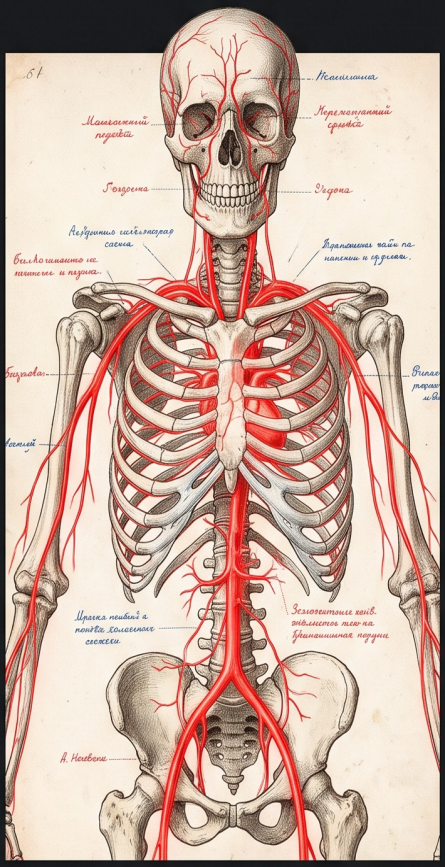
ГУГЛ В КРОВИ. АНАТОМИЯ ТОПА

Как устроен сайт, который попадает на первую страницу

А. Нехаев



ГУГЛ В КРОВИ. АНАТОМИЯ ТОПА

Как устроен сайт, который попадает на первую страницу — и что делать, если ваш лежит под ножом алгоритма

Автор: А. Нехаев
Бренд: Гугл в крови
Платформа: Литрес + Ozon
Версия: финальная версия v1.0, сентябрь 2026

ПРЕДИСЛОВИЕ. Почему анатомия, а не магия

SEO в 2026 — это не ритуал. Не заклинание. Не карго-культ «секретных методов», которые гуру продают на вебинарах за 49 900.

SEO — это тело. У него есть кости (техническое), мышцы (контент), сухожилия (ссылки), нервы (поведенческие), кожа (сниппеты), кровь (коммерция). Когда хирург оперирует, он сначала понимает анатомию. А потом режет. Не наоборот.

Эта книга — анатомический атлас. Не учебник по болезням и не список операций. Атлас: где что лежит, как связано, что без чего не работает.

Что вы получите:
- Язык, на котором можно думать о SEO без мистики
- Систему координат для диагностики сайта
- Операционные схемы для практиков
- Переводчик с «SEO-жаргона» на язык бизнеса — для владельцев

Что вы НЕ получите:
- Кнопку «в ТОП за неделю»

- Список из 300 каталогов, где разместиться
- Секретный метод, о котором не знает Google
- Мотивационную клюкву про «веру в успех»

Как читать:
- Вы SEO-практик — читайте подряд. Главы 1–8 — ваш операционный стол. Глава 9 — чек-лист аудита. Глава 10 — для случаев, когда нужно объяснить клиенту, почему «просто купить ссылок» не работает.
- Вы владелец бизнеса — читайте вступление, концовки глав («Владельцу на салфетке») и Главу 10. Остальное — справочно, когда подрядчик начнёт говорить на птичьем языке.

В путь.

ВСТУПЛЕНИЕ ДЛЯ ВЛАДЕЛЬЦЕВ. Если ты не SEO

Вы платите за SEO. Возможно, давно. Возможно, не видите результата. Возможно, отчёт подрядчика выглядит как китайская грамота: «проиндексировано 87 432 страницы», «получено 14 392 backlink», «позиции выросли на 23%».

И что с этим делать — непонятно.

Эта книга не сделает вас сеошником. Но она даст три вещи:

1. Язык. Вы сможете отличить «купили 50 ссылок на бирже за 5 000» от «восстановили ссылочный профиль после спам-атаки». На бытовом уровне — но достаточно, чтобы не платить за воздух.
2. Глаза. Вы увидите в отчёте подрядчика не цифры, а процессы. «Позиции выросли» — растущие позиции куда? По каким запросам? Сколько из них — коммерческие? Какой трафик это дало в деньгах?
3. Красные флаги. 7 признаков того, что подрядчик не хирург, а шаман с бубном. Они — в Главе 10. Если в вашем подрядчике 3 из

7 — бегите.

