Generative Engine Optimization (GEO): Visibility Metrics, Behavioral Effects, and Practical Interventions

Abstract

Generative engines (GEs) such as Google AI Overviews, Bing Copilot, and Perplexity are shifting user attention from linear search results to synthesized, source-attributed answers. This paper consolidates the current state of knowledge on Generative Engine Optimization (GEO): we define visibility metrics beyond traditional SEO ranking, summarize verifiability audits of GE answers, and compile behavioral evidence about click-through under AI summaries. Drawing on the KDD'24 GEO framework, we report that lightweight editorial changes—adding verifiable statistics, correct citations, and concise quotations—consistently increase a source's visibility within generated answers by double-digit percentages on benchmark tasks, with transferable gains on a deployed GE. We discuss domain-specific effects, limitations, and a reproducible GEO audit protocol.

Keywords: Generative Engine Optimization (GEO); generative engines; AI Overviews; visibility metrics; Position-Adjusted Word Count; G-Eval; RAG; zero-click behavior.

1. Introduction

Generative Engine Optimization (GEO) addresses how content can be structured and written so that generative engines (GEs) select, attribute, and surface it prominently in synthesized answers. Unlike SEO—where a single ranked listing governs visibility—GEs distribute attention across inline citations within the answer body and ancillary source modules. The GEO literature formalizes this as a shift from page rank to impression metrics within the answer itself.

2. Definitions and Formalization

We model a GE as a function $f\_GE(q, P) \rightarrow r$ that retrieves sources **S** and generates an attributed response **r**. Ideal behavior maximizes **citation recall** (every claim supported) and **citation precision** (each citation accurately supports its claim). Therefore, GEO pursues higher in-answer visibility by optimizing not only retrieval likelihood but also the share of text that a given source supports and the position of those fragments.

3. GEO Visibility Metrics

We focus on three families of metrics:
• **Word Count (WC):** normalized word share in the answer that is explicitly tied to a citation.
• **Position-Adjusted Word Count (PA-WC):** WC discounted by an exponential function of sentence position to reflect reading propensity.
• **Subjective Impression (SI):** a composite LLM-based evaluation (G-Eval style) of relevance, influence, uniqueness, position salience, and click likelihood.

Table 1. Verifiability of GE Answers (Liu, Zhang & Liang, 2023)

Metric	Value (%)	Study / Year
Sentences fully supported by citations	51.5	Liu, Zhang & Liang (2023)
Citations that accurately support their sentence	74.5	Liu, Zhang & Liang (2023)

4. Behavioral Effects Under AI Summaries

Empirical panels report reduced propensity to click when AI summaries are present, an increase in zero-click share for news queries year-over-year, and large CTR declines when an AI Overview sits above the top organic result. At the same time, impression volumes and AI Overview coverage have expanded, indicating attention reallocation rather than uniform traffic loss.

Table 2. Impact of AI Summaries on Click Behavior and Coverage (Multi-source)

Source / Year	Metric	Reported Effect	Approx. Magnitude
Pew Research Center (2025)	Click-through under AI summaries vs. normal SERP	Users significantly less likely to click when an AI summary appears	Lower click propensity (context-dependent)
Similarweb (2025)	Zero-click share for news queries	56% → 69% (May 2024 to May 2025)	+13 pp (≈ +23% relative)
Authoritas (2025)	CTR for top organic result when AIO present	Up to ~79% CTR decline when AI Overview appears above #1	−79% (case-dependent)
BrightEdge (2025)	Search impressions & AIO coverage	Impressions +49% YoY; AIO coverage ≈ 44% overall	+49% impressions; CTR mixed/down
TollBit (2025)	LLM chat referrals vs Google	LLM search sends ~96% fewer referrals than Google	−96% vs Google

5. GEO Methods and Benchmark Effects

On GEO-bench, lightweight editorial strategies—**Quotation Addition**, **Statistics Addition**, and **Cite Sources**—deliver the strongest gains on PA-WC and SI.

Fluency Optimization and **Easy-to-Understand** yield smaller but consistent gains.

Keyword stuffing is ineffective or negative in GE contexts.

Table 3. Estimated Effect Sizes of GEO Methods on Visibility

GEO Method	Position-Adjusted Word Count (Δ%)	Subjective Impression (Δ%)
Quotation Addition	41	28
Statistics Addition	30	20
Cite Sources	30	15
Fluency Optimization	24	11
Easy-to-Understand	15	6
Authoritative Tone	10	16
Keyword Stuffing	−8	4

