

Фирменная методика Prodactor: как мы исследуем нишевые сайты в ТОПе Яндекса

Оглавление

Фирменная методика Prodactor: как мы исследуем нишевые сайты в ТОПе Яндекса.....	1
Введение.....	1
Зачем мы это делаем.....	1
Как мы отбирали сайты.....	2
Как мы сохраняли сайты.....	3
Как мы обрабатывали сайты.....	3
Как мы анализировали данные.....	4
Что дальше.....	5
Открытые данные и приглашение.....	5

Введение

В информации про сайты для местного бизнеса полно мифов. Наверное, вы их слышали: «на сайте обязательно должен быть слайдер с фотографиями», «нужен онлайн-консультант, который встречает посетителя вопросом», «главная страница должна быть длинной и подробной». Эти убеждения кочуют из статьи в статью, из брифа в бриф. Но что, если проверить их не на словах, а на деле? Что, если взять не абстрактные советы, а реальные сайты, которые уже победили в конкурентной борьбе?

Мы решили не гадать, а исследовать. Прежде чем создавать готовые сайты для местного бизнеса, мы захотели досконально понять, как устроены успешные сайты, попавшие в ТОП Яндекса в органической выдаче. Для этого мы взяли конкретную нишу (в качестве первого примера — кафе), отобрали несколько десятков сайтов в городах-миллионниках, разобрали их до винтика и выяснили, что на самом деле работает.

В этой статье мы публично раскрываем свою методологию. Почему мы отбираем сайты именно так, как их сохраняем, как обрабатываем с помощью искусственного интеллекта и почему в итоге работаем только с лучшими сайтами — с теми, что попадают в ТОП.

Зачем мы это делаем

Наша логика проста и прагматична. Если сайт попал органическим способом на первую страницу Яндекса по конкурентному запросу в городе-миллионнике — значит, он что-то делает правильно. Поисковая система уже отобрала лучших. У них есть чему поучиться.

Простыми словами: мы не хотим изобретать велосипед. Мы хотим посмотреть на те велосипеды, которые едут быстрее всех, и понять, как они устроены. Это не копирование, это исследование лучших практик.

Технические детали: мы исходим из гипотезы, что позиции в органической выдаче коррелируют с качеством сайта с точки зрения пользователей и поисковых алгоритмов. Разумеется, мы исключаем факторы возраста домена и накопившейся за годы популярности — такие сайты мы отсеиваем на этапе отбора. Нас интересуют только современные решения, которые работают здесь и сейчас.

Цель исследования — не скопировать чужие сайты, а выявить глубинные закономерности. Что объединяет сайты из ТОПа Яндекса в конкретной нише? Какие страницы есть у всех? Какие функции встречаются чаще других? А что, вопреки стереотипам, вообще никто не использует?

Методика универсальна. Мы применяем её к любой нише местного бизнеса: кафе, автосервисы, салоны красоты, клининговые компании, стоматологии и другие. Результатами мы делимся открыто. Пусть пользуются все — владельцы бизнеса, маркетологи, веб-разработчики. Нам не жалко.

Как мы отбирали сайты

Самое важное в любом исследовании — чистота выборки. Если намешать всего подряд, результаты будут бессмысленными. Поэтому мы ввели жёсткие критерии отбора и безжалостно отсеивали всё, что могло исказить картину.

Простыми словами: мы не берём всё подряд. Мы отбираем только те сайты, которые действительно интересны владельцу обычного городского бизнеса в исследуемой нише. Никаких сетей, никаких точек при чужих заведениях, никаких одностраничников.

Технические детали: в выборку сознательно не попадают несколько категорий сайтов.

- **Одностраничники.** Это не сайты в полном смысле слова, а скорее визитки. Нам этот сегмент неинтересен, потому что он не отражает реальные возможности веб-разработки для бизнеса.
- **Сайты сетевых заведений.** У крупных сетей другие бюджеты, другие задачи, другие команды. Они не ориентир для обычного бизнеса.
- **Точки при торговых центрах, отелях, фитнес-клубах и других заведениях.** У таких проектов специфический трафик — они работают на случайных посетителях ТЦ или гостях отеля, и их сайты решают другие задачи. Нам нужны самостоятельные предприятия.
- **«Динозавры».** Сайты, которые попали в ТОП только за счёт возраста и накопленной популярности, но при этом выглядят как привет из прошлого. Мы ищем современные решения, которые работают сегодня.
- **Города вне списка.** Мы ограничиваемся миллионниками, чтобы выборка была однородной по уровню конкуренции. Сайт из города с населением 50–100 тысяч человек не сопоставим с сайтом из города-миллионника.

