

Sentryly — SEO Checklist

Sitemap, structured data, and submission checklist for Google and Bing.

Required basics

- Hosted at /sitemap.xml or another URL referenced from robots.txt.
Crawlers find sitemaps via robots.txt or direct submission.
- Includes only canonical URLs (no duplicates, no parameter variations).
Duplicates dilute the signal and slow indexing.
- Every URL is reachable and returns 200.
Soft-404s and broken URLs in a sitemap erode crawl budget.
- Last-modified dates reflect actual content changes.
Search engines use this to schedule re-crawls.

What to leave out

- Admin URLs, search-result pages, user-account pages.
These shouldn't be indexed and shouldn't be in the sitemap.
- URLs blocked by robots.txt or meta noindex.
Conflicting signals confuse crawlers.
- Tag pages, archive pages, and other thin content.
Either improve them or exclude them.

Submission and monitoring

- Sitemap submitted in Google Search Console and Bing Webmaster Tools.
Status updates show last read date and any parsing errors.
- Sitemap regenerated automatically on publish/update.
Manual updates fall behind quickly.
- Sitemap index used if you have more than ~50,000 URLs.
Single sitemap files have a 50,000-URL or 50 MB limit.
- IndexNow ping configured for fast discovery on Bing-family engines.
Modern publishers use both sitemap and IndexNow together.

Educational content from Sentryly (<https://www.sentryly.com>). Free to share with attribution. Verify time-sensitive details on the relevant vendor site.