

МОДУЛЬ – 2

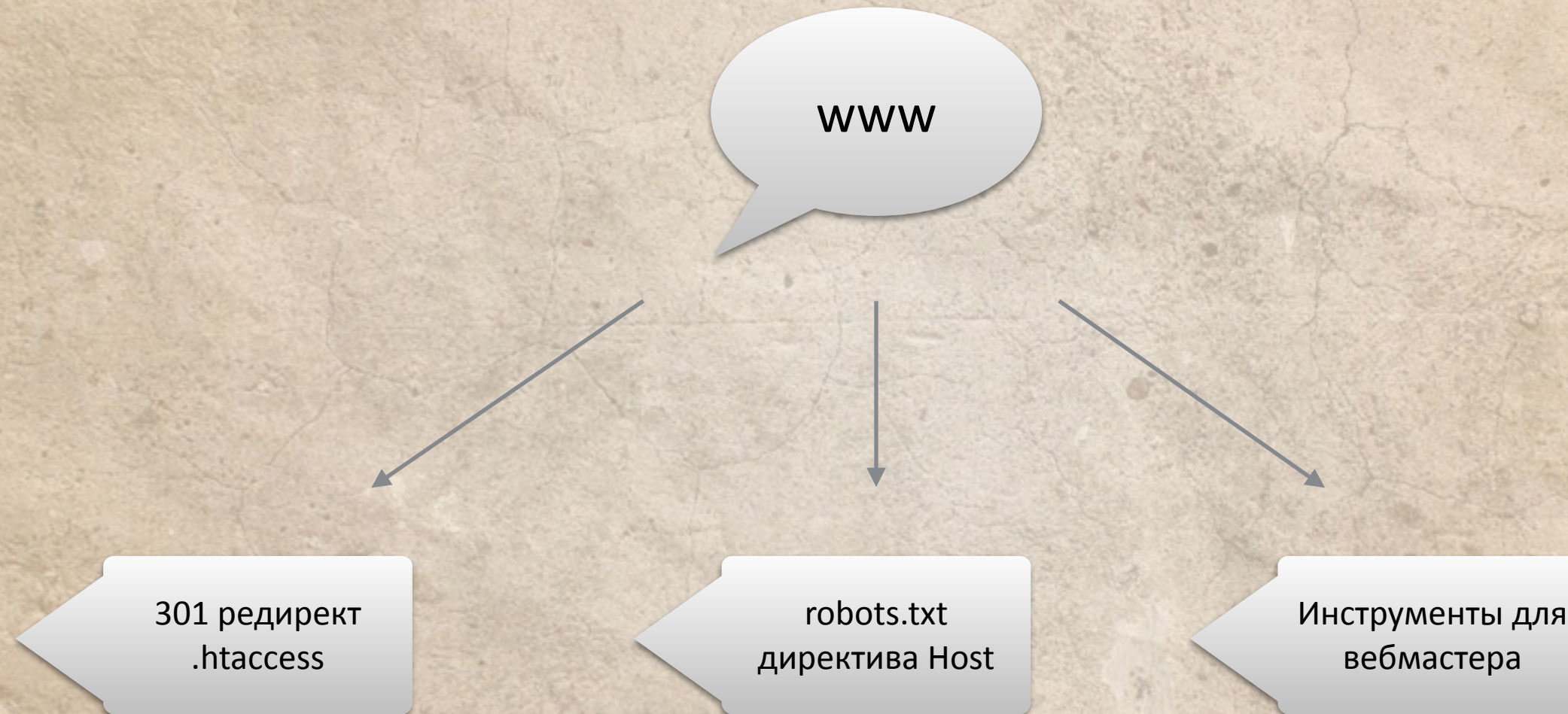
Техническая оптимизация



Техническая оптимизация

- Главное зеркало сайта
- Поиск аффилиатов
- Индексация сайта
- Мусорная и неполная информация
- Дубликаты
- Карта сайта
- Robots.txt и теги
- Валидация кода
- Правильность верстки
- Битые ссылки
- Уникальность основных элементов
- ЧПУ
- HTTP заголовки и скорость загрузки
- Сквозные блоки
- Возможные проблемы с хостингом

