SEO FOUNDATIONS

How search actually works in 2026: crawling, rendering, indexing, and ranking, the technical layer a developer controls, and what the AI-reshaped results page changed. Written to be read by a human or handed to an AI whole.

SELWYN UY · SELWYNUY.DEV

EDITION 2026.07

CONTENTS

IN THIS GUIDE

01 — THE STATE OF PLAY

SEARCH IS REPRICING, NOT DYING

IN PLAIN WORDS

Fewer people click through to real websites now, because Google and AI chatbots answer a lot of questions right on the results page. Search is not dying, it is changing shape. The new goal is to get your page mentioned inside those answers, and fast, plain, well-built pages are what make that possible.

SEO is going through its largest structural shift since mobile. SparkToro's 2026 clickstream study (Similarweb panel) found that 68.01% of US Google searches ended without a click to the open web in early 2026, up from 60.45% in 2024. Only 276 of every 1,000 searches now send someone to a site that is not Google's own. On queries where an AI Overview appears, Ahrefs measured a 58% lower clickthrough rate for the top-ranking page across 300,000 keywords.

That is the loss column. The other column: Seer Interactive's study of 5.47 million queries showed organic CTR on AI Overview queries rebounding 85% between December 2025 and February 2026, and pages cited inside the AI answer earning roughly 120% more organic clicks per impression than uncited pages. The channel is not dying, it is repricing. Raw traffic per ranking falls; citation becomes the new position one.

68%

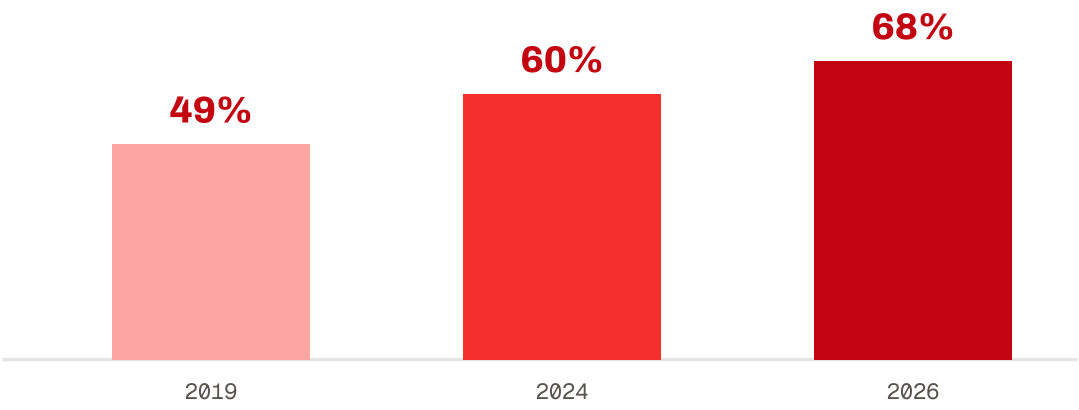
US Google searches ending with no open-web click, Jan-Apr 2026, up from 60.45% in 2024 (SparkToro / Similarweb panel, 2026).

-58%

CTR for the #1 result when an AI Overview is present, across 300K keywords (Ahrefs, Dec 2025 data).

+120%

More organic clicks per impression for pages cited inside the AI Overview vs uncited (Seer Interactive, 2026).



Share of US Google searches ending with no click to the open web (60.45% in 2024, 68.01% in 2026). SparkToro / Similarweb clickstream, 2026.