Что эта книга НЕ даст:

- Гарантию результата. SEO — это не покупка рекламы, где вы платите за клики. Это инвестиция в актив, который растёт со временем.
- Универсального решения. Каждая ниша, каждый сайт, каждая стадия развития требует своей операции.
- Замену подрядчику. Хороший хирург нужен. Эта книга — чтобы вы могли проверить его скальпель.

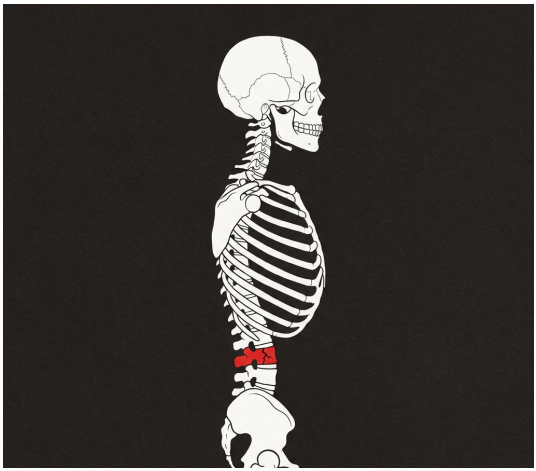
Семь красных флагов в отчёте подрядчика (расширенная версия — в Главе 10):

1. Отчёт заканчивается позициями, не доходит до трафика и денег
2. Все запросы в топ-10 — информационные, коммерческих нет
3. «Купили X ссылок» — без указания доноров и анкоров
4. Нет данных по индексации и техническим ошибкам
5. Рост позиций, но падение трафика (значит, сменили семантику)
6. Отчёт только по Яндексу или только по Google (без объяснения почему)
7. Подрядчик сам не может объяснить, что он сделал за месяц

Если хотя бы 3 из 7 про вашего подрядчика — эта книга для вас.

ЧАСТЬ I. ФУНДАМЕНТ ТЕЛА

[!] ГЛАВА 1. СКЕЛЕТ БЕЗ ТРЕЩИН. Техническое SEO как несущая конструкция



Без скелета нет тела. Трещина в позвоночнике — и мышцы висят, нервы не доходят до мозга, кровь течёт мимо органов. В SEO то же самое. Технические ошибки обрушивают всё остальное: контент, ссылки, поведенческие — всё становится бессмысленным, если Google не может корректно прочитать ваш сайт. Эта глава — про кости.

1.1. Индексация: что Google видит и почему 40% страниц не попадают в индекс

Цифра в заголовке не из воздуха. Аудит Avtoservis, начало 2024: 47% карточек товаров не были в индексе Google. Половина каталога. Семь тысяч страниц с уникальным контентом, фото, ценами —

мёртвый груз, который не приносил ни клиента, ни рубля. Причина — криво настроенный шаблон: в карточке товара случайно генерировался <meta name="robots" content="noindex">, потому что разработчик копировал шаблон для служебных страниц и забыл снять галку.

Это не уникальный случай. Это массовая болезнь.

Пять причин, почему страницы не попадают в индекс

Причина	Как диагностировать	Цена ошибки
noindex на уровне страницы или	Search Console → Pages → «Excl	Все страницы с этим шаблоном
robots.txt блокирует сканирова	Search Console → Pages → «Bloc	Google даже не знает, что эти
Canonical на несуществующий UR	Screaming Frog → Canonical → S	PageRank утекает в никуда
Дубликаты без canonical	Crawl-анализ → одинаковый titl	Google сам выбирает, какую вер
Soft 404 (страница существует,	Search Console → Pages → «Soft	Сканирование тратится на пусто

[x] Анти-паттерн: «Закрыли в robots.txt, чтобы не индексировалось»

Регулярный кейс от junior-сеошников. robots.txt не запрещает индексацию — он запрещает сканирование. Если Google знает о странице (например, по внешней ссылке), он её проиндексирует без сканирования. Результат: страница в индексе, без title, без description, без контента. Хуже, чем не было.

Правило: запрет индексации = <meta name="robots" content="noindex">. Запрет сканирования = robots.txt. Это разные инструменты для разных задач.

