GUIDES

Is your website open to ChatGPT and Perplexity?

By Erik Mäkitalo, founder of Mandalytics · Updated 3 November 2026 · 5 min read

THE SHORT ANSWER

Many websites block AI crawlers without knowing it, through a robots.txt line or a default setting in Cloudflare. And most AI crawlers cannot run JavaScript, so text that only appears after scripts load is invisible to them. You can check both for free in about ten minutes.

Check 1: does your robots.txt let AI tools in?

Open `yourdomain.com/robots.txt` in your browser. Look for lines that name these crawlers followed by `Disallow: /`:

Crawler	What it is for
OAI-SearchBot	Lets ChatGPT's search find and cite your pages
GPTBot	OpenAI's crawler for training its models
PerplexityBot	Lets Perplexity find and cite your pages
ClaudeBot	Anthropic's crawler
Google-Extended	A setting that controls whether Google's Gemini models may use your content; it does not affect normal Google Search

If you want to be named in AI answers, the search crawlers (such as OAI-SearchBot and PerplexityBot) should be allowed. Whether to allow the training crawlers is your own choice. A line like `User-agent: PerplexityBot` followed by `Disallow: /` means you are asking Perplexity to stay away. Remove it, or change it to `Allow: /`.



Most AI crawlers read only the raw HTML. Text that appears only after JavaScript runs is invisible to them.

Check 2: is something else blocking them?

If your site runs behind Cloudflare, open your Cloudflare dashboard and look for the setting that blocks AI bots or manages your robots.txt for AI crawlers. Some accounts have it switched on by default. Hosting firewalls and security plugins can do the same. If you are unsure, ask your web developer: "Are we blocking OAI-SearchBot, PerplexityBot or ClaudeBot anywhere?"

Check 3: can they read your text?

Most AI crawlers read only the raw HTML. To see what they see, open your most important page, right-click and choose "View page source" (not "Inspect"). Search the source for a sentence from the page.

- **You find your text:** great, AI crawlers can read it.
- **You find only an empty shell and scripts:** your page builds its text with JavaScript. Ask your developer for server-side rendering or a static version of your key pages.

What should I do next?

1. Allow the AI search crawlers in robots.txt and in any firewall or

Cloudflare setting.
2. Make sure your key pages show their text in the raw HTML.
3. Wait a few weeks, then ask your customers' questions in ChatGPT and Perplexity again and see whether you are named.

Then read how AI tools decide which sites to name for the next step.

Frequently asked questions

Should I allow GPTBot and other training crawlers?

That is your choice. Search crawlers like OAI-SearchBot and PerplexityBot decide whether you can be cited in answers. Training crawlers decide whether your content may be used to train models.

Will unblocking AI crawlers guarantee I am named?

No. It removes a barrier. Being named also depends on how useful, clear and well mentioned your site is.

Want these steps as a weekly to-do list for your own site?

Reading this is already a step forward. Mandalytics checks your website in Google and in AI tools every week, then sends the five tasks that will bring you the most visitors, each with the reason why and the steps to do it. You learn as you go, and every small win brings more visitors, more customers and, in the end, more revenue.

See how Mandalytics works

Where this comes from

- Vercel and MERJ: the rise of the AI crawler
- QuickSEO: what each AI platform cites