

Checking if Google Successfully Merged Your www and non-www Mirrors

You set up the 301s, added the rel=canonical tags, maybe even submitted a clean sitemap. Still, a nagging worry persists: did Google actually combine the www and non-www versions of your site into a single canonical entity? **Checking if Google Successfully Merged Your www and non-www Mirrors** demands more than a quick glance at whether both resolve—it requires verifying that Google’s indexing and ranking signals are all pointing to one URL structure, not split across two competing house numbers on the same street. The problem of mirror duplication chews through crawl budget, dilutes link equity, and confuses your own reporting. We see webmasters who’ve had both versions indexed for months without realizing it, and their core pages wobble between positions because Google can’t decide which one to show.

This isn’t an academic phantom. A 2023 audit of 10,000 small business sites by a popular SEO toolset found that roughly 28% of domains had both www and non-www variants discoverable in Google, with no clear canonical signal recognized. That’s nearly three in ten sites bleeding authority into a mirror they thought they’d closed. The good news: you can confirm the merge using a handful of indicators that are publicly accessible, cost nothing, and double as a health check for your entire canonical setup.

What It Means When Google ‘Merges’ Mirrors

Google doesn’t literally merge two pages into one database entry. Instead, it chooses one URL as the canonical and treats the other as a lesser duplicate that it might index but rarely serves. If you’ve declared via signals—301 redirects, canonical tags, internal links—that www is your preferred version, Google should eventually settle on www as the canonical, cluster all signals under it, and push the non-www variant into a limbo where it only surfaces in rare cases. But “eventually” can mean days, weeks, or never if your signals contradict each other. The canonicalization process relies on a scoring mechanism that weighs over a dozen factors, as explained in [Google’s official guide on consolidating duplicate URLs](#). The more consistent those signals, the faster the merge sticks.

Think of it like a postal address dispute. You’ve got two addresses that both lead to the same building, but the official city records only recognize one. If you keep handing out business cards with the wrong address, sending mail from both, and painting both street numbers at the curb, even the best sorting machines get confused. Google’s crawlers face similar confusion.

Tools That Cut Through the Ambiguity

Before you dig into logs, you need a short stack of instruments that show you exactly what Google sees and decides.

- **Google Search Console URL Inspection** - the single most authoritative signal. It tells you whether a URL is indexed, the user-declared canonical, and the Google-selected canonical. If the canonical selected for the non-www version points to the www version (or vice versa), the merge is happening. If both versions list themselves as their own canonical with no cross-reference, you have a problem.
- **A raw curl with Googlebot user-agent** - reveals the actual redirect chain and HTTP headers Googlebot encounters, including any `Link: rel="canonical"` header. This is the ground truth before any JavaScript rendering.
- **Your sitemap** - a quick sanity check to see if you've accidentally submitted both www and non-www URLs.
- **Third-party bulk index checkers** - like [SpeedyIndex's bulk Google index checker](#) can tell you at scale whether multiple mirror variants are still showing up in Google's index. This saves hours when you have thousands of URLs.

:::info If you don't have curl installed, online redirect checkers work, but always test with the Googlebot user-agent because some servers treat crawlers differently. :::

What Google Search Console Actually Shows You

Fire up the URL Inspection tool and punch in your domain's non-www home page (for example, <https://example.com>). After the tool loads, look at the "Coverage" section. If the page is indexed, you'll see a note: "URL is on Google." Click "Indexing" and examine "Google-selected canonical." If it says <https://www.example.com/>, then Google has already chosen the www version as canonical, and you're in good shape. If it displays <https://example.com> as its own canonical, this mirror is not merged yet. Repeat the inspection for the www variant—it should list itself as canonical, not some other page.

Force Google to Index Your Links 

A common situation we see: the non-www URL is indexed and the Google-selected canonical is the non-www, but the www URL also shows "URL is on Google" with its own canonical. Both are indexed, neither merged. The site will struggle because link equity gets split. To break the tie, you must amplify the signals for the version you want. Within Search Console, also visit the "Sitemaps" report. If you submitted a sitemap that includes both protocols, you're telling Google they're both important. Remove any sitemap that lists the unwanted variant.

:::warning A browser may show you the right final URL because it respects 301, but Googlebot might see a different redirect if you've got a CDN rule that serves a different

response to crawlers. Always verify with the Googlebot user-agent. ...

A Step-by-Step Inspection Flow

Run through this decision tree every time you suspect a mirror issue. The graph exposes where most setups break.

flowchart TD

```
A[Start: Choose canonical version (e.g. www)] --> B{Curl both mirrors with Googlebot UA}
B -- Redirects to canonical? -->|Yes| C[Check for canonical tag/header]
C -- Points to canonical? -->|Yes| D[Inspect both URLs in Search Console]
D -- Non-www canonical selected is www? -->|Yes| E[Merge likely successful]
B -- No --> F[Fix redirect to 301]
C -- No --> F
D -- No --> G[Check internal links & sitemap; fix signals]
G --> B
```

Curl Commands That Expose the Real Redirect and Canonical

The following snippet follows all redirects using Googlebot's identity and prints the final URL. Replace example.com with your domain.

```
# Follow redirects and show the final resolved URL as seen by Google
curl -s -o /dev/null -w "%{url_effective}\n" -L -H "User-Agent: Googlebot" https://example.com
```

Run this for both the www and non-www variants. If both end up at the identical canonical URL, you've at least passed the redirect layer. Next, check the rel="canonical" header (some servers serve it in the HTTP response):

```
# Check for a Link canonical header (HTTP header)
curl -s -I https://www.example.com | grep -i 'link:' | grep -i 'canonical'
```

And to verify the sitemap doesn't contain mirror URLs:

```
# Extract all URLs from the sitemap and check for duplicates curl -s  
https://example.com/sitemap.xml | grep -oP 'https?://[^\s'
```