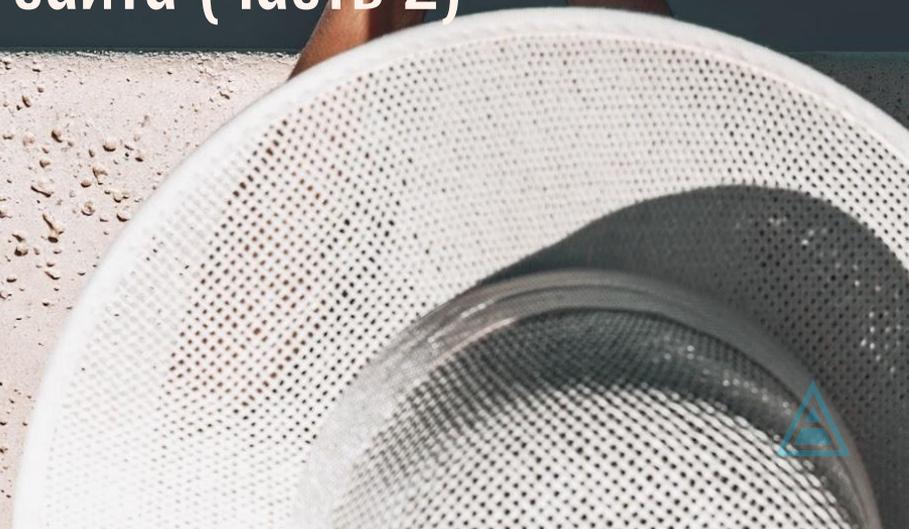


- занятие 8 -

Технический аудит сайта (часть 2)





01 Редиректы и их виды

02 Правила формирования URL

03 Дубли страниц и как их избежать

04 Атрибут canonical

05 Атрибут hreflang

06 Ограничение индексации страниц



01



Редиректы и их виды.



Что такое редиректы и их виды.

Редирект – перенаправление пользователя с одной страницы на другую (с одного URL на другой). Редиректы могут перенаправлять посетителей как на страницы того же сайта, так и на страницы другого сайта.

Можно выделить следующие основные виды редиректов/перенаправлений в интернете:

- 301 и 308 редирект – постоянное перенаправление. Используется когда одна страница полностью и навсегда переезжает на другой адрес;
- 302 и 307 редиректы – временное перенаправление;
- редирект с использованием Meta refresh. При этом, текущая вкладка браузера перезагружается.



Постоянный редирект (301, 308)

Когда используем:

- если страница переехала на новый URL-адрес.
- если обнаружены дубли страниц и их нужно “склеить”.

1. Проверяем настройку главного зеркала сайта:

- 301 редирект с http на https

`http://www.olx.ua/` на `https://www.olx.ua/`

- 301 редирект с www. на non-www.

`https://olx.ua/` на `https://www.olx.ua/` (или наоборот)

- 301 редирект с index.html на главную

`https://www.olx.ua/index.html` на `https://www.olx.ua/`

- 301 редирект с index.php на главную

`https://www.olx.ua/index.php` на `https://www.olx.ua/`



Постоянный редирект (301, 308)

Когда используем:

2. Проверяем возможные дубли внутренних страниц (выборочно на любой внутренней странице сайта):

