MERIDIAN SEARCH

AI SEARCH VISIBILITY

FREE RESOURCE

The AI Visibility Checklist

27 signals that determine whether AI systems cite your brand — and what to fix when they don't.

meridiansearch.io

HOW TO USE THIS CHECKLIST

A diagnostic, not a guarantee.

This checklist covers the 27 structural signals AI systems use to determine whether your brand is cited in generated answers. It is organized into five categories: Entity Recognition, Schema Markup, Content Architecture, Authority Signals, and Technical Infrastructure.

Work through each item and mark it as complete, partial, or missing. Items marked missing or partial represent citation gaps — structural reasons why AI systems may be citing competitors in your place.

A few important clarifications about what this checklist is and is not:

It is a self-audit tool, not a replacement for a technical investigation.

Several items — particularly in the Schema and Technical categories — require direct inspection of your page source, robots.txt, and Search Console data to assess accurately. Do not mark items complete based on assumptions.

Completing all 27 items does not guarantee citation.

AI citation is probabilistic, not deterministic. These signals improve the probability that AI systems will cite your brand for relevant queries. They do not guarantee it. Category competition, query intent, and training data recency also affect outcomes.

The items are not equally weighted.

Entity Recognition and Schema items tend to have the highest individual impact. If you're prioritizing, work through Categories 01 and 02 first before addressing Content Architecture and Authority signals.

> If this checklist surfaces gaps you don't have the internal resources to close, the AI Visibility Audit at meridiansearch.io provides a complete technical investigation, competitor benchmarking across 40–60 queries, and a prioritized 90-day roadmap.

CATEGORY 01 **Entity Recognition & Knowledge Graph**

AI systems build answers from entities they can identify and verify. Before any other signal matters, the system needs to establish that your brand is a real, distinct entity — not an ambiguous string of text.

SIGNAL 01

Consistent NAP data across all directories

Your brand name, address, and phone number must be identical across Google Business Profile, LinkedIn, Crunchbase, Apple Maps, Yelp, and every other directory where you appear. AI systems cross-reference these sources to verify entity legitimacy. A single inconsistency introduces ambiguity.

SIGNAL 02

Wikipedia or Wikidata entry exists

AI systems weight Wikipedia and Wikidata heavily as authoritative entity sources. A Wikidata entry with accurate properties (founded, industry, HQ, founder, official URL) gives AI systems a trusted anchor for your entity — even if your Wikipedia article is minimal.

SIGNAL 03

Google Knowledge Panel is claimed and accurate

A claimed Knowledge Panel signals to Google's systems — and to AI models trained on Google data — that your brand is verified. Inaccurate panels actively harm citation accuracy: the AI may cite you with wrong information, or avoid citing you to prevent surfacing bad data.

SIGNAL 04

Brand name is unambiguous and not shared with other entities

If your brand name is a common word, a place name, or shared with an unrelated entity, AI systems struggle to resolve which entity is being referenced. Disambiguation — through consistent modifier usage in your own content — is required to train AI systems toward the correct entity.

SIGNAL 05

Founding date, jurisdiction, and industry are consistently stated across sources

AI systems use structured attributes to anchor entities in context. Conflicting founding dates or industry classifications across your website, LinkedIn, Crunchbase, and directories create entity resolution failures that suppress citation confidence.

CATEGORY 02 **Schema Markup & Structured Data**

Schema tells AI systems what your pages are about without requiring inference. Pages with explicit structured data are consistently preferred over pages that require interpretation.

SIGNAL 06

Organization schema on homepage with complete properties

Your homepage Organization schema should include: name, url, logo, description, foundingDate, areaServed, contactPoint, and sameAs (linking to all authoritative profiles). Missing properties reduce the completeness score AI systems use to assess citation worthiness.

SIGNAL 07

sameAs links connect all authoritative profiles

The sameAs property in your Organization schema should list every authoritative external profile: LinkedIn, Crunchbase, Wikidata, GitHub, industry databases. This creates a verified entity graph that AI systems can traverse to confirm your brand's legitimacy.

SIGNAL 08

Service or Product schema on all relevant pages

Every service or product page should have explicit schema defining what is offered, who it's for, and what it costs (or a price range). AI systems use this to match your pages to specific category queries — without it, they infer from unstructured text, which is less reliable.

SIGNAL 09

FAQ schema on high-intent pages

Pages targeting question-format queries should include FAQPage schema with direct, complete answers. AI systems extract these answers verbatim for citation. An FAQ schema answer that directly addresses a search query is one of the highest-confidence citation signals available.

SIGNAL 10

BreadcrumbList schema on all non-homepage pages

BreadcrumbList schema helps AI systems understand your site architecture and the relationship between pages. It also improves how your URLs are rendered in AI-generated answers, which affects click-through when you are cited.

SIGNAL 11

No duplicate or conflicting schema blocks on any page

Duplicate JSON-LD blocks — even with identical content — introduce parsing ambiguity. AI systems encountering conflicting structured data on a single page often default to ignoring all schema on that page. Audit every page for schema conflicts before other optimizations.

CATEGORY 03 **Content Architecture for AI Extraction**

AI systems extract answers from content. The structure of that content determines whether your page is cited or skipped — regardless of how good the underlying information is.

SIGNAL 12

Primary answer appears in the first 100 words of every page

AI retrieval systems prioritize content that answers the implied query immediately. Pages that open with context, backstory, or navigation before reaching the answer are consistently deprioritized over pages that state the answer in sentence one. Restructure for answer-first architecture.

