

Пример 05 / SEO-аудит

Большой каталог: дубли и ошибки сканирования

Публичная обезличенная версия исходного отчёта. Названия компаний, домены, контакты и персональные данные заменены. Содержание проверок и выявленные показатели сохранены. Отдельные приложения со списками URL не публикуются.

Приложение:

Задача 1. Пагинация ведёт на ?page={page}

Лист приложения: «1 Пагинация».

Проблема. В листингах (разделы, бренды, подборки) номера страниц и стрелки «назад» и «вперёд» ведут на адрес с неподставленным плейсхолдером `?page=%7Bpage%7D`, то есть `?page={page}`. Такой адрес отдаёт 200, canonical у него указывает на раздел. Рабочая ссылка дальше одна: кнопка «Показать ещё», она ведёт на следующую страницу.

Последствия. Страницы раздела связаны в одну цепочку: со страницы N можно перейти только на N+1. Чтобы дойти до 30-й страницы, робот должен пройти 29 предыдущих подряд, и один сбой обрывает цепочку. Так и случилось при крауле: по номерам последних страниц в 388 разделах около 5 234 страниц листинга, краулер нашёл 1 999. В разделе /katalog 1 006 страниц, краулер дошёл до 7-й. 10 370 страниц товаров и брендов оказались на глубине 5 кликов и больше, самые дальние на глубине 12.

Где. Шаблон пагинации темы Компания-пример: блок nav со ссылками `a.gp-pages__item` и `a.gp-pages__arrow` внутри `main > section`. Рабочая ссылка: `div.gp-results__more > a`. В OpenCart плейсхолдер `{page}` заменяется в `system/library/pagination.php` при рендере; судя по результату, шаблон темы выводит URL до замены.

Пример. На странице <https://site-05.example/katalog/sumki/ryukzaki?page=3> ссылки «2», «4», «33» и обе стрелки ведут на <https://site-05.example/katalog/sumki/ryukzaki?page=%7Bpage%7D>.

Что сделать.

- Подставлять номер страницы во все ссылки пагинации: номера, стрелки, последнюю страницу. Ссылка на первую страницу остаётся адресом раздела без `?page`, она и сейчас верная.
- Для `page` с нечисловым значением (включая `{page}`), нулём или номером больше последней страницы отдавать 404.
- Кнопку «Показать ещё» оставить как есть.

Как проверить.

- В HTML разделов нет строк `%7Bpage%7D` и `{page}`.
- Запрос <https://site-05.example/katalog/sumki/ryukzaki?page=%7Bpage%7D> отдаёт 404.
- Со страницы 3 раздела «Рюкзаки» ведут рабочие ссылки на страницы 2, 4 и 33.

Задача 2. Слеш на конце адреса даёт редирект через http

Лист приложения: «2 Слеш-редиректы».

Проблема. Адрес со слешем на конце проходит два редиректа, и промежуточный идёт по http:

https://site-05.example/blog/ → 301 → http://site-05.example/blog → 301 → https://site-05.example/blog

Так ведут себя все 23 адреса со слешем, которые нашёл краулер. На каждом таком переходе робот делает лишний запрос, причём по незащищённому протоколу.

Что сделать.

- Правило, которое убирает слеш, должно сразу отдавать https: один 301 на https://site-05.example/blog. Пример правила приведён ниже.
- Убрать слеш на конце из 32 ссылок, перечисленных в таблице. Адрес /product-category/categories/ остался от старого сайта, его заменить на /katalog.