- 301 редирект с множеством слешей в URL

`https://www.olx.ua///dom-i-sad/` на `https://www.olx.ua/dom-i-sad/`

- 301 редирект с символами верхнего регистра

`https://www.olx.ua/DOM-i-sad/` на `https://www.olx.ua/dom-i-sad/`

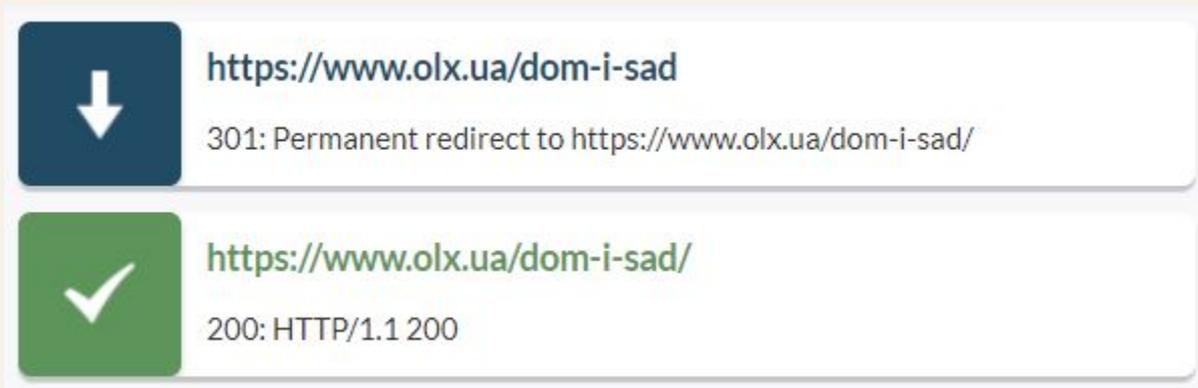
- 301 редирект с завершающих слешем

`https://www.olx.ua/dom-i-sad/` на `https://www.olx.ua/dom-i-sad/` (или наоборот)

На сайте не должно
быть ссылок на
редиректы.



Постоянный редирект (301, 308)



Redirect Path - с помощью расширения удобно отслеживать редиректы.



Meta и JS редиректы.

```
<meta http-equiv="refresh"  
content="...;url=..." />
```

Этот метатег направляет пользователя на новый URL через определенный период времени. Иногда он используется как простой способ переадресации, но поддерживается не всеми браузерами и может вводить пользователей в заблуждение. W3C не рекомендует использовать этот тег. Мы рекомендуем вместо этого применять переадресацию 301, выполняемую сервером.

- Иногда в браузерах отключен JavaScript с целью безопасности. Тогда обновление не сработает.
- Может не срабатывать кнопка «Назад» в браузере пользователя (в зависимости от версии браузера и прописанного кода).
- В результате может проиндексироваться не та страница. Ограничение свободы пользователя. Страницы перезагружаются автоматически, и это вызывает раздражение.
- Поисковый бот может определить страницу с meta-редиректом как спам.
- В старых версиях браузера такие переходы могут срабатывать некорректно.



02

Правила формирования URL.



Правила создания ЧПУ URL

1. без нижнего подчеркивания и пробелов;
2. в нижнем регистре;
3. без спецсимволов;
4. латиницей согласно правильной транслитерации (сena или tsena);
5. без спама;
6. сокращаем не в ущерб логике.



Оптимизация URL

- Структуру сайта лучше отображать в URL

/noutbuki/asus/core-i5

/noutbuki-asus-core-i5/

- Не стоит делать URL важных страниц с большим количеством папок

/noutbuki/igrovie/asus/core-i5

/catalog/c/noutbuki/igrovie/asus/core-i5

- Частотная категория во вложенности выше

/noutbuki/asus/core-i5

/noutbuki/core-i5/asus/



Оптимизация URL

- Короткие URL

`/vred-kurenia-posle-30`

`/statya-o-vrede-kurenia-posle-30-let`

- Статические ЧПУ адреса вместо динамических

`site.com/?sectionid=2&pageid=345 => site.com/noutbuki/asus/`

`site.com/?pageid=345§ionid=2 => site.com/noutbuki/asus/`



Динамические URL

- Короткие URL

`/vred-kurenia-posle-30`

`/statya-o-vrede-kurenia-posle-30-let`

- Статические ЧПУ адреса вместо динамических

`site.com/?sectionid=2&pageid=345 => site.com/noutbuki/asus/`

`site.com/?pageid=345§ionid=2 => site.com/noutbuki/asus/`



03



Дубли страниц и как их избежать.



Дубли – это отдельные страницы сайта, контент которых полностью или частично совпадает. По сути, это копии всей страницы или ее определенной части, доступные по уникальным URL-адресам.



Причины появления дублей

1. Автоматическая генерация дублирующих страниц движком системой управления содержимым сайта (CMS) веб-ресурса. Например:

`http://site.net/press-centre/cat/view/identifier/novosti/`

`http://site.net/press-centre/novosti/`

2. Ошибки, допущенные вебмастерами. Например, когда один и тот же товар представлен в нескольких категориях и доступен по разным URL.
3. Изменение структуры сайта, когда уже существующим страницам присваиваются новые адреса, но при этом сохраняются их дубли со старыми адресами. Например:
4. Добавление в URL GET-параметра, gclid или uclid метки.
5. Отсутствие настройки базовых редиректов.



Типы дублей

1. Полные дубли.

Это страницы с идентичным содержимым, доступны по уникальным, неодинаковым адресам.