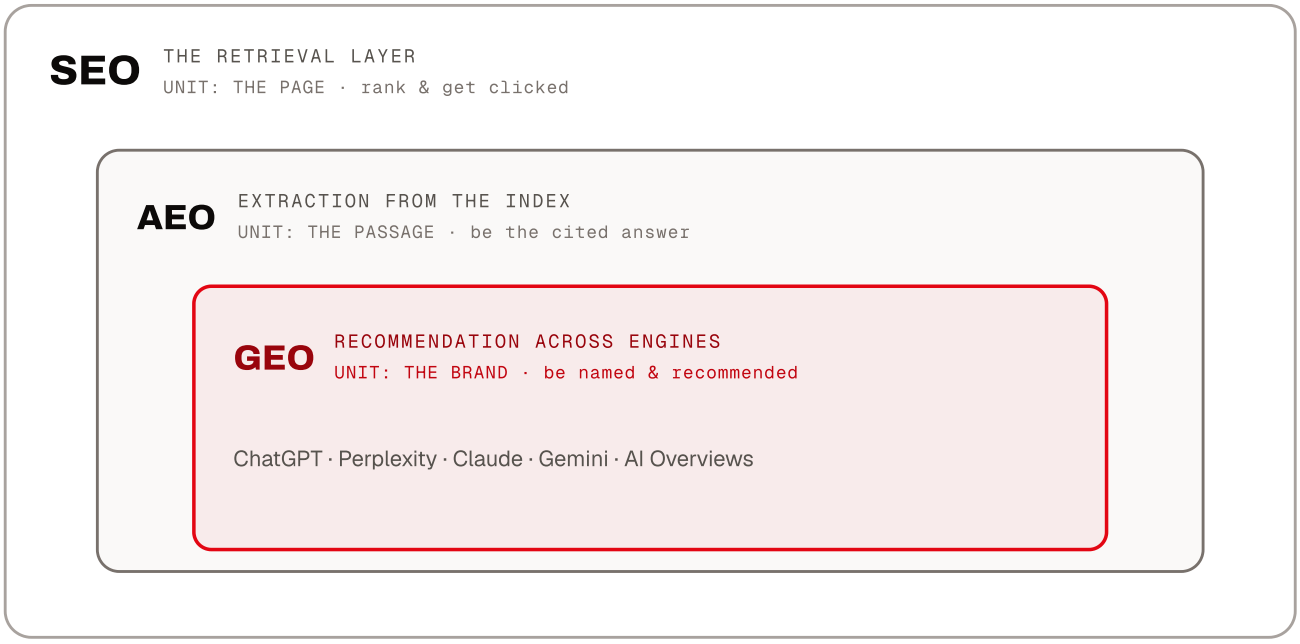
For a developer this is the moment technical fundamentals stop being commodity checklist work. Server-rendered HTML, clean indexing controls, fast pages, and clear entities now decide whether a site is retrievable by classic ranking and by AI answer engines at the same time. Google's own May 2026 guidance says answer and generative optimization "is still SEO". The discipline transfers; understand it deeply and skip the hacks.

02 — THE MAP

ONE FOUNDATION, THREE LAYERS

IN PLAIN WORDS

Three fancy names, one idea. SEO gets your page found by search engines, AEO makes one piece of your page the exact answer a search or chatbot quotes, and GEO gets your brand named and recommended by AI tools. They stack on top of each other, so you build the base first and the rest has something to stand on.



Concentric layers, not rivals: AEO builds on SEO, GEO builds on both. Each layer changes the unit you optimize.

SEO, AEO, and GEO are concentric layers on one shared foundation, not rival disciplines. **SEO** is the retrieval layer everything else sits on: every major answer and generative engine retrieves from a conventional crawler-built index. Google's index feeds AI Overviews, AI Mode, and Gemini; Bing's feeds ChatGPT search and Copilot; Perplexity runs its own crawler. **AEO** sits on top and changes the unit of optimization from the page to the passage: being the specific chunk an answer surface extracts and attributes. **GEO** sits on top of both and changes the unit again, from the passage to the brand: what generative engines already know about you, plus earned presence in the third-party sources they retrieve and cite.

A page cannot win AEO while failing SEO basics, and GEO without AEO-shaped content gets retrieved but not cited. AI Overviews are the surface where all three meet: SEO makes you retrievable, AEO makes you the extracted passage, GEO makes you the named and recommended brand.

READ THIS GUIDE IF

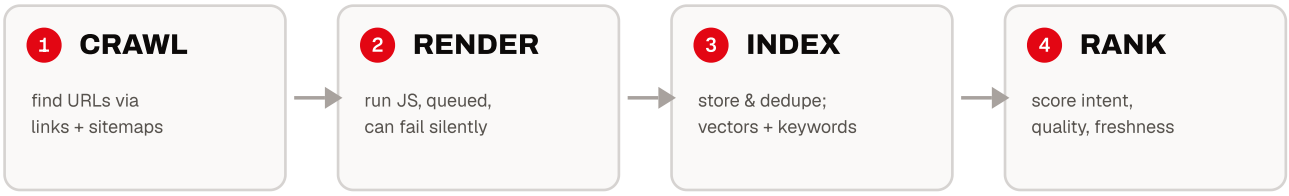
You need to be retrievable and ranked. If you already rank but are not the extracted answer, read **AEO Foundations**. If you need LLMs to name and recommend your brand, read **GEO Foundations**. All three at selwynuy.dev/guides.

