

CHECKLIST · FREE RESOURCE

AI SEARCH VISIBILITY CHECKLIST

*42 checks to get your brand cited by ChatGPT, Google AI Overviews,
Perplexity and Copilot*

Maison
Mint

AGENCY

2026

AI search visibility checklist

A growing share of your future customers will never see a results page. They ask a question, an AI answers it, and three or four brands get named. Everybody else is invisible. In our own accounts, AI referrals are still a small slice of traffic, but it is the slice that converts best, because the visitor arrives already told that you are the answer.

This checklist is the exact list we run when we audit a site for AI search (people call it GEO, generative engine optimization, or AEO, answer engine optimization; same thing). It is in the order we do it, not the order it is usually written about. Technical access first, because nothing else matters if the crawlers cannot read you. Then entity clarity, then content shape, then off-site signals. Tick what is done, fix what is not, and come back in a quarter.

One honest note up front: nobody controls which sources an AI cites. What you control is whether you are easy to find, easy to understand and easy to quote. That is what the 42 points below are about.

What is inside

1. Can the AI crawlers even read you?
2. Does the AI know who you are?
3. Is your content shaped to be quoted?
4. Do other sources agree that you exist?
5. Are you measuring it?
6. What to do first

Can the AI crawlers even read you?

Most sites we audit fail two or three of these without knowing. Start here.

- ☐ **robots.txt allows GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot, Google-Extended and Bingbot**
 - Check the file directly. Many security plugins and CDNs block AI crawlers by default, which is fine if you chose it and a disaster if you did not.
- ☐ **The firewall or bot protection does not challenge AI crawlers**
 - Vercel, Cloudflare and others can serve a JavaScript challenge to anything that looks automated. If a plain curl to your homepage returns a challenge page, so does the AI.
- ☐ **Your key content is in the HTML, not rendered only by JavaScript**
 - Most AI crawlers do not execute JavaScript. View source, not inspect element. If the text is not there, it does not exist for them.
- ☐ **llms.txt exists at the root and says who you are, what you do, where you are and what you charge**
 - A plain-text summary for language models. Keep every fact identical to the site.
- ☐ **Sitemap is current and submitted to Google Search Console and Bing Webmaster Tools**
 - Bing feeds Copilot and ChatGPT search. Half the sites we see have never been verified in Bing.
- ☐ **Pages load in under 2.5 seconds and return a clean 200**
 - Slow or flaky pages get crawled less and quoted less.
- ☐ **One canonical version of every URL (www or not, trailing slash or not)**
 - Duplicates split whatever authority you have.
- ☐ **Pages that matter are not behind a login, a cookie wall or a geo block**
 - Including the consent banner: it must not hide the content itself.

Does the AI know who you are?

AI systems think in entities: a brand, a person, a place, connected to each other. Your job is to make those connections unambiguous.

- ☐ **One brand name, spelled one way, everywhere**
 - Maison Mint, not Maison Mint Agency on one page and MaisonMint on another. Same for the founder's name.
- ☐ **Organization schema on the homepage with name, legalName, url, logo, address, founder and sameAs links**
 - sameAs should point to LinkedIn, Instagram, Google Business Profile, Crunchbase, Wikidata if you have it.
- ☐ **Person schema for the founder or key experts, linked to the Organization**
 - With jobTitle, knowsAbout, sameAs and a real photo. AI systems weight named people heavily.
- ☐ **An About page that states what you do in one plain sentence, where you are based, since when, and for whom**
 - Write it as if explaining to a stranger who has eight seconds.
- ☐ **Author boxes on every article with a name, a title and a link to a profile page**
 - Anonymous content gets cited less.
- ☐ **Google Business Profile verified, categories correct, description matches the site**
 - This is the single most used entity source for local questions.
- ☐ **Consistent name, address and phone on the site, GBP, LinkedIn, directories and the registry listing**
 - Even a different street abbreviation can split the entity.
- ☐ **A Wikidata item exists for the company (and the founder, if notable enough)**
 - Free, takes an hour, and it is a primary source for several AI systems.
- ☐ **Profiles on Crunchbase, Clutch, Sortlist or the equivalents for your industry, filled in and matching**
 - Third-party confirmation of the same facts.

Is your content shaped to be quoted?

AI answers are built from short, self-contained passages. Long, meandering pages lose; pages that answer in the first 60 words win.

