

СИСТЕМА ДЛЯ АНАЛИЗА ВНУТРЕННИХ ПАРАМЕТРОВ WEB - РЕСУРСА

Бондарчук Н.В.

Научный руководитель: Савельев А.О., кафедра ОСУ
Томский политехнический университет
nataliabondar4uck@yandex.ru

Введение

В современном мире трудно себе представить жизнь без интернета, каждый день мы заходим на множество сайтов и ищем необходимую нам информацию. Для любого бизнеса интернет, а в частности поисковые системы, является одним из эффективных способов привлечения клиентов. Поэтому в поисковых системах происходит борьба между сайтами-конкурентами за право быть в ТОП - 10 поисковых систем, таких как Яндекс и Google, а соответственно за потенциальных клиентов [1].

Исходя из этого было решено разработать систему, которая поможет владельцам бизнеса, маркетологам, PR - менеджерам и SEO - специалистам проанализировать внутренние параметры сайта, влияющие на его выход в ТОП - 10.

Параметры сайта и их характеристики

На производственной практике мною были исследованы характеристики внутренних параметров сайта, в результате чего мне удалось вывести сайт <http://dentalia.tomsk.ru/> в ТОП-10 в Яндекс и Google по некоторому количеству (около 50) ключевых запросов из стоматологической тематики.

Таким образом, были выделены параметры и их характеристики, по которым разработанная система рекомендует пользователю, что необходимо сделать на сайте для его продвижения. Одним из важных составляющих страницы сайта являются текстовые параметры (Таблица 1).

Таблица 1. Текстовые параметры

	Количество символов	Вхождение запроса
Title	60-80	1
Description	130-180	1-2
Keywords	2-15 слов	2-3
Текст	2000-3000	3-4

Текст на странице сайта, кроме перечисленных выше характеристик, имеет еще ряд важных характеристик, таких как наличие маркированных и нумерованных списков, наличие заголовков (H1, H2, H3 и т.д.), длина абзацев (500-1000 символов). Заголовки должны также включать в себя ключевой запрос, при чем заголовок H1 должен быть один на странице и содержать основной запрос.

На странице сайта должны присутствовать 2-3 изображения, которые должны весить не больше

300 Кб и иметь в названии ключевой запрос (например, если запрос «стоматология в Томске», то изображение может называться [stomatologia-v-tomske.jpg](#)). В заголовке и альтернативном тексте изображения также должны быть прописаны ключевые запросы.

Следующим параметром сайта было выделено время загрузки сайта. Сайт должен загружаться не более 1 секунды. При увеличении времени загрузки увеличивается и шанс, что пользователь покинет страницу до того, как она загрузится. То есть сайт потеряет потенциального клиента, а значит он будет неэффективным.

Для корректной индексации поисковыми роботами на сервере сайта должен быть файл `robots.txt`, в котором указано главное зеркало сайта, и информация о том, какие страницы открыты для индексации, а какие закрыты [2]. При отсутствии этого файла или при неправильном его заполнении поисковые роботы могут неправильно выбрать зеркало сайта и проиндексировать все страницы сайта. Таким образом в поисковой выдаче могут оказаться административные и приватные страницы сайта.

Так же для более быстрой индексации новых страниц сайта поисковыми роботами необходима XML-карта. XML-карта служит еще для распределения приоритетов страниц сайта при индексировании, в ней можно указать частоту обновления информации, и даты последнего обновления страницы. Размер XML-карты сайта не должен превышать 10 MB и не должен содержать больше 50000 URL. Если карта превышает данные показатели, надо разбить её на несколько файлов и создать файл индекса `sitemap` [3].

Для повышения CTR (кликабельности) сайта нужен `favicon`. Он позволяет улучшить узнаваемость сайта, а значит увеличить лояльность посетителя.

Следующим пунктом является ЧПУ (человекопонятные URL). Чтобы поисковые системы понимали, о чем страница сайта в них должны присутствовать ключевые запросы (например, <http://dentalia.tomsk.ru/services/lechenie-kariesa>, здесь поисковой системе и посетителю сайта сразу понятно, что речь пойдет о лечении кариеса).

Интерфейс системы

На рисунке 1 изображена вкладка «Текстовые и изображения». В этой вкладке пользователь вводит URL страницы сайта и ключевые запросы, после нажимает кнопку «Начать анализ».