03 — THE MACHINE

CRAWL, RENDER, INDEX, RANK

IN PLAIN WORDS

Before Google can show your page to anyone, four things have to happen. A Google robot visits the page, loads it like a web browser would, files it away in a giant library, then decides where it should appear when someone searches. If any of those four steps breaks, your page simply never shows up.



Anything you want found must be in the server HTML: the render step is queued and can fail silently. Source: Google, How Search Works.

Stage 1: Crawling

There is no central registry of URLs. Googlebot discovers pages through links from known pages and submitted XML sitemaps. Crawling is budgeted: crawl capacity (how hard Google can hit your server) times crawl demand (how much it wants your URLs, driven by popularity and staleness). Budget is a non-issue below roughly a million URLs, but infinite URL spaces like faceted navigation, calendar pages, and session parameters can burn it on any site.

Stage 2: Rendering

Googlebot works in two waves: it parses the raw HTML response first (links, metadata, visible text), then queues the page for the Web Rendering Service, an evergreen Chromium that executes JavaScript. The queue delay is variable and rendering can fail silently. The safe engineering position: everything you want indexed lives in the initial HTML response.

Stage 3: Indexing

Google canonicalizes duplicates (one URL per document cluster), extracts content and annotations, and stores documents in an index that pairs the classic inverted index with dense vector representations. Semantically related pages are retrievable without exact keyword overlap.

Stage 4: Ranking and serving

Query understanding models interpret intent; hundreds of signals score topicality, quality, freshness, usability, and context. Two non-documentation sources filled in the picture: in the DOJ antitrust trial, Google VP Pandu Nayak confirmed under oath that Navboost, fed by roughly 13 months of aggregated click data, re-ranks results, and the

2024 Content Warehouse leak showed fields like goodClicks, badClicks, and lastLongestClicks. Leaked fields prove the data is collected, not how it is weighted today.

RULE

Google measures whether your result satisfied the click. Titles and snippets must honestly match page content, and the page must resolve the query fast. A visitor bouncing back to the results page is evidence against you.

04 — YOUR LAYER

TECHNICAL SEO, MAPPED TO NEXT.JS

IN PLAIN WORDS

This is the plumbing part of SEO, the part you fully control in the code. It means making sure your words are actually inside the file the server sends out so robots can read them, telling Google which web address is the real one, and giving every page its own clean title. Get this plumbing right and everything else you do has a chance to work.

Rendering strategy is decision number one

Content must exist in the server HTML response. React Server Components, SSG, and ISR all produce indexable HTML by default. Pure client-side rendering, a "use client" page fetching data in useEffect, gambles on the render queue. Verify it the mechanical way: curl the URL and grep for your content, compare raw vs rendered HTML in Search Console's URL Inspection, and crawl the site twice in Screaming Frog (JS rendering off, then on) and diff the results.

Metadata and indexing controls

Use `generateMetadata` for title, description, canonical (`alternates.canonical`), robots directives, and Open Graph on every dynamic route. Ship `app/sitemap.ts` and `app/robots.ts` . Return real HTTP 404s with `notFound()` ; a 200-status "not found" component is a soft 404 and gets deindexed unpredictably. Moved content redirects with 301/308 in `next.config` or middleware, never a client-side `router.push` .

Duplication and canonicalization

One canonical URL per document: pick a trailing-slash policy, lowercase paths, strip tracking parameters. Faceted navigation needs a deliberate policy (canonicalize filter combinations to the base category, or noindex them). For i18n, emit hreflang via `alternates.languages` and link every language version to all others.

Crawl hygiene

robots.txt controls crawling, not indexing: a blocked URL can still be indexed from links, and a noindexed page must stay crawlable for the directive to be seen. Paginate with real `<a href>` links, not button-only onClick handlers, because Googlebot does not click. Monitor with Search Console's Page Indexing report; "Crawled, currently not indexed" and "Duplicate without user-selected canonical" are the rows that matter.

