FIELD NOTES FROM A PRESALES SOLUTIONS ARCHITECT

Google Search Console "Not Indexed" Warnings You Should Leave Alone — and the Two Checks That Tell You Which Ones Matter

A field note about a Search Console report that said 88 pages were not indexed on a site where every article was on Google — what each "not indexed" reason really means, which ones are correct behaviour, and the two checks I now run before I touch a single setting.

Daanial Farrukh · 12 November 2026 · https://quabin.com/search-console-not-indexed-warnings-leave-alone/

This one did not start in a customer room. It started with my own site.

I opened Google Search Console, went to Indexing › Pages, and read two numbers.

Indexed: 37.

Not indexed: 88.

More than twice as many pages "not indexed" as indexed. The grey bar was bigger than the green one. The first reaction is the same for everyone who sees that chart.

"Something is broken. Google is ignoring my site."

I have heard the same sentence from customers who run a company website on WordPress. "Search Console says half our pages are not indexed. Can you fix it?"

So I asked myself the question I ask them.

"Which pages? Name one."

That's the trap. The "not indexed" count is not a list of problems. It is a list of every URL Google found and decided not to keep — and on a normal WordPress site, most of those URLs should never be kept.

What "Not Indexed" Actually Means

Search Console's "not indexed" is a bucket, not an error. It holds every URL Google has seen — through links, the sitemap, scripts, feeds — that it did not add to the index, with a reason attached.

The flow looks like this:

Google finds a URL → crawls it (or not) → reads robots rules → decides → Indexed / Not indexed + reason

The reason is the whole story. The number on its own tells you nothing.

When I opened each reason on my own report, this is what was inside:

- **Excluded by 'noindex' tag — 36.** Tag archives such as /tag/tls/, the search page /?s=, and the author page.
- **Crawled - currently not indexed — 49.** RSS feeds such as /tag/tls/feed/, the WordPress emoji script wp-emoji-release.min.js, and one article.
- **Not found (404) — 2.** /* and /cdn-cgi/l/email-protection.
- **Discovered - currently not indexed — 1.** /category/uncategorized/.

That is 88. Of those 88, exactly one was a real article. Everything else was a page that a good SEO setup keeps out of Google on purpose.

Reason by Reason — Which Ones Are Correct

Here's the part that trips people up. Every one of these reasons sounds like a failure. Most of them are the site doing exactly what it should.

Excluded by 'noindex' tag

This is the biggest bucket on most WordPress sites, and the one people most want to "fix".

Rank Math (and other SEO plugins) put a noindex tag on tag archives by default. A tag archive is a list of posts that share a tag. On a small site, /tag/tls/ and /tag/ssl/ show almost the same posts. Google sees thin, duplicate pages. The plugin tells Google not to index them, and Google obeys.

Every time I add an article with four tags, I can create up to four new tag archives. That is why this bucket grew from 4 to 20 to 36 while the articles grew from a handful to two dozen. **The bucket grows because the site grows. That is not decay.**

The same goes for /?s= (internal search results) and the author archive on a one-author site. Neither should be in Google.

The wrong fix is to untick noindex on tags so the number goes down. That pushes dozens of thin, duplicate pages into the index and makes the whole site look weaker.

Crawled - currently not indexed

This reason sounds the worst: Google visited the page and still said no.

On my report it held three kinds of URL:

1. **Feeds.** Every tag and every post on WordPress has a /feed/ URL. It is XML for feed readers, not a page for people. Google crawls it and correctly does not index it.
2. **Scripts.** wp-emoji-release.min.js is a JavaScript file WordPress loads on every page. Google fetches it to render pages. It is not a page.
3. **One real article.**

The first two belong there forever. The third one is the only line in the whole report that deserved a second look — and even that turned out to be a reporting lag, which I come back to below.

Not found (404)

Two URLs:

- /cdn-cgi/l/email-protection — this is Cloudflare's email address obfuscation. When the feature is on, Cloudflare rewrites email addresses in your HTML into a link to this path and decodes them with a script. Google follows the link and gets a 404. It is not a page on your site and never was.
- /* — a literal asterisk. Something, somewhere, produced a link to it. A URL that never existed returning 404 is the correct answer.

A 404 on a URL that was never a page is not an error. It is the truth. A 404 is only a problem when it is on a URL that used to be a real page, or one your own menu or sitemap still links to.

Discovered - currently not indexed

One URL: /category/uncategorized/. It exists because one older post sits in the default "Uncategorized" category. Google knows the URL but has not bothered to crawl it.

The tidy fix is to move that post into a real category. It is optional. Nothing is broken.

Why the Report Looks Worse Than the Site

There is a second reason the chart misleads: it lags.

On my site the Pages report showed "Last updated" a week or more behind for several runs in a row. The counts stood still at Indexed 19 / Not indexed 54 for close to two weeks while I published an article almost every day. Then they jumped in one step to 37 and 88.

The article that sat in "Crawled - currently not indexed" was the clearest example. The report listed it there. When I inspected the same URL with URL Inspection, the answer was **"URL is on Google"**. The report was showing an older crawl. The live state was fine.

Plain takeaway: **the Pages report is a summary of the past. URL Inspection is the current answer for one URL.** When they disagree about a real article, believe the inspection.

How to Actually Check This

You do not fix this report by chasing the number. You check that every page that *should* be indexed is indexed, and that every page that is excluded is excluded on purpose.