Что должно быть в индексе, а что нет

Тип страницы	В индексе?	Инструмент
Карточка товара	[v] Да	Canonical на основную версию
Категория	[v] Да	Canonical на себя, noindex тол
Результаты фильтрации (цвет, р	[!] Зависит	Если даёт уникальный контент и
Страница поиска по сайту	[x] Нет	noindex, follow
Корзина, личный кабинет	[x] Нет	noindex, nofollow

Служебные (404, 500, политика	[!] Зависит	404 — noindex; политика — да,
Пагинация (?page=2)	[!] Зависит	Canonical на первую страницу и

1.2. Robots.txt и Sitemap: анатомия карты скелета

Robots.txt — это не «магический файл», который говорит Google «не лезь». Это карта скелета: какие кости трогать, какие нет. Файл лежит в корне сайта, читается один раз за сканирование, имеет жёсткий лимит — обычно обрабатываются первые 500-1000 директив.

Рабочий robots.txt для крупного каталога

```
# Главное правило: разрешаем всё, закрываем точечно
```

[x] Анти-паттерн: закрыть в robots.txt пол-сайта «на всякий случай»

Самая частая ошибка начинающих. Видел сайты, где в robots.txt закрыт /blog/ (потому что «там дубли») или /news/ (потому что «мало текста»). Google никогда не узнает об этих разделах — и они не индексируются. Через полгода владелец удивляется: «почему статьи не в поиске?»

Правило: если страница нужна в поиске — она должна быть доступна для сканирования. Закрывайте в robots.txt только:

1. Технические дубли (сортировки, фильтры без уникального контента)
2. Служебные разделы (корзина, кабинет, результаты поиска)
3. URL с параметрами аналитики и сессий

Sitemap.xml: перечень страниц, а не гарантия индексации

Sitemap не гарантирует, что страница попадёт в индекс. Он только сообщает о её существовании. Если страница из sitemap закрыта в robots.txt — Google её не просканирует. Если на ней noindex — Google

её не проиндексирует. Если она отдаёт 404 — Google её выкинет.

Для крупного каталога разбивайте sitemap по логическим разделам:

```
sitemap.xml          → индексный файл со ссылками на остальные
```

Лимит — 50 000 URL или 50 МБ на один sitemap-файл. Превысили — разбивайте.

1.3. Core Web Vitals 2026: что изменилось после отмены FID

В марте 2024 Google официально отключил FID (First Input Delay) и заменил его на INP (Interaction to Next Paint). Это не косметическая правка — это смена философии. FID измерял только первую задержку отклика, INP — все взаимодействия на странице, до ухода пользователя.

Метрики и пороги в 2026

Метрика	Что измеряет	Хорошо	Плохо	Инструмент
LCP (Largest Contentful Paint)	Скорость загрузки главного контента	≤ 2.5 сек	> 4.0 сек	CrUX, PageSpeed Insights
INP (Interaction to Next Paint)	Задержка отклика на взаимодействие	≤ 200 мс	> 500 мс	CrUX, реальные пользователи
CLS (Cumulative Layout Shift)	Визуальная стабильность	≤ 0.1	> 0.25	CrUX

Почему «у нас всё зелёное» — не гарантия

PageSpeed Insights показывает лабораторные данные (Lighthouse на эмуляции мобильного). CrUX — реальные пользователи (Chrome Users Experience). Между ними пропасть. Лаборатория показывает 95, а CrUX по тому же URL — 47 на мобильных. Причина: лаборатория тестирует на быстром интернете и новом устройстве, реальные пользователи сидят на 3G и трёхлетних телефонах.

Что делать:
1. Смотрите CrUX в Search Console → Experience → Core Web Vitals. Это правда.
2. Если 75% URL в зелёной зоне CrUX — ок. Если нет —

оптимизируйте.
3. Лабораторию используйте для поиска причин, а не для оценки результата.

INP: что болит и как лечить

INP — самая проблемная метрика для сайтов с тяжёлым JavaScript. Типичные виновники:

- Длинные задачи JavaScript (main thread блокируется > 50 мс) — scheduler.yield() или MessageChannel.postMessage() для разбивки
- Огромные обработчики событий — debounce, throttling
- Тяжёлый hydration в React/Vue — переход на SSR/SSG/Partial Hydration
- Сторонние скрипты (аналитика, чаты, реклама) — аудит и удаление неиспользуемых

[x] Анти-паттерн: «Поставили Cloudflare — и CWV зелёные»

CDN помогает с LCP (контент ближе к пользователю), но не лечит:
- Долгий TTFB вашего origin-сервера
- Тяжёлый JavaScript на клиенте
- Блокирующий рендеринг CSS
- Картинки без lazy-loading и в правильном формате

CDN — это бинт. Лечить надо саму рану.