05 — SPEED AS A SIGNAL

CORE WEB VITALS

IN PLAIN WORDS

Google measures three things about how your page feels to a real visitor: how fast the main content shows up, how quickly the page reacts when someone taps or clicks, and whether things jump around while it loads. Keep all three quick and steady, because a slow, jumpy page annoys people and quietly loses you visitors.

LCP	INP	CLS
≤ 2.5s	≤ 200ms	≤ 0.1
Largest Contentful Paint	Interaction to Next Paint	Cumulative Layout Shift
loading · at p75	responsiveness · at p75	visual stability · at p75

The "good" thresholds, all measured on real Chrome users at the 75th percentile. INP replaced FID in March 2024.

INP is the metric JavaScript-heavy sites fail most, and the one a React developer most needs to understand: long main-thread tasks from hydration, heavy event handlers, and third-party scripts all push it past the 200 ms line.

Honest framing on ranking impact: Google's documented position is that page experience acts more like a tie-breaker among comparably relevant results than a dominant factor. Do not promise ranking jumps from a Lighthouse score. The stronger arguments are that slow pages compound into abandonment and unsatisfied clicks, and conversion rates measurably track speed.

FIELD, NOT LAB

Rankings use field data (CrUX, the Chrome UX Report), not your local Lighthouse run. A perfect lab score can coexist with a failing real-user assessment. Check PageSpeed Insights (it shows both), the Search Console CWV report, and your own RUM via the web-vitals package or Vercel Speed Insights.

The Next.js playbook, by metric:

- **LCP:** next/image with `priority` on the above-the-fold hero, never lazy-load the LCP element, next/font to self-host fonts and kill render-blocking font CSS.

- **INP:** move logic into Server Components to shrink client bundles, code-split with next/dynamic, defer third parties with next/script's lazyOnload, break long tasks with startTransition.
- **CLS:** explicit dimensions on all images and video, reserve space for ads and embeds, animate with CSS transforms only.

06 — THE CONTENT SYSTEM

WRITE, PUBLISH, SCALE

IN PLAIN WORDS

This is how you actually make pages that people and search engines want. Start with what a person is really trying to do when they search, answer their exact question near the top, and add something no other page has, like your own numbers or screenshots. Then turn it into a simple weekly routine so your pages build on each other instead of standing alone.

Start from intent, not keywords

Every query has a dominant intent (informational, commercial investigation, transactional, navigational) and the live SERP is the spec: if the top results are comparison tables and Reddit threads, a 3,000-word essay is the wrong format. The workflow: seed terms from Search Console's real query data, expand with a keyword tool, cluster keywords by shared SERP results (pages rank for clusters, not single keywords), then map one primary cluster to one URL so pages never cannibalize each other.

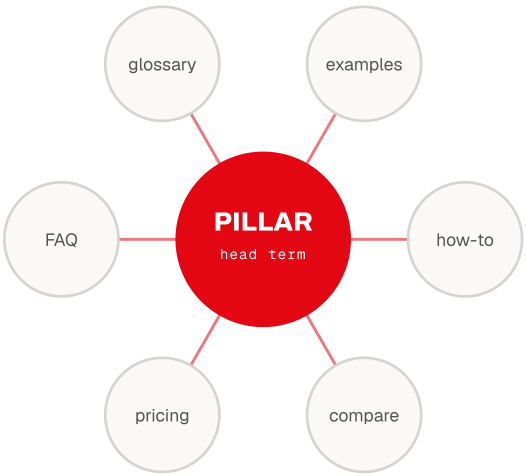
The on-page elements that still matter

Title tag: front-load the topic, keep it around 50 to 60 characters, expect Google to rewrite weak ones. Meta description: no ranking weight, but it is your ad copy for CTR. One H1; descriptive H2s and H3s that mirror the subquestions users ask, because those headings are exactly what snippets and AI systems extract. Readable slugs. Descriptive anchor text on internal links, never "click here".

Information gain is the 2026 differentiator

Generative systems compress consensus content into answers, so a page that restates the top ten results is structurally worthless. Add what does not exist yet: original data (run the benchmark, survey your users, publish the numbers), first-hand experience with screenshots and failure cases, expert quotes, and opinions with reasoning. This is also what Google's helpful content guidance literally asks for: people-first content demonstrating first-hand expertise.