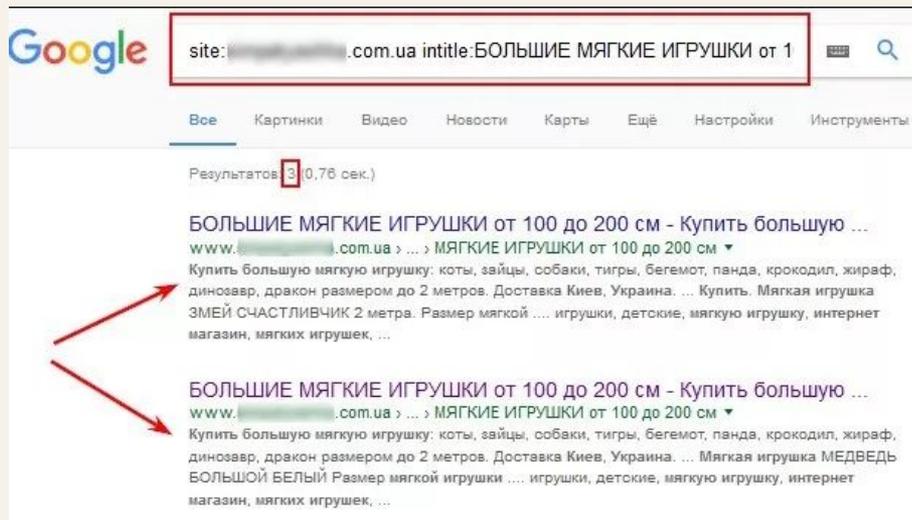
2. Частичные дубли.

В частично дублирующихся страницах контент одинаковый, но есть небольшие отличия в элементах.



К чему приводят дубли страниц на сайте

1. Проблемы с индексацией. При генерировании дублирующих страниц увеличивается общий размер сайта.
2. Изменения релевантной страницы в выдаче. Алгоритм поисковика может решить, что дубль больше подходит запросу.
3. Потеря ссылочного веса страниц, которые продвигаются. Посетители будут ссылаться на дубли, а не на оригиналы страниц.



Как избавиться от дублей

1. 301 редирект. Основной метод борьбы с дублями.
2. Файл robots.txt
3. Метатег `<meta name="robots" content="noindex, nofollow>` и `<meta name="robots" content="noindex, follow>`
4. Атрибут `rel="canonical"`



04



Атрибут
canonical.



Зачем нужен этот атрибут?

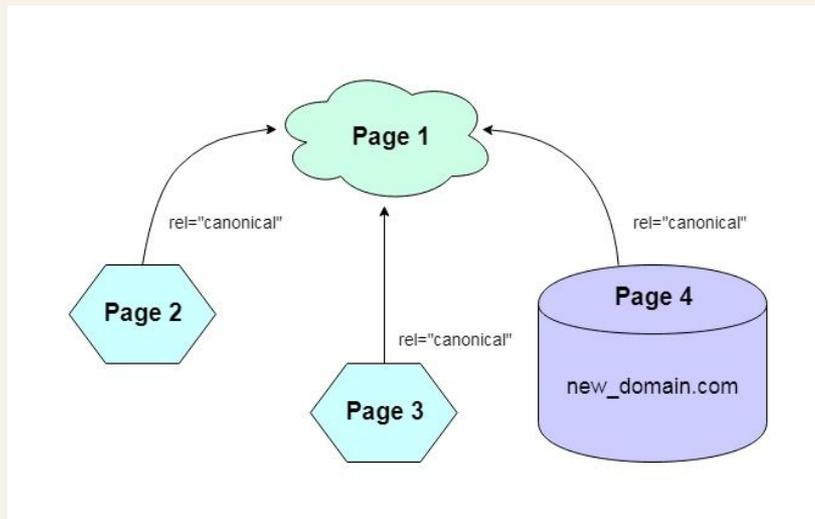
Атрибут `rel="canonical"` определяет среди похожих страниц сайта приоритетную – каноническую. Наличие страницы-первоисточника дает верный сигнал поисковой системе.

Каноническая ссылка также может указывать на основную страницу с другого домена, если на нем представлен аналогичный контент.



Когда применять

1. На всех страницах открытых к индексации.
2. Страницы сортировки
3. Страницы похожих товаров
4. Программа лояльности для клиентов и партнеров
5. Страницы пагинации
6. Работа с UTM-метками



Правила реализации

1. Одна страница – одна каноническая ссылка в `<head>`.
2. Убедитесь, что каноническая страница в индексе.
3. Не усложняйте структуру ссылок, выстраивая цепочки канонических URL.
4. При размещении ссылок с атрибутом `rel="canonical"` разными методами, например, страницы сайта и XML-карта, URL-адреса приоритетных страниц должны совпадать.
5. Выбирайте абсолютные ссылки вместо относительных. Так вероятность ошибки в каноническом URL сводится к минимуму.