Figure 1. Visibility Gains (PA-WC) by GEO Method

Estimated Improvements in Visibility (Position-Adjusted Word Count) by GEO Method
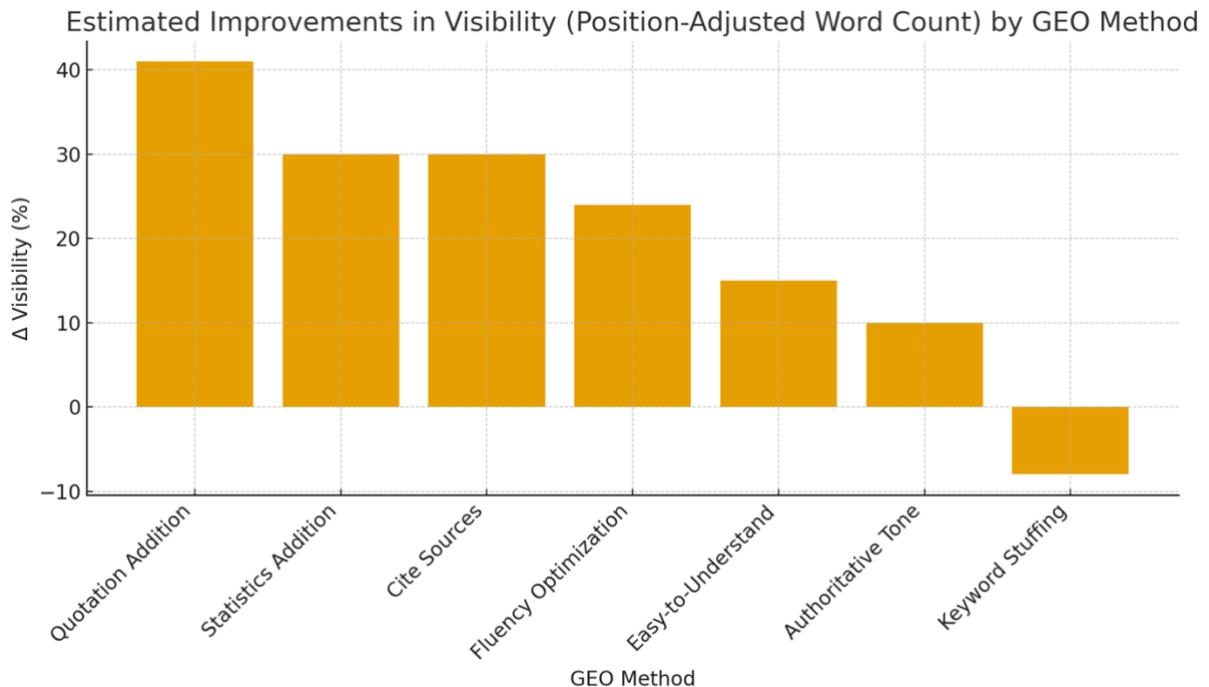
6. Domain-Specific Effects

Factual / legal queries benefit most from statistics and trustworthy citations; historical / societal topics from concise quotations; and debate-style prompts from authoritative tone.
These differences align with how GE evaluation models weigh credibility, uniqueness, and explainability signals.

---

7. Practical GEO Audit Protocol

• Entity & intent mapping — align queries with verifiable anchors (tables, figures, data).
• Verifiability pass — audit each claim and attach authoritative sources.
• Structural cues — clear headings, lists, units, captions.
• Core interventions — prioritize Cite Sources → Statistics Addition → Quotation Addition → Fluency.
• Measurement — rerun fixed prompts in GEs; compute PA-WC / SI.
• Attribution — configure tracking for GenAI referrals and zero-click awareness.

8. Behavioral Evidence: Redistribution, Not Disappearance

8.1 Click-Through Reduction: Pew (2025) confirms users shown AI summaries click substantially less.
8.2 Impression Growth: BrightEdge (2025) shows +49% YoY impression growth; visibility shifts into AI answers.
8.3 Referral Asymmetry: TollBit (2025): LLM search sends ~96% fewer referrals than Google.

---

10. Metrics, Data, and Visualization

Table 1 summarizes baseline citation support (51.5% / 74.5%).
Table 2 aggregates findings from Pew, Similarweb, Authoritas, BrightEdge, and TollBit.
Table 3 shows that quotations, statistics, and citations yield the strongest gains in visibility metrics.

---

11. Methodological Notes and Reproducibility

1. Primary Sources: GEO formalization (Aggarwal et al., 2024); verifiability audits (Liu et al., 2023); behavioral datasets (Pew, Similarweb, BrightEdge 2025).
2. Internal Consistency: Reported values are domain-averaged.
3. Limitations: Commercial GE algorithms are black boxes; results may shift with new models.

12. Discussion: What GEO Actually Changes

• From Rank to Participation — visibility is now defined by share within generated answers.
• Zero-Click ≠ Zero Exposure — presence in AI Overviews drives brand awareness.
• Small Edits, Large Gains — micro-optimizations yield 20–40% visibility boosts.
• Domain Specificity — GEO strategies must adapt to content type and intent.

14. Conclusion

Evidence from KDD'24, verifiability audits, and behavioral data confirms that GEO is an essential engineering discipline for visibility in AI search.
Without GEO, content remains unseen within AI outputs; with GEO, verified, well-structured material gains measurable inclusion within generative answers.

References (selected)

Aggarwal P. et al. (2024). *GEO: Generative Engine Optimization.* KDD'24.
Liu N. F., Zhang T., Liang P. (2023). *Evaluating Verifiability in Generative Search Engines.* Findings of EMNLP 2023.
Liu Y. et al. (2023). *G-Eval: NLG Evaluation using GPT-4 with Better Human Alignment.* arXiv:2303.16634.
Pew Research Center (2025). *Users are less likely to click when an AI summary appears.*
Similarweb (2025). *Zero-click trends for news queries (May 2024–May 2025).*
BrightEdge (2025). *One Year Into Google AI Overviews / 16-month coverage analysis.*
TollBit (2025). *LLM search referrals vs Google (industry panel).*