GEO Methodology Guide

Complete Audit Framework, Scoring System & Platform Strategies

Version 1.0 | February 2026
For clients, partners, and technical practitioners

Contents

1. Audit Workflow Overview

The GEO audit follows a three-phase process: Discovery, Parallel Analysis, and Score Aggregation.

Phase 1: Discovery & Reconnaissance

- Fetch homepage, detect business type (SaaS, Local, E-commerce, Publisher, Agency)
- Crawl sitemap (up to 50 pages, prioritized by navigation hierarchy)
- Collect page-level data: headings, word count, schema, links, images, meta tags
- Respect robots.txt; 30-second timeout per page; 1-second delay between fetches

Phase 2: Parallel Analysis

Five specialized analysis modules run in parallel, each producing a category score (0-100) with detailed findings:

Module	Focus	Key Metrics
Citability Analyzer	AI quotability	Passage self-containment, answer blocks, stat density
Brand Analyzer	Entity recognition	YouTube, Reddit, Wikipedia, LinkedIn presence
Technical Analyzer	Crawler access	robots.txt, llms.txt, rendering, meta tags
Content Analyzer	E-E-A-T signals	Author bios, credentials, freshness, depth
Schema Analyzer	Structured data	JSON-LD types, completeness, validation

Phase 3: Score Aggregation

The composite GEO score is computed as a weighted average:

GEO Score = (Citability x 0.25) + (Brand x 0.20) + (E-E-A-T x 0.20) + (Technical x 0.15) + (Schema x 0.10) + (Platform x 0.10)

2. AI Citability Analysis (25%)

This dimension measures how likely AI systems are to extract and cite passages from your content. It carries the highest weight because citation is the fundamental unit of GEO success.

What Gets Measured

- **Passage self-containment:** Can a paragraph answer a question on its own?
- **Answer block quality:** Are direct answers placed in the first 200-300 words?
- **Statistical density:** Fact-to-word ratio (target > 1:80)
- **Heading hierarchy:** Clean H1>H2>H3 structure (2.8x citation boost)
- **FAQ sections:** Question-answer pairs AI engines extract heavily

Optimization Targets

- Optimal passage length for citation: **134-167 words**
- Lead every section with a clear, concise answer (inverted pyramid)
- Include TL;DR statements so sections stand alone as answers
- Use precise statistics: "15%" beats "about 15%"
- Original/proprietary data is the **#1 citation magnet**

3. Brand Authority Assessment (20%)

AI systems need to trust an entity before citing it. Brand authority measures how recognizable and trustworthy your brand is across the platforms AI models draw from.

Signals Evaluated

- **Wikipedia/Wikidata presence:** Wikipedia accounts for 7.8% of ChatGPT citations
- **Reddit mentions:** Reddit dominates Perplexity (6.6% of citations)
- **YouTube presence:** #2 cited source on Google AI Overviews (18.8%)
- **LinkedIn company profile:** Professional authority signal
- **Review volume and velocity:** Active reviews signal living, trusted business
- **Cross-platform NAP consistency:** Character-for-character match everywhere

Key Insight

AI engines favor **earned media** (third-party coverage, reviews, industry mentions) over brand-owned content. A mention in a Reddit thread or YouTube review carries more citation weight than a page on your own website.

4. Content E-E-A-T Evaluation (20%)

Experience, Expertise, Authoritativeness, and Trustworthiness are the quality signals that differentiate citable content from noise.

Signal	What AI Looks For	Implementation
Experience	First-hand knowledge	Case studies, project photos, real examples
Expertise	Demonstrated skill	Author credentials, certifications, detailed how-tos
Authoritativeness	Recognition by others	Awards, press, third-party mentions, reviews
Trustworthiness	Verifiable accuracy	Sources cited, real data, transparent business info

Author Authority

Named experts with structured author bios and Person schema are significantly more likely to be cited. Anonymous content underperforms. Every piece of educational content should have a named, credentialed author.

5. Technical GEO Infrastructure (15%)

If AI crawlers cannot access your content, nothing else matters. Technical GEO ensures your site is crawlable, renderable, and machine-readable.

AI Crawler Access

Crawler	Platform	Priority
GPTBot	ChatGPT / OpenAI	Tier 1 (Critical)
ClaudeBot	Claude / Anthropic	Tier 1 (Critical)
PerplexityBot	Perplexity AI	Tier 1 (Critical)
Google-Extended	Gemini / Google AI	Tier 1 (Critical)
Bingbot	Bing Copilot	Tier 1 (Critical)
Amazonbot	Alexa / Amazon	Tier 2
Bytespider	TikTok	Tier 2
ChatGPT-User	ChatGPT Browsing	Tier 2
cohere-ai	Cohere	Tier 2

llms.txt

An emerging standard that helps AI systems understand your site structure. While 8/9 studied sites saw no measurable change, it remains a low-cost investment. Best practice: limit to 10-20 high-value pages, update quarterly.