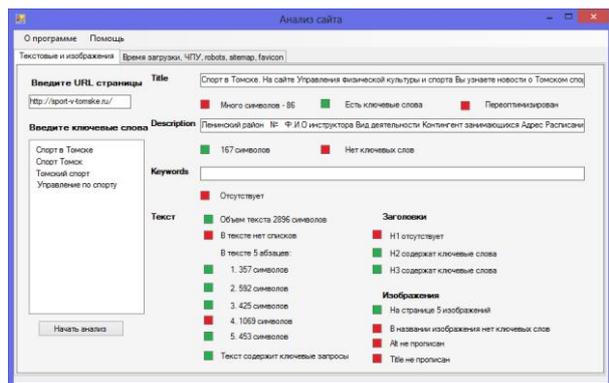


Рис. 1. Вкладка «Текстовые и изображения»

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

1. Отсутствует. – Если теги не прописаны.
- 2.1. Много символов – N. – Если символов больше 80 для Title, больше 180 для Description, больше 15 слов для Keywords.
- 2.2. Мало символов – N. – Если символов меньше 60 для Title, меньше 130 для Description, меньше 2 слов для Keywords.
- 2.3. N символов. – Количество символов в теге.
- 3.1. Есть ключевые слова. – Если в теге есть ключевые слова.
- 3.1. Нет ключевых слов. – Если в теге нет слов.
4. Переоптимизирован. – Если вхождений запроса больше 1 для Title, больше 2 для Description, больше 3 для Keywords.

При анализе текста в окне приложения выводится:

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

В результате анализа заголовков система выдает:

- 1.1. HN отсутствует. – Если на странице отсутствует HN.
- 1.2. HN содержит ключевые слова. – Если в HN есть ключевые слова.
- 1.3. HN не содержит ключевые слова. – Если в HN нет ключевых слов.
2. H1 используется больше одного раза. – Если на странице сайта больше одного заголовка H1.

При анализе изображений в окне приложения выводится:

- 1.1. На странице N изображений. – Количество изображений на странице.
- 1.2. Отсутствуют. – Если на странице нет изображений.
- 2.1. В названии изображения нет ключевых слов. – Если название изображения не содержит ключевые слова.
- 2.2. В названии изображения есть ключевые слова. – Если название изображения содержит ключевые слова.
- 3.1. Alt не прописан. – Если тег Alt не прописан.
- 3.2. Alt прописан. – Если тег Alt прописан.
- 4.1. Title не прописан. – Если тег Title не прописан.
- 4.2. Title прописан. – Если тег Title прописан.

Во вкладке «Время загрузки, ЧПУ, robots, sitemap, favicon» происходит анализ остальных параметров (Рисунок 2).

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

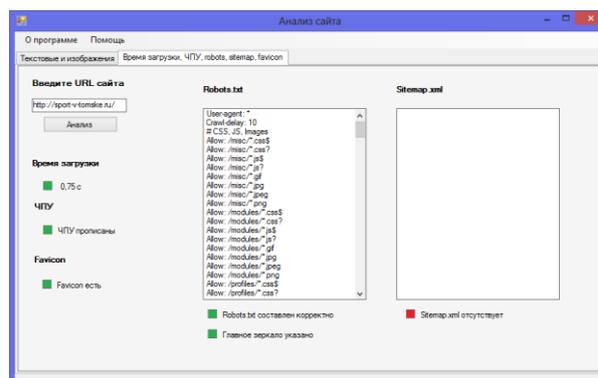


Рис. 2. Вкладка «Время загрузки, ЧПУ, robots, sitemap, favicon»

Заключение

В результате анализа внутренних параметров сайта разработанная система позволяет любому пользователю получить рекомендации именно по оптимизации его web-ресурса.

Список использованных источников

1. Ашманов И. С. «Оптимизация и продвижение сайтов в поисковых системах» - Издательский дом Питер, 2010 - 464 с.
2. Файл robots.txt: создание, установка, применение [Электронный ресурс]. – URL: <https://wildo.ru/fajl-robots-txt> (дата обращения 17.10.2016).
3. Правильный файл sitemap.xml или как создать карту сайта своими руками [Электронный ресурс]. – URL: <http://site-on.net/optimization/6-sitemap> (дата обращения 17.10.2016)