SIGNAL 13

H2 and H3 headings are written as questions or direct statements

AI systems use headings as query-matching signals. A heading written as "What is [your service]?" or "[Your service]: How it works" maps directly to question-format queries. Generic headings like "Overview" or "About Us" provide no matching signal.

SIGNAL 14

Content contains explicit, quotable definitions of your category and positioning

AI systems need a single, clearly stated sentence that defines what you do, who you serve, and what makes you different. This sentence becomes the basis for citations. If no such sentence exists in your content, AI systems construct one from inference — often inaccurately.

SIGNAL 15

Long-form content uses clear section structure with topic-specific subheadings

Unbroken long-form content without clear section delineation is difficult for AI systems to extract specific answers from. Each section should address one specific question or topic, with a subheading that signals what that section covers.

SIGNAL 16

Statistics, claims, and data points are linked to original sources

AI systems that cite specific statistics check whether those statistics are sourced. Unsourced claims reduce citation confidence. Every specific data point in your content should link to its primary source — not a secondary aggregator.

SIGNAL 17

Content is updated within the past 12 months

AI systems weight recency as a freshness signal, particularly for fast-moving categories. Content that has not been updated in over 12 months receives lower citation priority than equivalent content with a recent modification date. Systematically update high-priority pages.

CATEGORY 04 **Authority & Citation Signals**

AI systems infer trustworthiness from the sources that already trust you. Third-party validation — not self-declaration — is the primary driver of citation authority.

SIGNAL 18

Brand is mentioned by name in authoritative third-party content

Mentions in industry publications, analyst reports, reputable news outlets, and established blogs are the strongest authority signals available. AI systems trained on web data weight these mentions heavily. A single mention in a high-authority source outweighs dozens of self-published assets.

SIGNAL 19

Backlink profile includes links from authoritative industry sources

Links from authoritative domains remain a strong proxy for entity trust in AI training data. The source quality matters far more than quantity. Ten links from relevant, high-authority domains carry more weight than 500 links from low-authority directories.

SIGNAL 20

Brand appears in at least two industry-specific databases or directories

Category-specific databases (G2, Capterra, Clutch, industry associations, vertical directories) signal to AI systems that your brand is a recognized participant in its category — not a self-declared entrant. Presence in these sources is a prerequisite for category citation.

SIGNAL 21

Founder or leadership team has individual authority signals

When individual team members have published content, speaking credits, or citations in authoritative sources, those signals transfer to the brand entity. AI systems associate personal authority with organizational credibility, particularly for service businesses.

CATEGORY 05 **Technical Infrastructure**

Technical issues create silent barriers to AI citation. These are the infrastructure checks that most brands skip — and that explain why structurally sound content still fails to surface.

SIGNAL 22

llms.txt file exists at domain root

The llms.txt standard provides AI crawlers with explicit guidance on your most important pages, your entity context, and how you want your brand described. It is the most direct signal you can send to AI systems — and most brands have not implemented it.

SIGNAL 23

robots.txt does not block AI crawlers

Some robots.txt configurations inadvertently block GPTBot, ClaudeBot, PerplexityBot, and other AI crawlers. If AI systems cannot access your pages, they cannot cite them. Audit your robots.txt against the current list of AI crawler user agents.

SIGNAL 24

Canonical tags are correctly implemented with no .html suffix conflicts

Canonical URL mismatches — particularly .html vs. clean URL conflicts — cause AI systems to encounter duplicate content signals that reduce citation confidence. Every page should have a canonical tag that exactly matches its live, indexable URL.

SIGNAL 25

Page load speed under 3 seconds on mobile

Slow-loading pages are crawled less frequently and less completely by both search engines and AI crawlers. Core Web Vitals performance correlates with crawl depth. Pages that load in under 3 seconds on mobile receive more complete indexation than slower equivalents.

SIGNAL 26

og:title and og:description match page content accurately

Open Graph metadata is used by AI systems to understand page context before full content extraction. Titles or descriptions that are generic, keyword-stuffed, or do not accurately reflect page content reduce the accuracy of AI categorization for that page.

SIGNAL 27

Site has an XML sitemap submitted to Google Search Console

A submitted XML sitemap ensures AI systems that rely on Google's index have a complete map of your content. Pages not in the sitemap are discovered through crawl only — which is slower, less complete, and less reliable for new or recently updated content.

NEXT STEPS

What to do with what you found.

If you completed this checklist and identified gaps: the next step depends on how many items you marked incomplete and which categories they fall in.

1–5 gaps, all in Technical or Schema categories

These are execution items. Most can be resolved internally with a developer and a content editor working from this checklist. Prioritize Entity Recognition and Schema before anything else.

6–15 gaps across multiple categories

You have a structural visibility problem that requires coordinated remediation across technical, content, and authority tracks simultaneously. An AI Visibility Audit will map every cause, benchmark your position against three competitors, and produce a sequenced 90-day plan.

15+ gaps, or uncertainty about how to assess any item

Start with the audit. Attempting to fix a complex AI visibility problem without a complete diagnostic risks resolving low-impact items while missing the structural causes of your citation absence.

BOOK THE AI VISIBILITY AUDIT
meridiansearch.io/audit

GENERAL ENQUIRIES
hello@meridiansearch.io