90 Days of Measured AI Citations: What Answer Engines Actually Retrieve

Server-side citation data from a live, instrumented three-site portfolio — including a working local-service business — measured from the only vantage point that can't be gamed: the server's own access logs.

Published July 12, 2026 · Updated monthly · Window Feb 23 – Jul 12, 2026 · Drew M., IEO Engine™ (USPTO Serial 99676324) · ieoengine.com/research/measured-ai-citations.html

Every number in this report was measured, not estimated. It comes from reading raw server access logs on three live websites — a Sarasota pressure-washing company, a consumer photo app, and a methodology site — and identifying, verifying, and counting each time an AI answer engine pulled a page to build a response for a real user. Nothing here is a survey, a projection, or a vendor's dashboard. It is a ledger.

Finding 01 — The Zero-Click Gap Is Real, and It Is Enormous

Instrument (same 24 hours, Jul 11, all three sites)	Reading
Google Search Console	321 impressions · 0 clicks · 0.0% CTR
Server-side citation ledger	~821 answer-surface retrievals

> *In a single 24-hour period, AI answer engines retrieved these sites' pages roughly 2.5 times more often than Google showed them to human searchers — and Google delivered zero clicks.*

Finding 02 — Each Platform Has a Different Clock

Platform	Round trips	Median ingestion→citation lag
Perplexity	3	11 minutes
OpenAI / ChatGPT	10	9.2 days
Google (AI answers)	16	20.3 days
All platforms	30	6.9 days

Perplexity's near-instant loop (n=3 and growing) makes it the fastest feedback instrument in the industry. Google's answer layer is the slowest to adopt — but holds what it adopts.

Finding 03 — Rank and Citation Are Nearly Unrelated (r = 0.29)

Query (local-service site)	Google position	Times cited by AI
"roof cleaning lido key"	3.0	0 — never
"roof cleaning bayshore gardens"	5.3	0 — never

Query (local-service site)	Google position	Times cited by AI
"hoa paver sealing siesta key"	43	6

> *Pages in Google's top 5 that have never been pulled into an AI answer, while a page ranked #43 was cited six times. Answer engines select for clear, extractable answers backed by consistent facts — not for rank.*

Finding 04 — What Answer Engines Want From a Local Business

Page class	AI retrievals
Cost & pricing guides	205
Homepage	202
Long-form articles	96
Service & location pages	70

"How much does X cost" is the query class answer engines are built to serve. The traditional service-page-per-town architecture — the backbone of local SEO — is the least retrieved class on the site.

Finding 05 — Being Crawled Is Not Being Used

Across the portfolio, 1,660 AI-crawled pages have never once been retrieved for an answer. Documented extreme: on July 11, one platform's crawler ingested an entire 340+ page content library in 23 minutes, every request verified against the operator's published IP ranges. Answer-surface retrievals since: zero, clock running. Getting eaten is free. Getting used is earned.

Finding 06 — The Training Channel Is Separate From the Citation Channel

On July 11, Common Crawl — the dataset major LLM training corpora are built from — performed a verified crawl of the local-service site: 181 pages in 104 minutes, robots.txt honored. Its public index confirms 23 pages already in the May 2026 corpus, including two PDFs. Content taken by Common Crawl doesn't get quoted — it gets learned. And because it discovers by following links, it missed four of the site's six most-retrieved pricing pages: internal linking still decides what enters the training corpus.

Finding 07 — You Cannot Trust a User-Agent String

During this window the instrument caught live impersonation: a rented cloud VM announcing itself as an Anthropic answer agent, others claiming to be Google's and Anthropic's crawlers, and a probe session cycling user-agent variants to test defenses. Every count in this report is therefore verified against operator-published IP ranges (OpenAI, Google, Perplexity, Anthropic — 1,100+ prefixes, refreshed weekly) or forward-confirmed reverse DNS. A user-agent is a claim; the network is the evidence. These are floor numbers, not estimates.

METHOD — NAMED, NOT DISCLOSED

Measurements are produced by a private server-side instrument: citation events detected by proprietary log forensics, de-duplicated so multi-page fetches within one answer count once, filtered of operator traffic, tagged and excluded

during operator self-testing, and identity-verified as in Finding 07. Detection mechanics, scoring weights, and instrumentation fingerprints are deliberately not documented: publishing how the measurement works would let subjects game the measurement. Public totals update in delayed batches for the same reason. Findings are published; the instrument is not.

Cite as: Drew M., "90 Days of Measured AI Citations," IEO Engine Field Reports Nº 1, July 2026. ieoengine.com/research/measured-ai-citations.html — IEO Engine™, Intelligent Engine Optimization, USPTO Serial 99676324.