

PROPOSAL · INDEXATION · LARGE-SCALE SEO

Прискорення індексації великого сайту — 23 000+ сторінок

Діагностика причин повільної індексації + план дій під краул-бюджет і стек Nuxt + Cloudflare Pages — від команди, яка проходила це на власному сайті з 40 000+ сторінок.

з 2014

full-cycle studio

10+

років досвіду

400%

середній ROI

CA · US

клієнти

[Search Console / API](#)
[Серверні логи](#)
[Краул-бюджет](#)
[Nuxt / SSR](#)
[IndexNow](#)
[Sitemap-архітектура](#)

Ви просили не шаблон, а план — і це правильно. Тому одразу по суті. Важливо: ми проходили саме цю ситуацію на власному контентному проєкті у фінансово-оглядовій ніші з **40 000+ сторінок** — той самий симптом «виявлено, не проіндексовано». Назву проєкту публічно не показуємо (NDA), але механіка й рішення — ті самі, і нижче вони конкретно.

1. З чого починаємо діагностику

Не з сайту — з ДАНИХ. Сайт скаже «що», дані скажуть «чому».

1 Search Console: Page Indexing по причинах

Розділяємо «Discovered – currently not indexed» і «Crawled – currently not indexed» — це **два різні діагнози**. Перше = краул-бюджет (Google знає URL, але не заходить). Друге = заходив, але вирішив не індексувати (якість/дубль). Лікуються по-різному.

2 URL Inspection по вибірці

Беремо десятки сторінок: коли Google реально заходив (lastCrawlTime), що отримав, чи бачить контент. Часто видно, що Googlebot до більшості сторінок просто ще не дійшов.

3 Аналіз серверних логів — головне

Дивимось РЕАЛЬНІ візити Googlebot (за IP-діапазонами Google, а не за підставним User-Agent) — скільки разів на день і КУДИ він заходить. На новому домені з десятками тисяч сторінок це часто десятки візитів/день — звідси й повільна індексація.

4 Що бачить бот на Nuxt

Перевіряємо, чи отримує Googlebot готовий HTML, чи порожню оболонку (CSR). Для Nuxt на масштабі це критична точка.

2. Чому 23к сторінок індексуються повільно — що бачимо найчастіше

На новому домені такого розміру причини майже завжди з цього списку.

причина №1

Краул-бюджет і довіра нового домену

Google не витрачає бюджет обходу на 23к сторінок домену, якому ще не довіряє. Це головна причина «Discovered – not indexed»: URL знайдено, але черга на обхід ползе місяцями.

причина №2

Sitemap-індекс: дочірні карти не вичитуються

Для великих сайтів Google качає індекс-карту, але до дочірніх (де самі URL) доходить ліниво. URL рахуються «виявленими», але фактично не прочитані з карти.

причина №3 · ваш стек

Nuxt + Cloudflare Pages = ризик клієнтського рендера

Якщо контент домальовується в браузері (CSR), Googlebot на першому проході бачить майже порожню сторінку й відкладає індексацію. Для Nuxt на масштабі **SSR / пререндер обов'язковий**.

причина №4

Тонкий / near-duplicate контент

На оглядовому сайті частину сторінок Google може вважати близькими дублями або малоцінними → «Crawled – not indexed». Знижує й середню цінність домену, і бюджет.

причина №5

lastmod, якому Google не вірить

Якщо всі сторінки в sitemap мають однаковий свіжий lastmod (а контент не мінявся) — Google перестає довіряти датам і рідше перевіряє карту.

причина №6

Сторінки-сироти й глибина

Сторінки без вхідних внутрішніх посилань або глибше 3-4 кліків від головної — Googlebot до них майже не доходить, навіть якщо вони в sitemap.

3. Кроки з найбільшим ефектом

Порядок — за співвідношенням «зусилля → результат». Перші три дають найбільше.

1 Скоротити sitemap до ЯДРА, потім розширювати

Не подавати всі 23к одразу. Спершу пріоритетні (свіжі/цінні) сторінки → Google їх реально обробляє → довіра й бюджет ростуть → додаємо наступні хвилі. Саме це дало рух на нашому проєкті з 40k.