Главное зеркало



Аффилаты

- Сайты одной компании по одной тематике
- Одинаковая информация (!!!)
- Одинаковая контактная информация

Сервис – tools.promosite.ru

Индексация сайтов

- Различная индексация в Яндекс и Google
- Загружено роботом и сколько всего страниц в поиске
- Просматриваем страницы исключенные роботом
- Ошибки сканирования

Мусорная и неполная информация

- Страницы в разработки, незаполненные и т.д.
- Страницы с техническими ошибками
- Результаты поиска по сайту
- Страницы отдающие ответ 200 на несуществующие адреса
- Страницы закрытые в robots.txt и с ссылками
- **ДУБЛИКАТЫ**

Полные дубли

1. <http://site.ru/>
2. <http://www.site.ru/>
3. <http://site.ru/page/>
4. <http:site.ru/page>
5. <http://site.ru/index.php>

Причины появления

1. CMS и некорректная настройка модулей ЧПУ
2. Технические ошибки
3. Большие сквозные блоки
4. Страницы со служебной информацией
5. Контент представлен разными способами (.doc, .pdf)

Чем опасны дубликаты

1. Смена релевантных страниц
2. Потеря внутреннего ссылочного веса
3. Потеря общей уникальности контента на сайте
4. Санкции и пессимизация

Поиск дубликатов

1. А сколько должно быть страниц и сколько страниц в индексе?
2. Ищем дубликаты в SERP (операторы - «**site:**www.siter.ru», «**title:**(Заголовок документа)», «**mime:**pdf»)
3. Используем Хену и Page Weight
4. Инструменты для веб-мастера
5. сайтрепорт.рф и rooletka.ru

Как бороться с дубликатами

1. Руками
2. 301 редирект
3. rel = canonical
4. robots.txt
5. Мета тег robots

301 редирект

301 редирект с домена без WWW на домен с WWW:

```
Options +FollowSymLinks
RewriteEngine On
RewriteCond %{HTTP_HOST} ^site.ru$ [NC]
RewriteRule ^(.*)$ http://www.site.ru/$1 [R=301,L]
```

301 редирект с домена с WWW на домен без WWW:

```
Options +FollowSymLinks
RewriteEngine On
RewriteCond %{HTTP_HOST} ^www.site.ru$ [NC]
RewriteRule ^(.*)$ http://site.ru/$1 [R=301,L]
```

301 редирект – удаляем “/” на конце URL:

```
RewriteCond %{REQUEST_FILENAME} !-d
RewriteCond %{REQUEST_URI} ^(.+)/$
RewriteRule ^(.+)/$ http://www.site.ru/$1 [R=301,L]
```

301 редирект – добавляем “/” на конце URL:

```
RewriteCond %{REQUEST_FILENAME} !-f
RewriteCond %{REQUEST_URI} !\..{1,10}$
RewriteCond %{REQUEST_URI} !(\.)/$
RewriteRule ^(.*)$ http://www.site.ru/$1/ [L,R=301]
```

301 редирект – удаляем index.php для главной:

```
RewriteEngine On
RewriteBase /
RewriteCond %{THE_REQUEST} ^[A-Z]{3,9}\ /index.php HTTP/
RewriteRule ^index.php$ http://www.site.ru/$1 [R=301,L]
```

301 редирект со страницы:

```
Redirect 301 /old.html http://site.ru/new.html
```

Атрибут rel="canonical"

```
<link rel="canonical" href="http://www.site.ru/page"/>
```

Каноническая страница - это рекомендуем экземпляр из набора страниц с похожим содержанием

Внимание! Данная директива может быть проигнорирована!

Meta name="robots"

<meta name="robots" content="ЗНАЧЕНИЕ"/>

Варианты значений:

index — индексировать страницу.

noindex — не индексировать страницу.

follow — переходить дальше по ссылкам с этой страницы.

nofollow — не переходить по ссылкам с этой страницы

all — аналог одновременного использования index и follow.

Разрешено индексировать и переходить по ссылкам.

none — аналог одновременного использования noindex и nofollow.

Индексация полностью запрещена, по ссылкам не ходить.