Неправильно: `<link rel="canonical" href="shop.com/tables.html" />`

Правильно: `<link rel="canonical" href="http://shop.com/tables.html" />`



05



Атрибут
hreflang.



Атрибут hreflang.

Нужен для правильной индексации языковых и региональных версий сайта.\

Когда используется:

1. Переведены только пункты меню и сведения в подвале страницы, а основная информация оставлена на основном языке. Этот вариант чаще всего используется для разделов с пользовательским контентом – на форумах или страницах с отзывами.
2. На сайте используется один язык с некоторыми региональными различиями. Например, содержимое на английском немного различается для жителей США, Новой Зеландии и Австралии.
3. Все страницы проекта существуют на различных языках. К примеру, ресурс на итальянском перевели на французский и испанский языки.

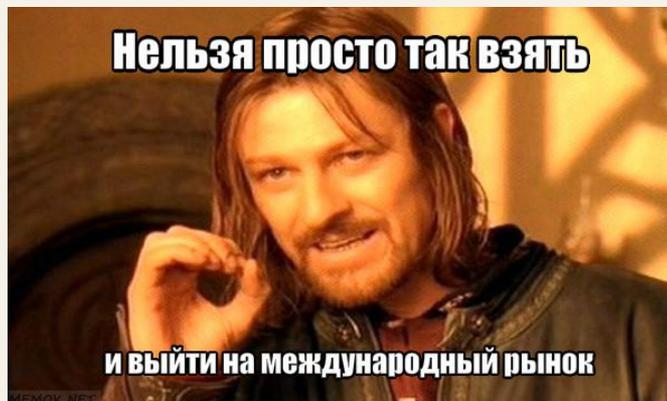


Атрибуты hreflang нужны, чтобы указать ссылки на страницы сайта, созданные на различных языках и в нескольких регионах. Они помогают поисковым системам отображать в выдаче результаты, оптимизированные под определенный регион, язык либо диалект.



Правила реализации

1. Теги будут приниматься во внимание поисковиками только если все языковые страницы ссылаются друг на друга. При этом на всех вариациях страницы должна присутствовать гиперссылка на саму себя.
2. Адреса необходимо указывать полностью, начиная с HTTP либо HTTPS.
3. Версии страниц могут размещаться на различных доменах.
4. Атрибут hreflang может содержать информацию только о коде языка либо дополняться данными о регионе, для которого создана определенная страница.



Правила реализации

В раздел `<head>` страницы добавить элемент `link`, который указывает на языковые версии веб-сайта.

Например на странице `site.com/page/` в `<head>` добавить

```
<link rel="alternate" hreflang="en" href="site.com/page/" />
```

```
<link rel="alternate" hreflang="ru-KZ" href="site.com/kz/page/" />
```

```
<link rel="alternate" hreflang="ru-RU" href="site.com/ru/page/" />
```

```
<link rel="alternate" hreflang="ru-UA" href="site.com/uk/page/" />
```



Hreflang x-default

Чтобы указать дефолтную страницу, можно использовать необязательное значение x-default.

Данный параметр добавляется в следующих случаях:

- при наличии на проекте IP-зависимых страниц, с которых пользователи перенаправляются на подходящую языковую версию сайта;
- для страниц, на которых реализована возможность динамического выбора контента на каком-либо языке;
- на страницы, где осуществляется выбор региона и языка.



06

Ограничение индексации страниц.



Зачем скрывать страницы от поисковых систем?

После релиза сайта его страницы сканируются роботами при первом обходе. Однако если дизайн и контент страниц пока ещё не оптимизированы для продвижения, рекомендуется на время доработок закрыть сайт от поисковиков. В каких ещё ситуациях нужен запрет на индексацию:

1. Создание мобильной версии на отдельном домене. В этом случае появляются дубли страниц, которые поисковики могут зафиксировать.
2. Тестирование сайта на другом домене. Если вы создали аналог основного сайта и проводите на нём тестовые работы, поисковые роботы могут воспринять страницы сайтов как дубликаты. В результате основной сайт может потерять позиции.
3. Смена дизайна, параметров, контента. На время работ по улучшению интерфейса и юзабилити закройте страницы сайта от индексации полностью или частично.



Что закрывать от индексации.

- Страницы сортировки.
- Страницы логина, авторизации, регистрации.
- Страницы без текстового наполнения.
- Корзину.
- Результаты поиска.
- Страницы тегов.



