Có nên chặn GPTBot? Mẫu robots.txt cho 5 bot AI phổ biến

Bản in A4 — đánh dấu vào ô khi đã kiểm tra / hoàn thành, ghi chú bất thường vào cột Ghi chú.

Ngày:	Đối tượng kiểm tra:	Người thực hiện:

Cách viết file robots.txt cho từng bot AI kèm mẫu dùng ngay

Kịch bản áp dụng	Dòng lệnh cần thêm vào cuối file `robots.txt`	Ý nghĩa thực tế
Kịch bản 1 (Cân bằng nhất): Chặn bot huấn luyện AI nhưng vẫn mở cho AI tìm kiếm	`User-agent: GPTBot` / `Disallow: /` / `User-agent: Google-Extended` / `Disallow: /` / `User-agent: CCBot` / `Disallow: /`	Không đưa dữ liệu vào kho huấn luyện của OpenAI, Google và Common Crawl; vẫn cho phép `OAI-SearchBot` và `PerplexityBot` tìm và dẫn link
Kịch bản 2: Chặn toàn bộ 5 bot AI phổ biến	Lần lượt khai báo `User-agent:` cho `GPTBot`, `Google-Extended`, `ClaudeBot`, `PerplexityBot`, `CCBot` và đặt `Disallow: /` ngay dưới mỗi tên bot	Đóng cửa hoàn toàn với cả 5 bot AI lớn, vẫn giữ mở bình thường cho `Googlebot` làm SEO
Kịch bản 3: Chỉ chặn một thư mục riêng tư	`User-agent: GPTBot` / `Disallow: /tai-lieu-noi-bo/` (thay bằng tên thư mục thật trên web của bạn)	Bot AI vẫn đọc được trang chủ và blog công khai, nhưng không được vào khu vực riêng tư

Trước khi chỉnh file robots.txt

✓	Hạng mục	Ghi chú
☐	Xác định rõ mục tiêu của website: mở rộng khách hàng qua AI hay bảo vệ nội dung trả phí.	
☐	Liệt kê các thư mục riêng tư cần che giấu (trang nội bộ, tài liệu khách hàng, khu vực thành viên).	
☐	Kiểm tra lượng truy cập thực tế của từng bot AI trên Cloudflare hoặc nhật ký máy chủ.	
☐	Sao lưu toàn bộ nội dung file `robots.txt` hiện tại ra một tệp văn bản riêng trên máy tính.	

Khi cập nhật cấu hình chặn bot

✓	Hạng mục	Ghi chú
☐	Phân biệt đúng tên bot huấn luyện (`GPTBot`, `Google-Extended`, `CCBot`) và bot tìm kiếm (`OAI-SearchBot`, `PerplexityBot`).	
☐	Kiểm tra kỹ để chắc chắn không viết nhầm `User-agent: *` đi kèm `Disallow: /`.	
☐	Giữ nguyên đường dẫn `Sitemap:` và các quy tắc dành cho `Googlebot` đang hoạt động bình thường.	
☐	Mở lại địa chỉ `tenmiencuaban.com/robots.txt` bằng tab ẩn danh ngay sau khi bấm lưu.	

Người thực hiện (ký, ghi rõ họ tên)	Người kiểm tra / xác nhận