Обычно http в заголовке Location появляется, когда Apache работает за nginx и считает запрос нешифрованным. Схему нужно задать в правиле явно, например в .htaccess:

```
RewriteCond %{REQUEST_FILENAME} !-d
RewriteRule ^(.+)/$ https://%{HTTP_HOST}/$1 [R=301,L]
```

Страница со ссылкой	Ссылки со слешем на конце
/brands (шапка и подвал), 21 ссылка	/akcii/, /blog/, /contact/, /dostavka-i-oplata/, /individualnoe-proizvodstvo/, /katalog/, /kto-nam-doveryaet/, /nashi-brendy/, /o-nas/, /originalnyj-stil/, /poligrafiya/, /populyarnye-voprosy/, /suveniry-dlya-vystavki/, /suveniry-pod-zakaz/, /tehtrebovaniya/, /tipy-nanesenij-logotipa/, /vidy-nanesenij/, /welcome-box/
Главная (ссылки-картинки в контенте)	/individualnoe-proizvodstvo/ezhednevniki-na-zakaz/, /suveniry-iz-kitaya/
/blog/brendirovannaya-produkciya-cto-eto-i-pochemu-ona-vazhna	/product-category/categories/ (4 ссылки)
/blog/vidy-podarocnoy-upakovki	/individualnoe-proizvodstvo/podarocnaya-upakovka-na-zakaz/, /individualnoe-proizvodstvo/podarki-iz-dereva/izgotovlenie-derevyannyh-shkatulok-yashchikov-i-korobov/
/blog/cto-takoe-tampopechat и /blog/tipy-nanesenij-logotipa/cto-takoe-tampopechat	/katalog/

Как проверить. Запрос https://site-05.example/blog/ отдаёт один 301 с Location: https://site-05.example/blog, следующий запрос отдаёт 200.

Задача 3. Старые разделы /gifts-ru/

Лист приложения: «3 gifts-ru».

Проблема. Два редиректа со старых адресов ведут на страницы 404:

Старый адрес	Сейчас	Ссылаются карточек	Правильная цель
/gifts-ru/odejda	301 на /catalog/odejda, там 404 (опечатка: catalog вместо katalog)	123	/katalog/odezhda
/gifts-ru/elektronika	301 на /catalog/elektronika, там 404	51	/katalog/tehnologii (раздел «Электроника»)

Ещё 61 раздел старого дерева /gifts-ru/ отдаёт 200 и дублирует основной каталог. Например, у /gifts-ru/sportivnye-tovary/massajery и /katalog/dlya-sporta/massazheri-2 одинаковый title, а /gifts-ru/novyy-god с title «Корпоративные подарки на Новый год» повторяет новогоднюю посадочную. На старые разделы ведут 2 674 ссылки с 1 221 карточки товара.

Где. Блок категорий в карточке товара (main > section > ul > li > a) у товаров, импортированных от Gifts. Анкоры: «Промо одежда», «Электроника и гаджеты», «Дом», «Отдых» и другие.

Что сделать.

- Исправить цели двух редиректов по таблице выше.
- В блоке категорий карточки выводить только разделы основного каталога.
- Снять с публикации разделы /gifts-ru/ и настроить с каждого 301 на соответствующий раздел каталога. В листе «3 gifts-ru» для 4 разделов кандидат найден по совпадению H1, у двух из них кандидатов два. Для остальных 57 цель согласовать с SEO до внедрения.

Как проверить. Повторный краул не находит адресов /gifts-ru/ с кодом 200, оба редиректа из таблицы ведут на страницы с кодом 200.

Задача 4. robots.txt закрывает товары по ошибке

Лист приложения: «4 robots-товары».

Проблема. Правила Disallow: /*cart (строка 460) и Disallow: /*route* (строка 445) срабатывают на любой адрес, где встречается это сочетание букв. Под них попали 20 товаров, страница бренда Routemark и одна картинка:

- /*cart: ручки и карандаши Carton (ruchka-sharikovaya-carton-plus-*, vechnyiy-karandash-carton-inkless-*), набор карандашей Cartoon, часы Carte Blanche, термкружки Decart, футболка Logicart, картинка cartholder-vinil-1060x723.webp;
- /*route*: дорожные сумки, папка и кошелёк Route, страница бренда /routemark.