С помощью каких инструментов

Показываем боту нужные страницы:

- robots.txt;
- sitemap;
- canonical;
- перелинковка.

Запрещаем боту ненужные:

- meta robots

```
<meta name="robots" content="noindex, nofollow>
```

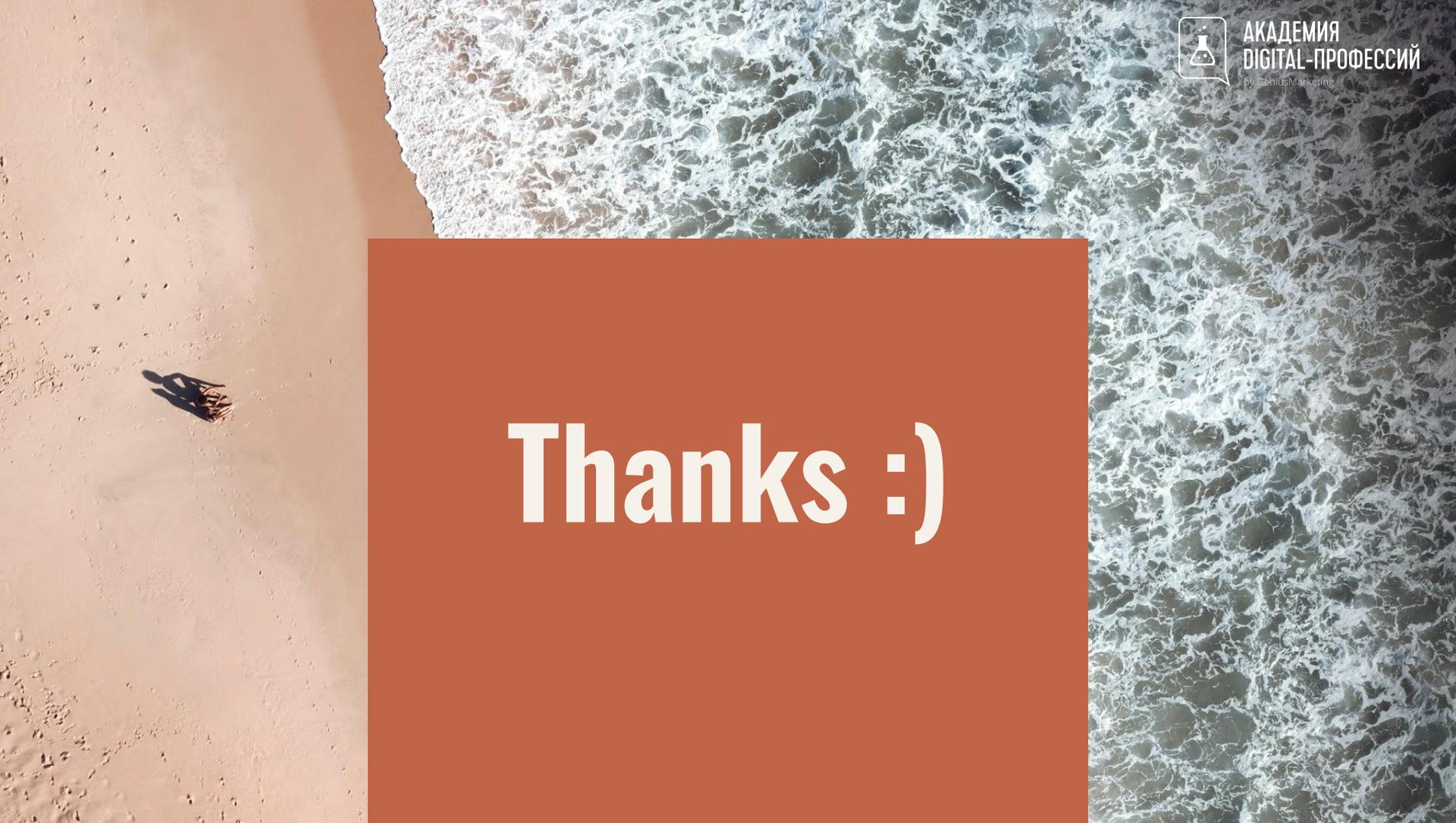
```
<meta name="robots" content="noindex, follow>;
```

- canonical;
- X-robots-tag;
- отдача боту 4xx код сервера.





АКАДЕМИЯ
DIGITAL-ПРОФЕССИЙ
by GeniusMarketing



Thanks :)