- ☐ **Key pages open with a short TL;DR block**
 - Three to five plain facts an AI can lift verbatim: what, for whom, where, since when, from what price. We put one on every pillar page.
- ☐ **Every important page answers its main question in the first two sentences**
 - Then explains. Not the other way round.
- ☐ **Headings are written as the questions people actually ask**
 - What does Google Ads management cost in 2026? beats Our pricing philosophy.
- ☐ **Each H2 section can be lifted out and still make sense on its own**
 - That is literally how it will be used.
- ☐ **Definitions are stated in one quotable sentence**
 - X is Y that does Z. AI systems love a clean definition.
- ☐ **Numbers, prices and dates are specific, not vague**
 - From 1,200 € per month is quotable. Affordable is not.
- ☐ **Tables and bulleted lists are used for comparisons and steps**
 - Structured content gets extracted far more often than prose.
- ☐ **FAQ sections exist on service pages and carry FAQPage schema**
 - Match the questions to real search and AI queries, not marketing questions.
- ☐ **Articles carry a visible published and updated date, and the updated date is real**
 - Freshness is a strong signal in AI Overviews. Do not fake it; actually update.
- ☐ **Original data, even small, is published with a method note**
 - Our own 12 accounts show X is citable. Industry average is Y, uncited, is not.
- ☐ **Claims about results carry a source or a clear our own data flag**
 - AI systems and readers both discount unsupported claims.
- ☐ **Every article has one clear topic and an internal link to its pillar page**
 - Topic clusters help the model understand what you are an authority on.
- ☐ **No AI-sounding filler**
 - If a paragraph could be deleted with no loss, delete it. Models trained on the web recognise and down-weight generic text.

Do other sources agree that you exist?

Citations are earned off-site. AI systems rely heavily on third-party mentions to decide whether a brand is real and relevant.

- ☐ **Mentions in at least three independent publications in the last 12 months**
 - Trade press, local business media, podcasts, partner blogs. Mentions, not just links.
- ☐ **Guest articles or expert quotes under the founder's name**
 - Bylines create the person-to-topic link that models pick up.
- ☐ **Listed in the roundups and best-of lists an AI would read for your category**
 - Search your category in ChatGPT and Perplexity and see which lists get cited. Get on them.
- ☐ **Reviews on Google, Clutch or the platform your buyers trust, with replies**
 - Review text is quoted directly in AI answers about agencies and local businesses.
- ☐ **LinkedIn company page and founder profile active and consistent with the site**
 - LinkedIn is crawled heavily and ranks for almost every B2B brand query.
- ☐ **Reddit, Quora or industry forum threads mention the brand accurately**
 - Reddit is a top source for several AI engines. Correct wrong information where you find it.
- ☐ **YouTube or podcast appearances with the brand in the title or description**
 - Transcripts are indexed and quoted.

Are you measuring it?

What you do not measure, you cannot improve, and AI search is still new enough that most of your competitors are not measuring it at all.

- ☐ **Ask ChatGPT, Perplexity, Gemini and Copilot the ten questions your buyers ask, monthly, and log who gets named**
 - A spreadsheet is enough. Note the prompt, the date, the brands and the sources cited.
- ☐ **GA4 has a channel group or exploration for AI referrals**
 - Filter session source for chatgpt.com, perplexity.ai, copilot.microsoft.com, gemini.google.com and claude.ai.
- ☐ **Search Console is checked for the question-style queries that trigger AI Overviews**
 - Queries starting with how, what, which, best and cost are the ones to track.
- ☐ **Server logs or a bot analytics tool show which AI crawlers visit and what they fetch**
 - If GPTBot never fetches your pricing page, start there.
- ☐ **A quarterly review is in the calendar**
 - AI search changes monthly. The checklist is only useful if it runs again.

What to do first

If you only fix five things: open robots.txt and the firewall to AI crawlers, write an llms.txt, put a one-sentence definition and an FAQ on every service page, add Organization and Person schema with sameAs links, and start the monthly log of who gets named for your ten questions. That alone puts you ahead of most of your market.

If you want us to run this audit on your site and hand you the ranked fix list, that is what our SEO and GEO service is. Book a free 30 minutes at mintagency.eu/contact and bring your domain.

Book a free 30 minutes

Written by Sarah Johanna Ferara and the Maison Mint team, Tallinn. Figures marked as our own data come from client accounts we manage and are illustrative. Free to share; please link to mintagency.eu/resources when you do. © Mintagency OÜ 2026.