1.4. Мобильная версия и mobile-first: где ломается позвоночник

Google официально перешёл на mobile-first indexing ещё в 2023 году. Это значит: для ранжирования используется мобильная версия вашего сайта. Если её нет или она хуже десктопной — вы в проигрыше.

Три варианта мобильной версии и что с ними не так

Вариант	Плюсы	Минусы
Адаптивный дизайн (один URL, C	Google рекомендует. Один URL —	Если CSS тяжёлый — медленный р
Отдельная мобильная версия (m.	Проще делать «лёгкую» версию д	Google требует точного соответ
AMP (Accelerated Mobile Pages)	Исторически — быстрая загрузка	В 2026 почти мёртв. Google бол

Рекомендация: адаптивный дизайн с оптимизацией под мобильные. Точка.

Что чаще всего ломается на мобильных

- Межстраничная реклама и попапы, которые перекрывают контент (Google может понизить за это)
- Нечитабельный текст без зума (размер шрифта < 12px на мобильном)
- Кликабельные элементы слишком близко друг к другу (минимальный тач-таргет — 48×48 px)
- Контент, который не помещается без горизонтальной прокрутки
- Отсутствие viewport meta тега или неправильный initial-scale

1.5. JavaScript-рендеринг и краулинговый бюджет больших каталогов

Этот параграф — для тех, у кого сайт больше 10 000 страниц. Если у вас блог на WordPress с 50 статьями — пропускайте. Если у вас интернет-магазин на 100 000+ SKU или каталог статей на React — читайте внимательно.

Проблема: Google не видит то, что рендерится на клиенте

Googlebot умеет рендерить JavaScript, но делает это отложенно (deferred rendering) и с ограничениями:

- Two-wave indexing: первая волна — то, что видно без JS (HTML, SSR). Вторая волна — после рендеринга. Между волнами может пройти от нескольких дней до нескольких недель

• Краулинговый бюджет ограничен: Google не будет ждать, пока ваш React отрендерит 50 000 SKU

• Сторонние скрипты (аналитика, чаты, реклама) Googlebot может не выполнить или выполнить с ошибками

Что делать: SSR, SSG, ISR

Подход	Что это	Когда использовать
CSR (Client-Side Rendering)	HTML пустой, контент появляетс	[x] Плохо для SEO. Никогда для
SSR (Server-Side Rendering)	HTML генерируется на сервере п	[v] Хорошо для динамических ст
SSG (Static Site Generation)	HTML генерируется заранее при	[v] Идеально для блога, катего
ISR (Incremental Static Regene	Гибрид: статика + перегенераци	[v] Для крупных каталогов с ре

Рекомендация для каталога на 50 000+ страниц: Next.js с SSG для категорий и блога, SSR для карточек товаров, ISR для всего остального. Если стек — Vue/Nuxt — то же самое через nuxt generate и nitro.

[x] Анти-паттерн: «У нас SPA на React, но мы добавили prerender.io»

Prerender-сервисы (prerender.io, prerender.cloud) — костыль. Они создают статическую копию для Googlebot, но:
- Если сервис упадёт — Google увидит пустые страницы
- Каждая страница — отдельный запрос к вашему серверу, удваиваете нагрузку
- Кэш может протухать, и Google получит устаревший контент

Это временное решение. Полноценный SSR/SSG — единственный правильный путь.

Краулинговый бюджет: что это и как его не сжечь

Краулинговый бюджет — это количество URL, которое Googlebot готов просканировать на вашем сайте за определённое время. Зависит от:

- Авторитетности домена (новые сайты получают меньше)
- Скорости ответа сервера
- Количества «полезных» страниц

Сжигатели бюджета, которые надо вырезать

Что сжигает	Как лечить
Бесконечная фасетная навигация	Canonical на основную категори
Сессионные ID в URL (?sid=abc1	Cookies вместо URL-параметров
Поиск по сайту с выдачей в инд	noindex на результатах поиска
Простыни календарей (?d=2026-0	Canonical на сегодняшнюю дату
Редирект-цепочки (3+ редиректо	Прямой 301 с исходного URL на
Страницы с 5xx или soft 404	Чистка или правильный ответ се

Фасетная навигация: 50 000 URL — нормально или катастрофа?

