CHECKLIST · AI AT WORK

How to stop AI companies training on your website

Published 8 November 2026. Read the full story, sources and any updates at digitaladvisors.com.au/journal/ai-training-opt-out-settings/

- [ ] **Open your robots.txt.** Visit yourdomain.com.au/robots.txt and see which crawlers it already names.
- [ ] **Decide on training.** If you'd rather AI companies didn't train on your site, add Disallow rules for Google-Extended, GPTBot, ClaudeBot and meta-externalagent.
- [ ] **Leave the search crawlers alone.** Don't block Googlebot, Bingbot, OAI-SearchBot or Claude-SearchBot unless you're happy to lose that search traffic.
- [ ] **On Cloudflare, use Disallow for training.** Under Security, Settings, filter by Bot traffic, then AI Crawler and Agent access policies, avoid Block if you want to stay in Google and Bing.
- [ ] **Only exclude AI Overviews on purpose.** The Search generative AI setting in Search Console removes you from AI Overviews and AI Mode, and doesn't stop training.
- [ ] **Check for NoArchive or NoCache tags.** If your site uses them to keep pages out of Bing's AI, set Bing's new preferences before 1 March 2027.
- [ ] **Recheck in February 2027.** Bing starts reading the shared __usage_ai-training line on 1 February 2027, and Google plans to in early 2027.

NOTES

General information only, not legal, tax or financial advice. Check the full story and its sources before acting.