SAMPLE REPORT - produced from public files only; not a paid client engagement.

AI Search Readiness Audit

Site: https://example-agency.com **Date:** 2026-10-13 **Prepared by:** CiteScore Team

72/100

Good

A. AI crawler access: 15/30
B. Discoverability: 20/20
C. Structured data: 15/25
D. On-page signals: 22/25

Executive summary

Biggest issue: robots.txt blocks GPTBot, ClaudeBot, Google-Extended, Applebot-Extended, CCBot. These crawlers supply AI models and answer engines (OpenAI, Anthropic, Google Gemini, Apple, Common Crawl), so those systems currently can't read this site's pages. A /llms.txt file is published. No FAQPage schema on the homepage.

Top fixes

#	Priority	Effort	Fix
1	High	Low	Search/answer bots (OAI-SearchBot, ChatGPT-User, Claude-SearchBot, PerplexityBot) are allowed, which is good. Training bots are not. Decide whether to keep the Disallow: / rules for GPTBot, ClaudeBot, Google-Extended in robots.txt. If brand visibility in AI answers matters more than opting out of training, remove them. While they are blocked, Gemini (Google-Extended) and future OpenAI/Anthropic models learn nothing about the site from its own pages.
2	High	Medium	Add a visible FAQ section (5-8 real buyer questions) with matching FAQPage JSON-LD. Answer engines often quote Q&A blocks.
3	Medium	Medium	Mark up each core offering with Service schema (name, description, areaServed, provider) on its service page.
4	Medium	Low	Rewrite 2-4 H2s as the questions buyers actually ask (e.g. 'How long does SEO take?') and answer each in the first 40-60 words.
5	Medium	Medium	Publish one original, citable asset (benchmarks, pricing ranges or a case study with numbers). Answer engines prefer citing specific data.

Findings and evidence

Category	Check	Pts	Evidence
A. AI crawler access	GPTBot can crawl	0/5	Blocked: User-agent: GPTBot -> Disallow: /
A. AI crawler access	OAI-SearchBot can crawl	5/5	Allowed: User-agent: OAI-SearchBot (no root disallow)
A. AI crawler access	ClaudeBot can crawl	0/5	Blocked: User-agent: ClaudeBot -> Disallow: /
A. AI crawler access	PerplexityBot can crawl	5/5	Allowed: User-agent: PerplexityBot (no root disallow)
A. AI crawler access	Google-Extended can crawl	0/5	Blocked: User-agent: Google-Extended -> Disallow: /
A. AI crawler access	Bingbot can crawl	5/5	Allowed: User-agent: Bingbot (no root disallow)
B. Discoverability	Homepage returns 200	4/4	HTTP 200
B. Discoverability	robots.txt present	3/3	found
B. Discoverability	XML sitemap	6/6	https://example-agency.com/sitemap.xml -> HTTP 200, 91 <loc> entries
B. Discoverability	/llms.txt	7/7	HTTP 200, 2925 chars; first line: # Example Agency

Category	Check	Pts	Evidence
C. Structured data	Any valid JSON-LD	5/5	1 blocks, 0 unparseable; types: ImageObject, Organization, Person, PostalAddress, SpeakableSpecification, WebPage, WebSite
C. Structured data	Organization / LocalBusiness	6/6	present
C. Structured data	WebSite / WebPage	4/4	present
C. Structured data	FAQPage	0/6	not found on homepage
C. Structured data	Service / Product / Article	0/4	not found
D. On-page signals	Title tag 10-65 chars	4/4	44 chars: "Example Agency | The Embedded Growth Office"
D. On-page signals	Meta description 50-170 chars	4/4	139 chars: "Example Agency builds and operates the growth infrastructure between current revenue and the next level throu"
D. On-page signals	Exactly one H1	4/4	1 H1: "Turn growth into a system."
D. On-page signals	3+ H2 sections	3/3	9 H2
D. On-page signals	Question-style headings	0/3	0 headings phrased as questions
D. On-page signals	Open Graph tags	3/3	3/3 present (og:title, og:description, og:image)
D. On-page signals	Canonical + lang	2/2	canonical: https://example-agency.com/; lang: en
D. On-page signals	Robots meta not noindex	2/2	robots meta: none

How the score is calculated

The score is out of 100 and comes from four weighted categories, using only public files fetched on the audit date: **A. AI crawler access (30)**, split evenly across GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot, Google-Extended and Bingbot. A full block scores 0, a partial block half. **B. Discoverability (20)**: homepage 200 (4), robots.txt (3), XML sitemap (6), /llms.txt (7). **C. Structured data (25)**: any JSON-LD (5), Organization (6), WebSite/WebPage (4), FAQPage (6), Service/Product/Article (4). **D. On-page signals (25)**: title (4), meta description (4), single H1 (4), H2 structure (3), question headings (3), Open Graph (3), canonical + lang (2), no noindex (2). Grades: 80+ Strong, 65-79 Good, 45-64 Needs work, under 45 Poor.

Limits: this audit checks the homepage and site-wide files only. It does not query ChatGPT, Perplexity or Google AI Overviews, measure traffic, or guarantee citations. Robots rules are read as written. Some CDNs or firewalls also block bots at the network level, which a file check can't see.