Additional Technical Factors

- Server-side rendering (SSR) for all public content pages
- Sub-3-second page load time
- Valid canonical URLs and no duplicate content
- IndexNow protocol for Bing/Copilot indexing speed

6. Schema & Structured Data (10%)

Google's May 2025 guidance explicitly recommends JSON-LD for AI content. Schema has evolved from a nice-to-have to **critical infrastructure** that acts as connective tissue between websites and AI agents.

Schema Type	Purpose	Priority
LocalBusiness	Business identity, NAP, hours	Essential
Organization	Corporate entity recognition	Essential
FAQPage	Q&A pairs AI engines extract	Essential
Service	Service offerings with descriptions	High
Product	Product data with pricing	High
Review + aggregateRating	Social proof and trust signals	High
Article	Blog/content with author attribution	High
HowTo	Step-by-step processes	Medium
Person	Author credentials and expertise	Medium
ContactPoint	Contact information structure	Medium
BreadcrumbList	Site navigation hierarchy	Standard

Schema contributes approximately **10% of Perplexity's ranking factors** (Qwairy). While the direct weight is moderate, schema enables AI systems to understand entity relationships that influence all other dimensions.

7. Platform Optimization (10%)

Only 44% of AI citations come from owned websites. 48% come from community platforms. Your presence on the platforms AI models cite from is essential.

YouTube

18.8% of top 10 on Google AI Overviews; 13.9% on Perplexity

Reddit

46.7% of Perplexity's top 10 cited sources; community trust signal

Wikipedia

47.9% of ChatGPT's top 10; entity recognition foundation

LinkedIn

Professional authority; 1.3% of Google AI Overviews

Houzz / Yelp / BBB

Industry-specific platforms for local business authority

8. Platform-Specific Strategies

Optimizing for ChatGPT

- Prioritize encyclopedic, factual content with dense statistics
- "Best X" blog lists account for 43.8% of all cited page types
- Wikipedia presence is the strongest single signal (47.9% of top 10)
- Authoritative tone with source citations

Optimizing for Perplexity

- Community presence (Reddit, forums) is the dominant signal
- Recency matters: Perplexity favors recently published content
- YouTube content is heavily cited (13.9% of top 10)
- 2.76x more sources per query than ChatGPT (broader citation)

Optimizing for Google AI Overviews

- Existing search rankings still matter (but E-E-A-T amplifies weak positions)
- Balanced mix of social and professional signals
- YouTube is #2 cited source (18.8% of top 10)
- Schema markup directly influences content understanding

9. The E-E-A-T Revolution

The single most important shift in the GEO era: **entity authority has replaced domain authority**. Traditional DA has collapsed to just r=0.18 correlation with AI citations. What matters now:

Old Signal	New Signal	Why
Domain Authority (DA)	Entity Recognition	AI needs to identify you as a known concept
Backlink quantity	Cross-platform presence	NAP consistency across all platforms
Anchor text optimization	Earned media coverage	Third-party trust > self-promotion
Keyword density	Author credentials	Named experts with Person schema
Link building campaigns	Review velocity	Active reviews signal living business
Content length	Verifiable claims	AI systems must trust before citing

Critical insight: Pages ranking #6-10 with strong E-E-A-T get cited **2.3x more often** than #1 pages with weak authority. You do not need to outrank the competition; you need to out-trust them.

10. Content Strategy for AI Citations

What Gets Cited Most

- **Original/proprietary data** — First-party research is the #1 citation magnet
- **Precise statistics** — "15%" beats "about 15%"
- **FAQ content** — All AI engines heavily leverage Q&A; pairs
- **Direct answers first** — Place response in first 200-300 words
- **Entity specificity** — Name + service + location in content
- **"Best X" lists** — 43.8% of ChatGPT cited page types

Content Structure Best Practices

- Lead with clear, concise answer (inverted pyramid)
- Clean H2/H3 hierarchy signals passage topics to AI systems
- TL;DR statements so sections stand alone as answers
- Include FAQ sections on every page
- "Last updated" timestamps for freshness signals
- Topic clusters with interlinking (30% higher citation rates)

11. Measurement & Tools

GEO measurement requires new tools and metrics beyond traditional SEO analytics:

Tool/Metric	Purpose	Category
Share of Model (SoM)	Brand frequency in AI responses	Primary KPI
Siftly	AI citation tracking across platforms	Citation Monitoring
Gauge	Share of Model measurement	Citation Monitoring
Profound	Citation pattern analysis	Research
Qwairy	Provider citation behavior studies	Research
GA4 (AI referral)	AI-referred traffic attribution	Traffic Analytics
Manual testing	ChatGPT + Perplexity query monitoring	Ongoing QA

Key Performance Indicators

- **Citation frequency:** How often AI cites your content per relevant query
- **Share of Model:** Your brand's share of AI-generated responses vs competitors
- **Citation sentiment:** How positively AI presents your brand
- **AI-referred traffic:** Sessions from AI platforms (track in GA4)
- **Entity recognition:** Whether AI correctly identifies your brand and attributes