SEO Birdy

Technical SEO Checklist

32 checks that keep your site crawlable, indexable and fast, in the order search engines meet it.

FREE GUIDE · NOVEMBER 2026 · SEOBIRDY.COM



How to use this checklist

Work top to bottom. Each group protects one stage of how search engines process a page: discover and crawl, then index, then render and serve. Fix anything in groups 1 to 4 before spending time on speed or schema.

Site: _____ **Date:** _____

Checked by: _____ **Next audit:** _____

1. Crawling and robots.txt

- [ ] **robots.txt responds** at `/robots.txt` with 200 (or 404). Never 5xx or a timeout.
- [ ] **No** `Disallow: /` for all user agents. Important sections are allowed.
- [ ] **CSS and JavaScript** needed to render pages are not disallowed.
- [ ] `Sitemap:` **line** with the full sitemap URL.
- [ ] **AI crawlers decided on purpose:** OAI-SearchBot (ChatGPT search), GPTBot (training), Google-Extended (Gemini; doesn't affect Google Search).
- [ ] **Crawlable links:** important pages linked with `<a href>`, within about three clicks of home.
- [ ] **No orphan pages:** every page you want indexed has internal links to it.

2. XML sitemap

- [ ] **Valid file:** UTF-8, absolute URLs, at most 50,000 URLs or 50 MB per file.
- [ ] **Clean:** only canonical URLs that return 200 and aren't noindexed.
- [ ] **Accurate** `<lastmod>`**.** Google ignores `<priority>` and `<changefreq>`.
- [ ] **Submitted** in Google Search Console and Bing Webmaster Tools.

3. Status codes and redirects

- [ ] **Key pages return 200:** homepage, services, locations, top posts.
- [ ] **Removed pages return 404 or 410.** No soft 404s; no mass redirects to the homepage.
- [ ] **301 (or 308) for permanent moves,** in one hop. No chains or loops.
- [ ] **Internal links point to final URLs,** not to redirects.
- [ ] **Server errors investigated** in Search Console's Crawl Stats (host status).

4. Indexing and canonicals

- [ ] **No stray noindex** in a robots meta tag or `X-Robots-Tag` header on pages you want indexed.
- [ ] **Intentional noindex** on thank-you, cart, checkout and internal search pages, and those pages are not blocked in robots.txt.
- [ ] **One site version:** http, https, www and non-www all 301 to one, path by path.
- [ ] **Canonicals correct:** self-referencing, pointing to 200 URLs, matching the sitemap and internal links.
- [ ] **Duplicates consolidated:** tracking and filter parameters, trailing slashes, print versions.
- [ ] **Index status checked:** Page indexing report, then URL Inspection for key pages.

5. Mobile and rendering

- [ ] **Mobile parity:** same content, headings, links, alt text and structured data as desktop.
- [ ] **Viewport tag:** `<meta name="viewport" content="width=device-width, initial-scale=1">`
- [ ] **Renders cleanly:** main content and links visible in URL Inspection's tested page; no noindex in the original HTML that JavaScript is meant to remove.

6. Security and speed

- [ ] **HTTPS everywhere:** valid certificate on auto-renew, no mixed content, http redirects to https.
- [ ] **Core Web Vitals good** at the 75th percentile: LCP within 2.5 s, INP 200 ms or less, CLS 0.1 or less.
- [ ] **Fast server response:** HTML arrives in about 1.5 seconds or less.

7. Structured data, international and Bing

- [ ] **Structured data valid** in the Rich Results Test and Search Console enhancement reports.
- [ ] **hreflang** on every language version, including a self-reference (only if you have several languages or regions).
- [ ] **Bing Webmaster Tools verified;** IndexNow set up if your CMS or host supports it.
- [ ] **Monitoring on:** alerts for noindex, page errors, robots.txt changes, sitemap loss and SSL expiry.

Status code cheat sheet

Code	What Google does	Use it for
200	Passes content to indexing (not guaranteed)	Every live page
301, 308	Follows; strong signal the target is canonical	Permanent moves
302, 307	Follows; weak signal	Changes of a few days at most
304	Content unchanged; no effect on indexing	Caching
404, 410	Not indexed; indexed URLs removed; crawled less over time	Removed pages with no replacement
Soft 404	"Not found" content returning 200; reported in Search Console	Never: return a real 404
401, 403	Googlebot never logs in; not indexed	Private pages
429, 5xx	Crawling slows; persistent errors drop URLs	Fix fast
robots.txt 5xx	Crawling stops 12 hours, then last good copy for up to 30 days	Keep robots.txt at 200

Google follows up to 10 redirect hops, but every hop is a chance to fail. Redirect straight to the final URL.

Safe robots.txt for a small business site

```
User-agent: *
Disallow: /wp-admin/
Allow: /wp-admin/admin-ajax.php

Sitemap: https://www.example.com/sitemap.xml
```

robots.txt controls crawling, not indexing. A disallowed URL can still be indexed from links, without a description. To keep a page out of search, use `<meta name="robots" content="noindex">` and leave the page crawlable so search engines can see the tag. Google supports only `user-agent`, `allow`, `disallow` and `sitemap`; `crawl-delay` is ignored.

Five-minute checks you can run in a browser

What	How
robots.txt	Open `yourdomain.com/robots.txt` and read it
Sitemap	Open `/sitemap.xml` (or the URL in robots.txt); spot-check 10 URLs

What	How
Status code	Chrome DevTools, Network tab, reload; first row shows the status
noindex	View source (Ctrl+U) and search for `noindex`; check headers with `curl -I URL`
Canonical	View source, search for `rel="canonical"`
Site versions	Type http://, https://, with and without www; all should end on one URL
Index status	`site:yourdomain.com` search, then URL Inspection in Search Console

Monthly routine

- [ ] Read the Page indexing report: any new reasons under "Why pages aren't indexed"?
- [ ] Check the Core Web Vitals and HTTPS reports in Search Console.
- [ ] Inspect your five most important URLs with URL Inspection.
- [ ] Re-read robots.txt and confirm the sitemap still loads.
- [ ] Confirm the SSL certificate's expiry date and auto-renewal.
- [ ] After any plugin, theme or hosting change, rerun groups 1 to 4 the same day.

Change log

Date	What changed	Checked after



Free tools from SEO Birdy
Website SEO Audit, AI Visibility Audit and Fan-Out Map at seobirdy.com/tools. The full guide behind this checklist: www.seobirdy.com/blog/technical-seo-checklist