Architecture



Topic clusters: one pillar for the head term, a spoke per subtopic, dense links both ways. Internal linking is the most controllable authority lever you have.

Topic clusters: a pillar page for the head term, spokes for subtopics, dense bidirectional internal links with descriptive anchors. Internal links are the most controllable authority lever a developer has. Every important page within three clicks of the homepage, zero orphans.

NAMED SPAM POLICIES

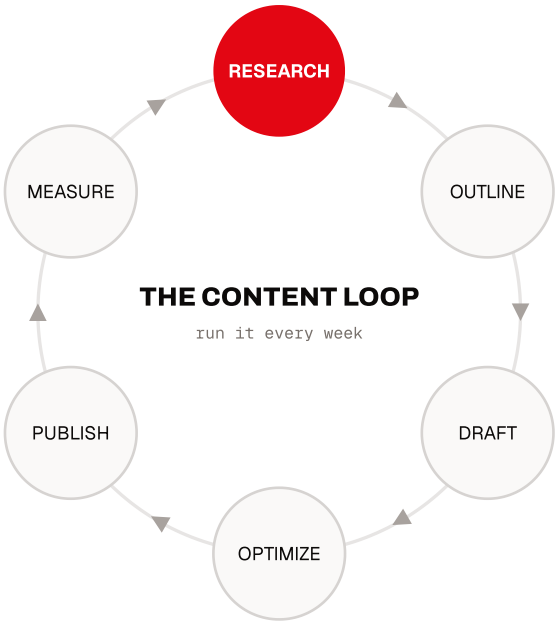
Scaled content abuse (mass-produced pages without value, AI or human), site reputation abuse (parasite content on rented subdomains), expired domain abuse, and doorway pages are all explicitly named in Google's spam policies. Raters are instructed to score low-effort mass-produced content Lowest.

How to write a page that ranks and gets cited

The skeleton of a page built to rank and be cited: one H1, a byline for E-E-A-T, an answer-first intro, question-led sections, your own data, internal links, and schema.

Write the page in that order and it satisfies a human, a ranking system, and an answer engine at once. Lead with the answer, not the backstory. Phrase headings as the questions a reader actually types, then answer each in the first 40 to 60 words before you elaborate. Put one thing on the page that exists nowhere else, your own number, test, or screenshot, because a page that only restates the top ten is worthless to a model that already summarized them. Attach a real byline and date, link out to your primary sources, and link in to three related pages. That is the whole craft; the rest is doing it repeatedly.

The production loop



A repeatable weekly cycle anyone can run: research a question, outline to the live SERP, draft answer-first, optimize, publish, measure, then feed what you learn into the next page. The system is the loop, not any single post.

This is the part that makes it a system rather than a hobby. Each stage is small and handoff-able: research picks one question from the Search Console query data, outlining reads the top three live results and matches their format, drafting fills the page template above, optimizing runs the checklist, and measuring waits a month before judging. Run it once a week and the cluster compounds; run it once and quit, and you never leave the dip.

The pre-publish checklist

The optimize stage in one page. This is the SOP to hand a writer or VA; nothing ships until every box is true.

PRE-PUBLISH CHECK	WHY IT MATTERS
One H1, topic front-loaded, under 60 characters	The title is your click and your topic signal
Every H2 phrased as a real question	Headings are exactly what snippets and AI lift
A 40-60 word answer directly under each question	The extractable unit for snippets and AI Overviews
One thing no competitor has (data, test, screenshot)	Information gain is the 2026 differentiator
Author byline and a visible date	E-E-A-T: proof a real, credible person wrote it
Links out to 3+ related pages, descriptive anchors	Internal links are your top controllable authority lever
Linked in from 3+ existing pages	A new page with no internal links is an orphan
Canonical set, alt text on every image	Avoids duplication dilution, captures image search
JSON-LD valid in the Rich Results Test	Entity clarity for the Knowledge Graph and AI
Content present in the server HTML (curl and grep)	Client-only content may never be crawled

How to scale

Scaling is not writing more; it is running the loop more times without dropping the bar. Three levers. **Cluster, don't scatter:** pick one topic, publish the pillar and its spokes as a batch, and interlink them, so each new page compounds the cluster instead of standing alone. **Templatize:** the page anatomy and the checklist above are the SOP; hand them to a writer or VA and the output stays consistent without you in every draft. **Gate on quality, not volume:** one rule keeps scale from becoming the scaled-content abuse Google penalizes, which is that every page must pass the information-gain check. A page that cannot pass it does not ship, however fast you need the volume.