2 Гарантувати SSR / пререндер для Nuxt

Щоб Googlebot отримував готовий HTML з контентом, а не порожню оболонку. На Cloudflare Pages це вирішується конфігурацією рендера.

3 Внутрішня перелінокка hub-and-spoke

Підняти orphan-сторінки в межах 2-3 кліків від сильних: тематичні хаби, блоки «схоже», логічні перехресні посилання. Це прямий канал, яким бот доходить до глибини.

4 Чесний lastmod + IndexNow

lastmod тільки коли контент реально змінився. Плюс IndexNow (підтримується на Cloudflare) для миттєвого пінгу пріоритетних URL.

5 Прибрати/злити малоцінні сторінки

Near-duplicate і тонкі — об'єднати або закрити. Менше «шуму» → вища середня цінність домену → Google виділяє більше бюджету на решту.

6 Щотижневий контроль за логами й GSC

Дивимось, чи реальний Googlebot почав заходити глибше і чи росте частка проіндексованих — рухаємось за даними, а не наосліп.

Головне, що відрізняє великий сайт: тут не «полагодити одну помилку», а **керувати краул-бюджетом і довірою** — подавати Google рівно стільки, скільки він готовий якісно обробити, і розширювати в темпі його довіри.

4. Які інструменти використовуємо

Search Console — Page Indexing по причинах + URL Inspection API на вибірці (реальний статус і дата обходу).

Аналіз серверних логів — реальні візити Googlebot за IP Google (відсіюємо підставні User-Agent), частота й глибина обходу.

Краулер з JS-рендером — що насправді бачить бот на Nuxt (готовий HTML чи порожня оболонка), карта внутрішніх посилань і глибини.

Корпоративні enterprise-сервіси — для семантики, беклінків і моніторингу позицій (кросс-перевірка з реальними даними GSC).

Наш досвід саме з цим

Чому ми впевнені в плані, а не теоретизуємо.

Власний контентний проєкт у фінансово-оглядовій ніші, **40 000+ сторінок**: той самий симптом «виявлено, не проіндексовано», реальний Googlebot — десятки візитів на день на десятки тисяч сторінок, sitemap-індекс качався, а дочірні карти не вичитувались. Розібрали по логах і GSC, скоротили подачу до ядра, навели рендер і перелінокку — індексація зрушила. Назву проєкту не показуємо публічно (NDA), але це **не теорія — ми це вже проходили**. Для вашого сайту на Nuxt + Cloudflare Pages підхід той самий, з поправкою на стек.

WebCoreLab — full-cycle digital studio · Toronto, Canada

Дизайн, розробка, SEO/GEO і DevOps під одним дахом — **від ідеї до продажну й підтримки**. На ринку **з 2014**, **10+** років у digital, **20+** років сукупного досвіду команди.

Працювали з масштабними контентними сайтами — технічне SEO, індексація, вимірюваний результат.

ПАРТНЕРСТВА ТА ВИЗНАННЯ

Релевантні кейси — SEO та масштаб

Проєкти, близькі за інженерною суттю — SEO-інфраструктура, великі каталоги й платформи. Усі посилання живі.

E-COMMERCE · CATALOG

Message Warehouse

Великий каталог, чиста структура, органівка (US).

→ конверсія каталога вверх

GEO · SEO

TechFlow

Гео-видимість і органічний трафік, B2B.

→ **рост видимості в ніше**

GEO · LOCAL

GreenLeaf

Локальний каталог + органівка по регіонах.

→ **локальные заявки**

GEO · HEALTHCARE

MedCore

Структурований каталог, технічне SEO.

→ **органіка + доверие**

SEO · DATA

Ahrefs case

SEO-інфраструктура й контент-рушій під масштаб.

→ **масштаб органіки**

ENTERPRISE · LOGISTICS

XPO Logistics

Високонавантажена платформа, тисячі сторінок.

→ **стабільність под нагрузкой**

PLATFORM · SCALE

LatinAmericanCupid

Платформа на мільйони користувачів і URL.

→ **масштабируемость**

BRAND · E-COMM

Detox

Бренд-магазин (US), чиста індексована структура.

→ **конверсія и скорость**

SERVICE · BOOKING

Custom Maids Toronto

Локальний сервіс (Toronto), органічні заявки.

→ **поток заявок**

[Дивитись усі кейси →](#)