Зависит от того, есть ли там уникальный контент. Если фильтр «красные шины Michelin размера R15» даёт реальную выдачу из 8 товаров и эти товары никогда не появятся в основной категории — оставляйте. Это длинный хвост, он может приносить 10–15% трафика. Если фильтр создаёт пустую страницу или страницу с одним товаром — закрывайте. Это сжигание бюджета.

Правило для фасетов:
1. Если страница даёт уникальную выдачу ≥ 3 товаров — оставляем в индексе
2. Если страница пустая или даёт дубль основной категории — noindex + canonical
3. Если страница фильтра даёт выдачу, которая полностью входит в основную категорию — закрываем в robots.txt

Блок «Владельцу на салфетке»

5 метрик, которые вы должны видеть в ежемесячном отчёте вашего подрядчика:

1. Процент страниц в индексе Google и Яндекса (Search Console +

Яндекс.Вебмастер). Если меньше 70% известных страниц — у вас техническая болезнь.

2. LCP и INP для мобильных в CrUX. Если CrUX не собирается — у вас настолько мало мобильного трафика, что Google не может измерить. Это тоже сигнал.

3. Количество ошибок в Search Console → Pages → Excluded. Любое число больше 0 — это деньги, утекающие в трубу.

4. Количество проиндексированных страниц vs известных. Если разрыв растёт — что-то блокирует сканирование.

5. TTFB (Time to First Byte) вашего сервера. Больше 600 мс — и никакой CDN не спасёт. Меняйте хостинг.

Если подрядчик не даёт вам эти 5 цифр каждый месяц в читаемом виде — вы платите не за SEO, а за воздух.

Артефакты главы

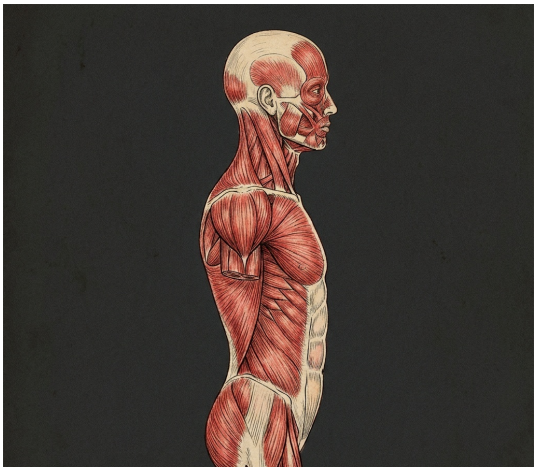
[v] Чек-лист экспресс-аудита технички (1 страница):

- [ ] Search Console → Pages: процент страниц в индексе
- [ ] Search Console → Core Web Vitals: статус мобильных
- [ ] robots.txt проверен в Search Console → robots.txt Tester
- [ ] Sitemap.xml загружен, ошибок нет
- [ ] CrUX по основным типам страниц
- [ ] Сканирование Screaming Frog: дубли, noindex, canonical
- [ ] PageSpeed Insights по 5 ключевым URL
- [ ] Проверка редирект-цепочек (не длиннее 1)
- [ ] Проверка фасетной навигации на дубли
- [ ] Проверка TTFB через WebPageTest

Если хотя бы 3 пункта красные — начинайте с этой главы. Не с контента, не со ссылок. С костей.

Глава 2 → «Мышцы, которые тянут ТОП». Контент и семантическое ядро. Если скелет собрали — пора наращивать мясо.

ГЛАВА 2. МЫШЦЫ, КОТОРЫЕ ТЯНУТ ТОП. Контент и семантическое ядро



Кости без мышц — труп. Контент — это мышечная масса: тексты, заголовки, структура, E-E-A-T. Здесь же — семантическое ядро как нервная сеть запросов. Не «пишите тексты», а «какой мышечный каркас делает вашу категорию сильнее конкурентов».

2.1. Семантическое ядро: сбор, кластеризация, приоритизация

Семантическое ядро — это не список из 5000 ключей в Excel. Это карта мышц вашего сайта: какие группы запросов тянут трафик, какие висят без нагрузки, каких групп вообще нет.