SMALL MOVES, BIG IMPACT

The highest return per hour, in order: add a 40-60 word answer to the top of your ten best-ranking pages; internal-link every new post from three older ones; refresh the date and the stats on last year's top page; and give your five money pages a real author byline. None takes an afternoon; each one moves a needle.

07 — TRUST

E-E-A-T AND CORE UPDATES

IN PLAIN WORDS

Google wants proof that a real, trustworthy person who actually knows the topic wrote your page, shown through things like a named author, a short bio, and a visible date. E-E-A-T is just shorthand for experience, expertise, authority, and trust. A few times a year Google reshuffles all its rankings in what it calls a core update, and after a drop the fix is usually better content over time, not a quick tweak.

E-E-A-T is not a score and not a direct ranking factor. It is the rubric the roughly 16,000 contracted quality raters use to evaluate results, and their ratings calibrate the machine-learned quality systems, so it shapes what the algorithms are trained to reward. Four letters, with trust as the umbrella the other three feed:

E — **Experience**

First-hand proof you actually used or did the thing. Added in 2022; it matters most for reviews and how-to content.

E — **Expertise**

Depth of knowledge on the topic, shown by who wrote it. A named author tied to a real bio page is the signal.

A — **Authoritativeness**

Whether others in the field treat you as a go-to source. Off-site corroboration outweighs anything you claim yourself.

T — **Trustworthiness**

The umbrella: accurate, transparent, safe. Visible dates, an editorial policy, and honest sourcing build it.

YMYL (Your Money or Your Life) topics carry the strictest bar. The September 2025 rater guidelines update expanded the category to "YMYL Government, Civics & Society" and added rating examples for AI Overview responses. Signals a developer can actually build: real author bylines linked to bio pages, Person and Organization schema tying authors to their sameAs profiles, visible dates, an editorial policy, about and contact pages, and citations to primary sources. Off-site corroboration outweighs on-site claims: independent evidence that the entity is who it says it is.

Core updates: cadence and recovery

2025 shipped three broad core updates (March, June, December); 2026 has already shipped March core, March spam, and May core updates. Google's own recovery doctrine: after a core update there is usually nothing to "fix"; the update re-assessed content broadly, improvements take months, and recovery typically registers at a subsequent core update. The triage protocol: annotate update dates in analytics, segment losses (site-wide vs template vs topic), compare winners and losers in your SERPs, then invest in content quality and pruning rather than technical micro-tweaks.

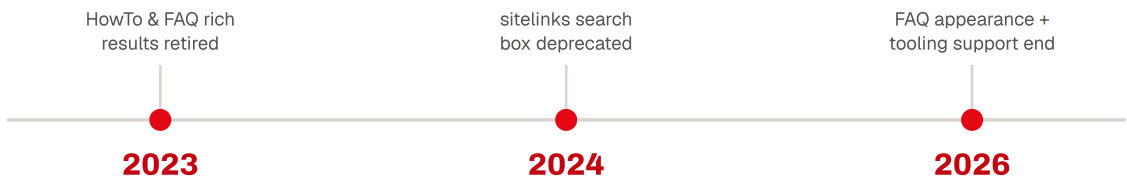
08 — MARKUP AND AUTHORITY

STRUCTURED DATA AND LINKS

IN PLAIN WORDS

Structured data is a hidden bit of code that spells out plainly what your page is, like "this is a product" or "this is a recipe", so machines understand it at a glance. Links are other websites pointing to yours, once treated as gold but now worth far less than simply being mentioned by name. Spend your effort on clear labels and honest mentions, not on chasing a pile of links.

Structured data: fewer rich results, a second job



What Google has removed. Unsupported markup earns nothing, but is not penalized. Ship the schema that still returns a result.

Use JSON-LD, rendered in a Server Component from the same data that renders the page, so the markup never drifts from what a visitor sees. The catch is that most published schema advice is stale: the rich results it promises were retired years ago (above). Unsupported markup is not penalized; it simply does nothing.

