pronedra.ru/alternative/2017/03/22/buduscheealternativnoy-energetiki/>, свободный. – Загл. с экрана (дата обращения: 7.02.2017).

41. 8 необычных источников альтернативной энергии для дома, офиса и отдыха [Электронный ресурс] – URL: <http://www.novate.ru/blogs/280-415/31040/>, свободный. – Загл. с экрана (дата обращения: 7.02.2017).

42. 10 источников энергии будущего [Электронный ресурс] – URL: <https://rodovid.me/energy/10-istochnikov-energii-buduschego.html>, свободный. – Загл. с экрана (дата обращения: 16.02.2017).