Step 1: Make your list of pages that matter

On most sites this is short: the homepage, a handful of service or about pages, and every article. Your post sitemap is the easiest source for articles.

```
curl -s https://example.com/post-sitemap.xml | grep -o "<loc>[^<]*</loc>"
```

Count them, so you know what "all of them" means:

```
curl -s https://example.com/post-sitemap.xml | grep -c "<loc>"
```

If an article is missing from the sitemap, that is a real problem — but a different one. A stale sitemap is a plugin cache issue, not a Google issue, and Google can still find the post through your internal links.

Step 2: Confirm each important page says "index"

Look at the robots meta tag the page actually serves:

```
curl -s https://example.com/your-article/ | grep -io '<meta name="robots"[^>]*>'
```

Expected output for an article:

```
<meta name="robots" content="index, follow, max-snippet:-1, max-video-preview:-1, max-image-preview:large"/>
```

Now run the same command on a tag archive:

```
curl -s https://example.com/tag/tls/ | grep -io '<meta name="robots"[^>]*>'
```

Expected output:

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

There you should see noindex. If an article shows noindex, that is the one real fault worth fixing — usually a per-post setting in the SEO plugin sidebar, or a "Discourage search engines" tick under Settings › Reading.

Also check the HTTP header, because some setups send it there instead:

```
curl -sI https://example.com/your-article/ | grep -i x-robots-tag
```

No output is fine. noindex here on an article is a fault.

Step 3: Inspect the real articles, not the report

This step is where most tutorials stop and where mine starts.

In Search Console, click the "Inspect any URL" box at the top, paste the article URL, and press Enter. Wait for "Retrieving data" to finish — on my site that takes 30 to 40 seconds, and typing in the box while it runs breaks the result.

You get one of three answers:

- **URL is on Google** — done. Whatever the Pages report says, this article is indexed.
- **URL is not on Google** or **URL is unknown to Google** — click **REQUEST INDEXING** and wait for the green "Indexing requested" dialog.
- An error such as a redirect, a 404 or "Excluded by noindex" on an article — that is a real fault. Go back to Step 2.

Two notes from doing this every day. A bookmarked deep link straight to the inspection page often gives a Google 404; going through the search box at the top always works. And the first REQUEST INDEXING click sometimes answers "Something went wrong" — a second click a few seconds later has worked every time for me.

Request once per URL. Pressing the button every day for the same URL does not move it up the queue.

Step 4: Read each "not indexed" reason once, then leave it

Open Indexing › Pages, click each reason, and look at the example URLs. Sort every URL into one of two piles:

- **Expected:** tag archives, /feed/ URLs, .js and .css files, /?s=, author pages on a one-author site, /cdn-cgi/ paths, junk URLs that were never pages.
- **Look closer:** any article, service page or the homepage.

Only the second pile gets work: inspect it (Step 3) and request indexing if the inspection says it is not on Google.

If the "look closer" pile is empty, the report is healthy, whatever the grey bar says.

The Checklist I Run Every Day

1. **The new article is in the sitemap** — curl -s https://example.com/post-sitemap.xml | grep -c "<loc>" went up by one.
2. **The new article says index** — the robots meta check in Step 2 shows index.
3. **The new article is requested** — URL Inspection, then the green "Indexing requested" dialog.
4. **One or two older articles are spot-checked** — URL Inspection says "URL is on Google".
5. **Every article URL in a "not indexed" bucket is inspected** — and requested once if it is really missing.
6. **Tag, feed, script and /cdn-cgi/ URLs are left alone** — no noindex changes, no removal requests.

With this, my own count went from "88 not indexed, panic" to "every article on Google, nothing to do" in about five minutes a day.

What I Tell Every Customer Now

When someone shows me a Search Console chart with a big grey bar, I don't open settings. I ask one question.

"Is any page you actually want people to find in that list?"

If the answer is "no, it's all tags and feeds", they're in good shape. If the answer is "I don't know", we're not done yet — and the next ten minutes go to Step 3, one real URL at a time.

I write about the questions I ask in real customer rooms — the ones that change architectures and surface the problems no one else noticed. If you found this useful, drop your email below and I'll send the next one when it's ready.

FAQs

Why does Search Console show more pages not indexed than indexed?

Because it counts every URL Google found, not only your pages. On WordPress that includes tag archives, RSS feeds, search results, author pages and script files. On a small site with many tags, these easily outnumber the real pages. The count matters far less than whether any real page is in the list.

Should I remove the noindex tag from WordPress tag archives?

No, not on a typical small site. Tag archives are thin lists of posts that overlap with each other, so SEO plugins such as Rank Math noindex them by default. Forcing them into the index adds duplicate, low-value pages and does not help your articles rank.

What does "Crawled - currently not indexed" mean for an RSS feed?

It means Google fetched the feed and decided not to index it, which is correct. A feed is XML for feed readers, not a page for searchers. Leave it alone. Only act if a real article or service page appears in this bucket, and inspect that URL first because the report can lag.

Why is /cdn-cgi/l/email-protection showing as a 404 in Search Console?

That path belongs to Cloudflare's email address obfuscation. Cloudflare rewrites email addresses on your pages into links to this path and decodes them with a script, so Google follows the link and gets a 404. It was never a page on your site and does not need fixing.

The Pages report says an article is not indexed, but URL Inspection says it is on Google. Which is right?

URL Inspection. The Pages report is a summary that can be a week or more behind, while inspection checks the current state of one URL. If inspection says "URL is on Google", the article is indexed and nothing needs to be requested.

How often should I click Request Indexing for the same URL?

Once. Repeated requests for the same URL do not speed it up. Request it after you publish or after a real fix, then check again with URL Inspection a few days later.

Read online: https://quabin.com/search-console-not-indexed-warnings-leave-alone/