

Robots.txt Checklist

A robots.txt checklist verifies that your robots.txt file is accessible, valid, and allows search engines and the AI crawlers you choose to reach important content, while blocking only what you intend.

12 checks · 10 sections · For: Developers, technical SEOs

Site: _____ Checked by: _____ Date: _____

FILE ACCESSIBILITY

☐	Check file location	File is at the root and returns 200	High
☐	Check file size	File is under 500 KiB	Low

SYNTAX

☐	Validate syntax	Directives parse correctly with no typos	High
--------------------------	-----------------	--	------

USER-AGENT RULES

☐	Check user-agent groups	Groups are correctly scoped	Medium
--------------------------	-------------------------	-----------------------------	--------

ALLOW AND DISALLOW

☐	Check disallow rules	No important paths blocked	High
☐	Check CSS and JS access	Rendering resources are not blocked	High

SEARCH CRAWLERS

☐	Check Googlebot and Bingbot	Search crawlers can access key sections	High
--------------------------	-----------------------------	---	------

AI CRAWLERS

☐	Check AI crawler rules	GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot rules match policy	Medium
--------------------------	------------------------	--	--------

SITEMAP DECLARATION

☐	Declare sitemap	Sitemap directive uses an absolute URL	Medium
--------------------------	-----------------	--	--------

WILDCARDS

☐	Test wildcards	* and \$ patterns match only intended URLs	Medium
--------------------------	----------------	--	--------

TESTING

☐	Test key URLs	Key URLs tested against live rules	High
--------------------------	---------------	------------------------------------	------

DEPLOYMENT

☐	Check after deploy	Production file is not the staging file	High
--------------------------	--------------------	---	------

Common mistakes

- Shipping a staging Disallow: / to production
- Blocking CSS and JavaScript
- Using robots.txt to remove pages from the index
- Copying AI crawler blocks without a policy decision

- Forgetting that rules are case sensitive

When to run this checklist

- Before and after every deployment
- Before launching a site
- When setting an AI crawler policy
- When pages unexpectedly drop from search

SiteAuditLint workflow

Crawl, find, fix, re-crawl, compare, monitor. Run the same checks automatically at siteauditlint.com, free for up to 100 URLs per crawl.