Что сделать. Сузить правила, чтобы они закрывали только служебные адреса OpenCart:

вместо Disallow: /*route*

Disallow: /*route=

вместо Disallow: /*cart

Disallow: /cart\$

Disallow: /cart?

Задача 5. Битые ссылки на удалённые товары и /blog/null

Лист приложения: «5 Битые ссылки».

Проблема. На сайте 160 внутренних ссылок на 6 адресов с кодом 404:

Адрес (404)	Страниц со ссылкой	Где стоит ссылка
/bokal-dlya-martini-bistro	68	Карусель похожих товаров в карточках (div.gp-carousel__track article.gp-tile) и листинг раздела «Аксессуары для вина»
/vizitnitsa-enamel-rozovaya	3	Листинги визитниц в /vip/lichnie-aksessuari/vizitnici; у ссылки текст другого товара (визитница Pierre Cardin)
/bloknot-real-hard-work	2	Листинги: раздел «Для нефтяников и шахтёров» ?page=2, бренд /hard-work ?page=2
/zont-trost-alessio-cherniy	2	Листинги: раздел «Зонты-трости» ?page=6, бренд /matteo-tantini ?page=2
/korobka-my-warm-box-temno-sinyaya	2	Листинги: раздел «Подарочная упаковка» ?page=4, подборка /sdelano-v-rossii ?page=7
/blog/null	4	Картинки с src="null" в теле 4 статей блога

Ещё два редиректа ведут на 404, они описаны в задаче 3.

Что сделать.

- Не выводить удалённые и отключённые товары в листингах и каруселях. Проверить, что кэш листингов и связи «похожие товары» обновляются после импорта.
- С адресов товаров, снятых с продажи насовсем, ставить 301 на родительский раздел.

Как проверить. Повторный краул не находит внутренних ссылок на страницы с кодом 404.

Задача 6. Картинки с кириллицей в имени файла отдают 404

Лист приложения: «6 Картинки 404».

Проблема. 51 картинка из кэша WebP отдаёт 404. У всех в имени исходного файла есть кириллица «гб», например /image/cachewebp/catalog/gifts/products/7537.44.8гб_1_1000x1000-456x456.webp. Это флешки Twist Color на 8 и 16 ГБ: на 20 страницах (карточки и листинг флешек) у них нет фото.

Что сделать. Переименовать исходные файлы латиницей, обновить пути в базе и транслитерировать имена файлов картинок при импорте.

Задача 7. Статьи блога открываются по двум адресам

Лист приложения: «7 Дубли статей».

Проблема. 8 статей отдают 200 по двум адресам, у каждой копии canonical указывает на неё саму. Хаб /tipy-nanesenij-logotipa и раздел /akcii ссылаются на адреса с разделом в пути, HTML-карта /sitemap ссылается на короткие.

Оставить	Перенаправить 301
/blog/tipy-nanesenij-logotipa/cto-takoe-dtf-pechat-i-zachem-ona-nuzhna	/blog/cto-takoe-dtf-pechat-i-zachem-ona-nuzhna
/blog/tipy-nanesenij-logotipa/cto-takoe-lazernaya-gravirovka	/blog/cto-takoe-lazernaya-gravirovka
/blog/tipy-nanesenij-logotipa/cto-takoe-sublimacionnaya-pechat	/blog/cto-takoe-sublimacionnaya-pechat
/blog/tipy-nanesenij-logotipa/cto-takoe-uf-pechat	/blog/cto-takoe-uf-pechat
/blog/tipy-nanesenij-logotipa/cto-takoe-tampopecthat	/blog/cto-takoe-tampopecthat
/blog/tipy-nanesenij-logotipa/cto-takoe-shelkografiya	/blog/cto-takoe-shelkografiya
/blog/akcii/gornolyzhnaya-akciya-kant	/blog/gornolyzhnaya-akciya-kant
/blog/akcii/suveniry-s-besplatnym-naneseniem	/blog/suveniry-s-besplatnym-naneseniem

Что сделать.