What still earns visible rich results: Product, Review, Article, Recipe, Event, JobPosting, LocalBusiness, VideoObject, BreadcrumbList, Organization, Person. The second job in 2026 is entity disambiguation: schema feeds the Knowledge Graph and helps machine consumers, including AI systems, resolve who and what your pages are about. Validate with the Rich Results Test and the schema.org validator.

Links: overrated, not dead

Google's Gary Illyes said links are no longer a top-3 ranking signal, and in 2024: "We need very few links to rank pages." Google's documentation demoted links from "an important factor" to "a factor" in March 2024. What still works: a few relevant, editorially given links from authoritative pages; digital PR built on original data that journalists cite; and unlinked brand mentions, which increasingly matter because LLM-based systems learn brand-topic associations from plain text. The GEO guide carries the anchor dataset here: across 75,000 brands, branded mentions correlate with AI visibility at 0.664 versus 0.218 for backlinks. What fails: paid link networks and exchanges at scale. Internal linking remains the highest-ROI link work a developer can do.

09 — THE NEW SERP

THE AI-RESHAPED SERP

IN PLAIN WORDS

SERP just means the search results page, the list you see after you search. Now AI often writes its own answer at the top and pulls quotes from several pages to build it, sometimes for questions your page never set out to answer. Write clear sections that each fully answer one question, and you become one of the pages it quotes.

AI Overviews (generative answers above organic results) appeared on somewhere between 16% and 48% of queries in 2025-2026 depending on the tracker: Semrush's 10-million keyword panel saw 6.49% in January 2025, a 24.61% peak in July, and roughly 15.69% by November, while BrightEdge's B2B-heavy set reports around 48%. Both are methodology artifacts; the trend is what matters. AI Mode, the fully conversational tab, rolled out to all US users in May 2025.

The key mechanism is **query fan-out**: the system decomposes one query into many synthetic sub-queries, retrieves passages (not whole pages) for each, and synthesizes a cited answer. Full mechanics live in the GEO guide. Three consequences: pages get cited for queries they never targeted; passage-level clarity wins, so write self-contained sections under descriptive headings that lead with the answer; and covering adjacent subquestions increases your retrieval surface. Two independent studies now converge on roughly a third of AI Overview citations coming from top-10 organic results (Ahrefs: 37.9% in early 2026, down from 76% in mid-2025; Surfer: 32%), which means citation retrieval is a partially separate game from classic ranking. The craft of winning extraction is the AEO guide's territory.

GOOGLE'S OFFICIAL RULES

No special requirements beyond being indexed and snippet-eligible. The controls are the existing ones: nosnippet, data-nosnippet, max-snippet, noindex. Google-Extended opts out of Gemini training, not Search AI features. And llms.txt is consumed by no major engine: of roughly 38,000 domains with a valid file, 97% saw zero requests for it in the measured month (Ahrefs). The AEO guide carries the full debate.

10 — INSTRUMENTS

MEASUREMENT IN 2026

IN PLAIN WORDS

These are the free tools that tell you whether any of this is working, like Google Search Console showing which searches led people to you. The big change is what to count: instead of only raw clicks, watch whether AI answers mention you, whether more people search for your brand by name, and whether visitors actually turn into customers.

Search Console is ground truth for queries, impressions, and clicks: segment branded vs non-branded with regex filters, and watch for impressions-up-clicks-down divergence, the signature of AI Overview exposure (AI feature traffic is folded into the "Web" type with no breakout). GA4 for landing-page sessions and conversions, with referral segments for chatgpt.com, perplexity.ai, and gemini.google.com. CrUX and PageSpeed Insights for field vitals. Bing Webmaster Tools, because Bing feeds ChatGPT search. Screaming Frog or Sitebulb for audits, server logs to see what Googlebot and the AI bots actually fetch.

The KPI shift: track citations and share of voice in AI answers, branded search volume, and qualified conversions per session, not raw clicks. AI-citation tracking is a new tool category (Semrush AI toolkit, Ahrefs Brand Radar, Profound, Peec AI); the GEO guide covers the tools and the sampling methodology that keeps their numbers honest. The full stack, and how to wire GA4 and Tag Manager end to end, lives in The Stack companion.