После всех фильтров у нас остаётся несколько десятков сайтов, которые и ложатся в основу исследования. Все они отобраны по единым правилам:

- **Города:** Новосибирск, Екатеринбург, Казань, Красноярск, Нижний Новгород, Челябинск, Омск, Самара, Уфа, Ростов-на-Дону, Воронеж, Волгоград, Пермь (список может дополняться в зависимости от ниши).
- **Поисковый запрос:** максимально широкий запрос по нише (например, для кафе — «кафе сайт», для салонов красоты — «салон красоты», для автосервисов — «автосервис сайт»).
- **Глубина:** только первая страница органической выдачи Яндекса. Никакой рекламы, только прямые ссылки на сайты. Для каждого города получается собрать от 3 до 6 ссылок.

Как мы сохраняли сайты

Когда несколько десятков кандидатов отобраны, встаёт технический вопрос: как зафиксировать их состояние, чтобы можно было спокойно анализировать, не боясь, что владелец сайта завтра всё переделает? Живые сайты меняются постоянно: правки вносятся ежедневно, дизайны

обновляются, контент добавляется и удаляется. Нам нужен «слепок реальности» — точная копия каждого сайта на момент исследования.

Простыми словами: мы скачиваем все отображенные сайты целиком, чтобы они не изменились, пока мы их изучаем. После этого у нас появляется архив, к которому можно вернуться через месяц или через год.

Технические детали: для сохранения мы используем утилиту `single-file`. Это инструмент, который позволяет сохранить список веб-страниц сайта в набор файлов вместе с текстом и структурой.

Команда выглядит так:

```
single-file --urls-file "C:\Prodactor\download\links.txt" --output-directory "C:\Prodactor\download" --browser-wait-until=load --browser-load-max-time=30000 --block-stylesheets --block-images --remove-scripts
```

Где `C:\Prodactor\download\` — папка для сохранения файлов сайта.

Разберём ключевые параметры, потому что они важны для понимания чистоты эксперимента.

`--browser-wait-until=load` и `--browser-load-max-time=30000` говорят утилите: не спеши, подожди до 30 секунд, пока сайт полностью загрузится, выполни весь JavaScript, покажи весь динамический контент. Это гарантирует, что мы сохраняем сайт именно в том виде, в котором его видит обычный посетитель.

`--block-stylesheets` и `--block-images` отключают сохранение внешних стилей и изображений. Нам важна структура, тексты, ссылки, формы, а не картинки и красоты, которые занимают много места. В итоговом файле остаётся чистая и лёгкая суть сайта.

`--remove-scripts` удаляет исполняемые скрипты. Это делает сохранённый файл абсолютно безопасным и статичным — он больше не пытается никуда отправить запросы, ничего не грузит, просто лежит в архиве, идеальный для анализа.

На выходе мы получаем несколько десятков компактных наборов файлов, каждый из которых содержит полную структуру сайта со всеми текстами, но без картинок, стилей и тяжеловесных скриптов. Причём все файлы получены практически одновременно, что позволяет говорить о снимке состояния исследуемых сайтов.

Как мы обрабатывали сайты

Следующий этап — самый трудоёмкий. Иметь несколько десятков папок с файлами мало, нужно превратить их в структурированные данные, пригодные для подсчётов и сравнения. Делать это вручную — недели кропотливой работы. Здесь нам на помощь приходит искусственный интеллект.

Простыми словами: мы научили нейросеть разбирать сайты по полочкам и заполнять анкеты (Описи). Получается быстро, единообразно и без ошибок.

Технические детали: для каждого сохранённого набора файлов мы запускаем ИИ со специально разработанным промптом. Промпт содержит подробную инструкцию: что смотреть, как группировать, какие разделы заполнять. На выходе нейросеть выдаёт структурированный документ — мы назвали его **Опись сайта** — по одному для всех шаблону.

