

STANDARD OPERATING PROCEDURE · WW-021

Diagnose and fix pages that are not indexed

Work through Search Console's Page indexing report, separate expected exclusions from real problems, fix the cause for important pages and validate the fix with Google.

TECHNICAL SEO

VERSION 1.0	EFFECTIVE 5 October 2026
NEXT REVIEW 5 April 2027	TIME 2 hours
DIFFICULTY Advanced	RELATED WW-004 · WW-008
AUDIENCE Technical SEO specialists and developers with Search Console access	
TOOLS Google Search Console · URL Inspection tool · A web browser	

Purpose

A page that is not in Google's index cannot appear in search results, however good it is. Search Console lists every reason pages are left out, but many of those reasons are expected and harmless. This procedure separates the important pages that should be indexed from intentional exclusions, fixes the cause for each important page and asks Google to validate the fix. It is finished when every important page is indexed or has a documented fix awaiting validation.

Scope

In scope

- Reviewing every reason in the Page indexing report
- Inspecting individual URLs and fixing noindex, robots.txt, canonical, redirect, 404 and quality causes
- Starting validation in Search Console

Out of scope

- Manual actions and security issues (check those reports separately)
- Content rewrites beyond what is needed to resolve thin or duplicate pages
- Site migrations (see WW-008)

Roles and responsibilities

Role	Responsibility
Technical SEO specialist	Analyses the report, diagnoses each cause and starts validation.
Developer	Fixes templates, robots.txt, canonicals and server errors.
Content owner	Decides whether thin or duplicate pages should be improved, merged or removed.

Prerequisites

- Full access to a verified Search Console property, ideally a **Domain** property (WW-004).
- A submitted XML sitemap containing only the pages that should be indexed.
- A list of the site's important pages, such as service, product and location pages.

Procedure

Review the report

Get an overview before fixing anything.

1. In Search Console, open **Indexing**, then **Pages**.

2. Use the filter at the top to switch from **All known pages** to **All submitted pages**. These are the pages you told Google you want indexed.
3. Under **Why pages aren't indexed**, list each reason with its page count in the work ticket.

Separate expected exclusions from problems

Most sites have exclusions that are correct.

1. Mark these as usually expected, after spot-checking a few URLs in each:
 - **Alternate page with proper canonical tag**
 - **Page with redirect**, for old URLs that redirect correctly
 - **Excluded by 'noindex' tag**, for pages such as thank-you, basket and account pages
 - **Not found (404)**, for pages that were deliberately removed
2. Mark these as likely problems that need investigating:
 - **Blocked by robots.txt** or **Excluded by 'noindex' tag** on an important page
 - **Server error (5xx)** and **Soft 404**
 - **Duplicate without user-selected canonical** and **Duplicate, Google chose different canonical than user**
 - **Crawled - currently not indexed** and **Discovered - currently not indexed** on important pages

Diagnose each important page

Inspect individual URLs to find the exact cause.

1. Click a reason, then click an affected URL and choose **Inspect URL**.
2. Click **Test live URL** and read the result for crawl permission, indexing permission, the canonical Google selected and the rendered page.
3. Open the URL in a browser, view the page source and check for `noindex` in a robots meta tag, the `rel="canonical"` value and the HTTP status.

Fix the cause

Apply the fix that matches the reason.

1. **Blocked by robots.txt**: remove or narrow the `Disallow` rule that matches the page, then check the file at `/robots.txt`.
2. **Noindex on an important page**: remove the `noindex` meta tag or `X-Robots-Tag` header from the template or page setting.
3. **Server error (5xx)**: ask the developer to check server logs for the URL and fix the error.
4. **Soft 404**: add real content to the page, or return a proper `404` or `410` status if it should not exist.
5. **Duplicate or wrong canonical**: set a self-referencing canonical on the preferred page, point duplicates to it, and make internal links and the sitemap use only the preferred URL.
6. **Crawled or discovered but not indexed**: improve the page so it is clearly useful and different from other pages on the site, and add internal links to it from relevant, well-linked pages.

[!WARNING] Before removing a `Disallow` rule or `noindex` tag, check why it was added. Some exist to keep private, duplicate or staging pages out of search.

Validate the fix

Ask Google to recheck.

1. For a single important URL, use **URL Inspection**, **Test live URL**, then **Request indexing**.
2. For a reason where all affected pages are fixed, open that reason in the report and click **Validate fix**.
3. Check the validation status weekly until it shows **Passed** or lists pages that still fail.

Quality checks

- ☐ Every reason in the report is marked as expected or problem, with spot-checked examples.
- ☐ No important page is blocked by robots.txt or noindex.
- ☐ Important pages have self-referencing canonicals and appear in the sitemap.
- ☐ Validation has been started for every fixed reason.
- ☐ Remaining unindexed important pages have a documented next step.

Troubleshooting

Symptom	Likely cause	Fix
Validation fails	Some URLs still have the issue	Inspect the failed examples, fix them and start validation again.
Google ignores the chosen canonical	Internal links, sitemap or redirects point elsewhere	Make every signal point to the same preferred URL.
Live test passes but page still not indexed	Google has not recrawled, or sees little value	Request indexing, add internal links and improve the content.
Many parameter URLs in the report	Filters or tracking create crawlable variants	Canonicalise to the clean URL and stop linking to parameter versions.

Records

- Save the reason counts before and after, the fixes made and the validation outcome with the work ticket.
- Note intentional exclusions in the client's technical SEO record so future audits do not repeat the work.

References

- [Page indexing report \(Search Console Help\)](#)
- [URL Inspection tool \(Search Console Help\)](#)
- [Block search indexing with noindex \(Google Search Central\)](#)
- [How to specify a canonical URL \(Google Search Central\)](#)

- [Robots.txt introduction and guide \(Google Search Central\)](#)

WRIGHT WEBSITES · WW-021 · V1.0 · COVERED BY THE RESOURCES LICENCE AT
WRIGHTWEBSITES.CO/LEGAL/RESOURCES-LICENCE