- Настроить 301 по таблице. Оставляем адреса, на которые ссылаются хаб и раздел акций.
- В HTML-карте /sitemap и в sitemap.xml ставить оставленные адреса.
- Чтобы дубли не появлялись снова, модуль блога должен отдавать статью только по одному адресу, по её основному разделу.

Как проверить. <https://site-05.example/blog/cto-takoe-tampopecthat> отдаёт 301 на <https://site-05.example/blog/tipy-nanesenij-logotipa/cto-takoe-tampopecthat>.

Задача 8. Разделы-дубли

Листы приложения: «8 Дубли title разделов», «8б Одноимённые H1 разделов».

Проблема. 7 пар разделов с одинаковым title, а раздел «Футболки» есть в «Одежде» дважды:

Раздел	Адрес 1	Адрес 2
Колорит (Color it)	/katalog/podarochnie-nabori-10/nabori-color-it	/katalog/kollekcii/serii/color-it
Наборы сотрудника (Welcome pack)	/katalog/podarochnie-nabori-10/nabori-welcome-pack	/katalog/kollekcii/serii/welcomepack

Раздел	Адрес 1	Адрес 2
Массажеры	/katalog/dlya-sporta/massazheri-2	/gifts-ru/sportivnye-tovary/massajery (уходит по задаче 3)
Пылесосы	/katalog/tehnologii/pilesosi-2	/katalog/avtomobilnye-aksessuary/dlya-salona/pilesosi
Ремни	/katalog/muzhskie-aksessuary/odezhda-2/remni-2	/vip/lichnie-aksessuari/muzhskie-aksessuari/remni
Сумки на пояс	/katalog/dlya-sporta/sumki-dlya-sporta/sumki-na-poyas-2	/vip/portfeli-i-sumki/sumki-na-poyas-3
Халаты	/katalog/odezhda/halati-3	/vip/lichnie-aksessuari/bannii-prinadlezhnosti/halati-2
Футболки	/katalog/odezhda/futbolki (id 688)	/katalog/odezhda/futbolki-2 (id 1233)

Кроме того, у 154 разделов H1 совпадает с H1 раздела из другой ветки: «Вазы», «Зонты», «Карандаши», «Книги» и другие.

Что сделать.

- Если у пары один и тот же набор товаров (вероятно, Color it, Welcome pack и «Футболки»), оставить один раздел, второй убрать из меню и перенаправить 301 на первый.
- Для одноимённых разделов в разных ветках добавлять в title название родительского раздела или ветки, например «Ремни VIP».

Задача 9. Шаблоны title, H1 и description товаров

Лист приложения: «9 Дубли title товаров».

Проблема.

- У 999 товаров в 415 группах одинаковый title. Это варианты цвета или размера с общим названием: например, 13 карточек «Косметичка с возможностью полноцветной печати». H1 и description у них тоже совпадают.

Что сделать. Если название товара не уникально, добавлять в title, H1 и description различитель: цвет, размер или артикул («..., синий, арт. 20411.30»).

Задача 10. Страницы без description

Лист приложения: «10 Без description».

Проблема. У 181 страницы нет meta description: 176 страниц брендов (/sols, /prodir, /parker, /hugo-boss и др.), список /brands и служебные /privacy, /consent, /cookie-consent. Title у страниц брендов состоит из одного названия бренда: «Sol's», «Prodir».

Что сделать. Добавить шаблоны для страниц брендов, например title «{Бренд}: товары с логотипом | Компания-пример» и description «{Бренд}: товары с нанесением логотипа оптом. Каталог, цены и расчёт тиража в Компания-пример.»

Задача 11. 70 брендов закрыты в robots.txt

Лист приложения: «11 robots-бренды».

Проблема. 70 страниц брендов закрыты строками 487-556 вида Disallow: /victorinox\$. Закрытые страницы не попадают в индекс, а список брендов /brands продолжает на них ссылаться.

Что сделать. Удалить строки брендов из robots.txt. Бренды и номера строк перечислены в листе приложения.