Вот что входит в Опись:

- **Раздел 0. Общие данные** — название компании, город, профиль деятельности, адрес сайта, дата сбора.
- **Раздел 1. Все страницы сайта** — полный список всех страниц.
- **Раздел 2. Шапка сайта** — какие ссылки и семантические блоки используются в шапке.
- **Раздел 3. Футер сайта** — та же информация для подвала.

- **Раздел 4. Главная страница** — все ссылки с главной, смысловые блоки, визуальные элементы.
- **Раздел 5. Функциональность** — все формы обратной связи, онлайн-запись, корзина, интерактивная карта и другие функции, которые исполняет сайт.
- **Раздел 6. Товары (услуги) и цены** — есть ли описания, указаны ли цены, их точность и актуальность.
- **Раздел 7. Внешние ссылки** — социальные сети, ссылки на внешние ресурсы.
- **Раздел 8. Обратная связь** — телефоны, email, мессенджеры, формы обратной связи, отзывы, голосования и другие способы связи с пользователями.
- **Раздел 9. Юридическая информация** — ИНН/ОГРН владельца сайта, политика конфиденциальности, соглашение о персональных данных, лицензии, оферты и прочая юридическая информация.
- **Раздел 10. SEO-данные** — ключевые слова, метатеги, структура текстов, URL, перелинковка.

Важное уточнение: описи — наш внутренний рабочий инструмент. Мы их не публикуем и никому не показываем.

Как мы анализировали данные

Когда все описи готовы, начинается математика (статистика). Мы перестаём быть исследователями и превращаемся в счетоводов. Задача проста и одновременно сложна: подсчитать, сколько раз встречается каждая страница, каждый блок, каждая функция, каждое ключевое слово.

Простыми словами: мы просто считаем, что встречается часто, а что — редко. И оказывается, что реальность сильно отличается от стереотипов.

Технические детали: для каждого элемента мы фиксируем частотность. Например (на примере ниши кафе), страница «Контакты» оказалась у всех сайтов — 100% попадание. Страница «Меню» — у большинства. А вот страница с подробным описанием отдельного блюда — только у каждого седьмого. Это уже пища для размышлений.

Но самое главное: после подсчёта мы работаем не со всей массой данных, а только с лидерами. Единичные случаи нас не интересуют. Если какой-то элемент встретился один-два раза — это случайность, а не закономерность. Мы выделяем **ТОП** — самые частотные позиции — и в дальнейшем оперируем только ими.

Примеры того, как это выглядит:

- В шапке сайта мы выделяем ТОП-9 ссылок и ТОП-6 блоков.
- На главной странице — ТОП-11 ссылок, ТОП-7 смысловых блоков, ТОП-4 визуальных блока.
- В футере — ТОП-8 ссылок и ТОП-7 блоков.
- В функциональности — ТОП-4 функции.

Этот подход — работа только с лидерами, только с тем, что встречается у многих — позволяет нам отсеять шум и увидеть реальную картину. Не «нам кажется», не «дизайнер так захотел», а объективные данные «как у большинства», основанные на анализе лучших.

Что дальше

Вся эта методология — не самоцель. Она нужна, чтобы отвечать на главные вопросы: каким должен быть идеальный сайт в конкретной нише? Какие страницы обязательны, а без каких можно обойтись? Что реально работает в SEO, а что — пустая трата времени?

Ответы на эти вопросы — в наших исследованиях по конкретным нишам (кафе, салоны красоты, автосервисы, клининг и другие).

В каждом таком исследовании мы показываем:

- ТОП обязательных страниц.
- Что на самом деле должно быть в шапке и футере.
- Какие функции нужны, а какие — лишние.
- Как обстоят дела с ключевыми словами, метатегами, URL и перелинковкой у лидеров.

Открытые данные и приглашение

Мы не прячем свои исследования. Если вы предприниматель, маркетолог или веб-разработчик — берите нашу методологию, проводите свои анализы. Проверяйте свои гипотезы, смотрите, как устроены конкуренты в вашей нише.

Если хотите, чтобы мы исследовали вашу нишу — пишите, обсудим. Мы открыты к сотрудничеству.

[Ссылка на архив.](#)

Дата публикации: апрель 2026 г.