Три ошибки, которые делают ядро бесполезным:
1. Собрали всё подряд — 50 000 запросов, 80% из которых нулевые или околонулевые

2. Не кластеризовали — все запросы свалены в кучу, стратегия «одна страница = один запрос»
3. Не приоритизировали — все запросы равны, нет понимания, что важнее

Как собирать ядро в 2026

Каналы сбора в порядке убывания ценности:

Канал	Что даёт	Инструменты
Search Console → Performance →	Реальные запросы, по которым в	Search Console, Яндекс.Вебмаст
Конкуренты из топа	Запросы, которые приносят день	Serpstat, Keys.so, SpyWords, A
Подсказки поиска	Длинный хвост, который реально	Serpstat, парсеры подсказок
Яндекс.Вордстат и Google Trend	Объём спроса, сезонность	wordstat.yandex.ru
Внутренняя аналитика	Запросы, по которым уже приход	Метрика → Отчёты → Источники →

Анти-паттерн: покупать готовое ядро у «агентства». В 95% случаев это мусор: половина запросов нерелевантна, четверть — нулевки, ещё четверть — не та интентность.

Кластеризация: один запрос = один URL, но не один текст

Кластеризация — это группировка запросов по интенту и посадочной странице.

Пример для интернет-магазина автозапчастей:

Кластер	Запросы	Посадочная
Тормозные колодки (бренд)	«тормозные колодки TRW», «коло	Карточка категории бренда
Тормозные колодки (модель)	«колодки на Киа Рио», «тормозн	Подбор по авто
Информационный	«как заменить тормозные колодк	Статья в блоге
Сравнительный	«TRW или Bosch колодки», «что	Сравнительная статья

Правило: один кластер = одна страница. Если вы делаете две страницы под один кластер — каннибализация. Google выберет одну, вторую утопит.

Приоритизация: не все запросы равны

После кластеризации отсортируйте кластеры по:

1. Коммерческий интент (готовы купить) → выше приоритет
2. Объём спроса (wordstat) → средний вес
3. Конкуренция в топе (кто там сейчас) → обратный вес (если все сильные — дольше выводить)
4. Связь с продуктом (есть товар в наличии) → высокий приоритет
5. Текущая позиция (если уже 15-20 место, легче дотянуть до 10) → бонус

Формула приоритета = (коммерция × объём × релевантность) / конкуренция. Это не точная наука, но порядок даёт.

2.2. LSI, TF-IDF, BERT и MUM: как Google реально читает текст в 2026

Забудьте LSI (Latent Semantic Indexing) — этот термин умер лет 8 назад. Google давно не использует «скрытое семантическое индексирование» в смысле, в каком его продают копирайтерам.

Что реально работает в 2026:

BERT (Bidirectional Encoder Representations from Transformers)

BERT — нейросеть, которая понимает контекст слов в предложении. С 2019 года Google использует её для большинства англоязычных запросов, с 2021 — для русского.

Что это значит для текста:
- Google понимает, что «купить iPhone» и «iPhone купить дешево» — это разные интенты
- Понимает, что «Apple» в запросе «Apple акции» — это компания, а в «Apple iPhone» — бренд
- Понимает порядок слов и их взаимосвязь

Практический вывод: нет смысла вставлять ключ в текст «для плотности». Google смотрит на смысл, а не на вхождения. Текст

«купить iPhone недорого в Москве» без ключа вхождениями работает лучше, чем «iPhone iPhone iPhone купить Москва недорого iPhone».

MUM (Multitask Unified Model)

MUM — следующее поколение, мультимодальное (понимает текст, изображения, видео). Запущено в 2021, в 2026 — один из ключевых компонентов для сложных составных запросов (наряду с BERT и RankBrain, не заменяет их).

Что умеет MUM, чего не умеет BERT:
- Отвечает на сложные составные запросы («у меня анемия, какие витамины пить»)
- Синтезирует информацию из нескольких источников
- Работает с 75 языками одновременно
- Понимает визуальный контент

Практический вывод: для сложных тем (здоровье, финансы, право) — нужен авторитетный, структурированный контент, который MUM может использовать как источник для ответа. Это пересекается с E-E-A-T.

TF-IDF: что осталось от него

TF-IDF (Term Frequency — Inverse Document Frequency) — старая метрика, но до сих пор используется как один из факторов. Измеряет: насколько часто термин встречается в вашем тексте относительно других текстов в корпусе.

Практический вывод: если термин встречается в вашем тексте так же часто, как в топ-10 конкурентов — это норма. Если в 5 раз чаще — переоптимизация, Google может понизить. Используйте Serpstat или Keys.so для анализа TF-IDF топ-конкурентов и калибровки плотности.

2.3. E-E-A-T: где граница между реальным сигналом и