Задача 12. Ссылки и картинки поставщика в описаниях товаров

Листы приложения: «12 Ссылки на external-02.example», «126 Внешние картинки».

Проблема.

- В описаниях 186 товаров стоят 275 ссылок на сайт поставщика external-02.example: на его карточки, страницы брендов и поиск. 154 из них без nofollow. Пример:
`https://site-05.example/beysbolka-shade-up-oranjevaya` ссылается на `https://external-02.example/id/34830`.
- Таблицы размеров в описаниях одежды подгружаются картинками с external-03.example и external-04.example: 21 картинка, около 1 900 вставок.

Что сделать.

- При импорте вырезать из описаний ссылки на external-02.example, оставляя текст ссылки, и вычистить уже загруженные описания.
- Скачать таблицы размеров на свой сервер и заменить src.

Задача 13. Нет alt у миниатюр галереи товара.

Проблема. В карточках 185 460 вставок картинок без alt, почти все это миниатюры галереи 150×150. Основное фото и картинки в листингах alt имеют.

Что сделать. В шаблоне карточки выводить alt у миниатюр: «{Название товара}, фото {N}».

Задача 14. Страница /brands на стандартном шаблоне OpenCart

Проблема. Список производителей /brands свёрстан стандартным шаблоном OpenCart (div#product-manufacturer, колонки col-sm-3), а не темой сайта. Отсюда другая шапка и подвал с 21 ссылкой со слешем (задача 2), пустой атрибут lang, нет description, title «Производители». На /brands ведут хлебные крошки «Бренды» на 451 странице брендов, а в меню стоит /nashi-brendy («Наши бренды»). Получается две страницы с одной задачей.

Что сделать. Оставить /nashi-brendy, с /brands поставить 301, ссылку «Бренды» в хлебных крошках вести на /nashi-brendy. Если /brands нужен, перевести его на шаблон темы.

Задача 15. Адреса товаров со спецсимволами

Лист приложения: «14 URL со спецсимволами».

Проблема. Генератор адресов переносит в URL лишние символы:

- 159 адресов с «quot» и «039» из HTML-сущностей " и ', например /ryukzak-doss-dlya-noutbuka-156quot;
- 110 адресов с символом ², например ...-plotnostyu-320-g-m²-...;

- 18 адресов с неразрывными пробелами (%C2%A0), например /mediapleer%C2%A0%C2%A0media%C2%A0atv%C2%A04k;
- 3 адреса с буквами Š и í.

Ещё у 12 брендов в адресе подчёркивание (/paul_bocuse). Эти адреса можно не менять.

Что сделать.

- В генераторе адресов перед транслитерацией декодировать HTML-сущности, удалять кавычки и спецсимволы, неразрывные пробелы заменять дефисом. В адресе оставлять только a-z, 0-9 и дефис.
- Регенерировать около 290 существующих адресов, со старых поставить 301.

Задача 16. Мелкие правки

Лист приложения: «16 Прочее».

- /contact: title «Контакты для нанесения логотипа - купить в Компания-пример» и description по шаблону раздела. Задать для страницы отдельные мета-теги в коде (повтор п. 9 аудита 08.10).
- Иконка входа в шапке (/login) без текста: добавить aria-label="Вход".
- Ссылки на соцсети в шапке и подвале ведут на редиректы: <https://social.example/community> на <https://social.example/community> на <https://social.example/community>. Поставить конечные адреса.
- 5 300 адресов картинок с двойным слешем, например /image/cachewebp/catalog/oasis/products//7fb8d783ffe6509f-800x800.webp. Картинки открываются, но путь собран без папки товара.

О публичном примере

Условные адреса *.example сохраняют технический смысл примеров, но не ведут на сайты клиентов. Внешние ссылки на рабочие таблицы удалены; сами приложения в этот PDF не включены. Это исторический отчёт, а не утверждение о текущем состоянии какого-либо сайта.