11 — THE PLAYBOOK

DO THIS, IN ORDER

IN PLAIN WORDS

This is the whole guide turned into a checklist you can follow from the top down. It comes in three levels, from just getting found, to earning a ranking, to staying ahead, and each step names the tool to use and how you know it worked. If you are new, start at level one and do not skip ahead.

Everything above, as an ordered plan. Follow it top to bottom if you are starting out; jump to your level if you are not. Every step names the tool and how you know it worked.

LEVEL 1 STARTER

get found at all

1

Verify your site in Google Search Console.

`search.google.com/search-console` — add your domain, then confirm ownership with the DNS TXT record it gives you.

Done when: the Page Indexing report starts listing your pages (allow a few days).

2

Ship and submit a sitemap.

Add `app/sitemap.ts` and `app/robots.ts`, then submit /sitemap.xml under Search Console > Sitemaps.

Done when: the sitemap status reads "Success" with your page count.

3

Give every page an intentional title and description.

Export `metadata` (or generateMetadata) from each route.

Done when: View Source shows your own title and description, not the layout default.

4

Confirm your content is in the HTML, not just the JavaScript.

Run `curl -s https://yoursite/page | grep "your headline"`.

Done when: the headline appears in the raw response. If not, the page is client-only and crawlers may miss it.

LEVEL 2 PRACTITIONER

earn the ranking

5

Set a canonical on every indexable page.

Add `alternates.canonical` pointing at the clean URL (one trailing-slash policy, no tracking params).

Done when: URL Inspection's "Google-selected canonical" matches the one you declared.

6

Map one keyword cluster to one URL.

Pull real queries from the Search Console `Performance` report, group by intent, split or merge pages so none compete.

Done when: no two URLs rank for the same head term (that is cannibalization).

7

Fix your worst Core Web Vital.

Open `pagespeed.web.dev` on a key page, read the field data, fix the failing metric (LCP: add priority to the hero image).

Done when: the Search Console CWV report moves that URL group to "Good".

8

Add structured data where it earns a rich result.

Render JSON-LD in a Server Component for Product, Article, or Organization as applicable.

Done when: search.google.com/test/rich-results reports a valid, eligible result.

9

Wire real-user monitoring.

Install the `web-vitals` package or Vercel Speed Insights to log LCP, INP, and CLS from actual visitors.

Done when: you have a p75 field trend to watch, not a one-off lab score.

10

Audit rendering at scale.

Crawl the site in `Screaming Frog` twice, JavaScript off then on, and diff the results.

Done when: both crawls surface the same content; any gap is a render-queue dependency.

11

Publish information gain, not consensus.

Ship one page built on original data you own: a benchmark you ran, a dataset you gathered.

Done when: it earns links or citations you did not ask for.

12

Instrument the AI-search divergence.

Build a Search Console filter for impressions-up-clicks-down queries, and add GA4 referral segments for `chatgpt.com, perplexity.ai, gemini.google.com`.

Done when: you can see AI exposure and AI referrals apart from classic organic.

12 — ANTI-PATTERNS

THE MISTAKES THAT COST THE MOST

IN PLAIN WORDS

A plain list of the errors that quietly wreck your results, like hiding your words in code that robots cannot read, or chasing links that no longer count for much. Read it as a list of things to avoid, so you do not pour months into effort that works against you.

- Shipping indexable content only after client-side hydration and trusting the render queue. Anything important goes in the initial server HTML, verified with curl.
- Optimizing lab Lighthouse scores instead of 75th-percentile field data. Rankings use real users.
- Adding FAQ or HowTo schema in 2026 expecting rich results. Those displays are gone; markup that mismatches visible content violates guidelines.
- Publishing scaled AI-generated content without editorial review or added value. That is the named "scaled content abuse" spam policy.
- Soft 404s (HTTP 200 for missing pages), client-side redirects for moved URLs, and robots-blocking a page while expecting its noindex to work.
- Chasing backlink volume when Google needs "very few links". The leverage moved to digital PR, original data, brand mentions, and internal linking.

- Measuring health by rankings and raw clicks in a SERP where 68% of searches end without a click. Track citations, divergence, branded demand, conversions.
- Treating AI search as a separate discipline with hacks like llms.txt instead of extending fundamentals: indexable HTML, answer-shaped passages, entity clarity, and letting the AI crawlers you want citations from actually crawl.

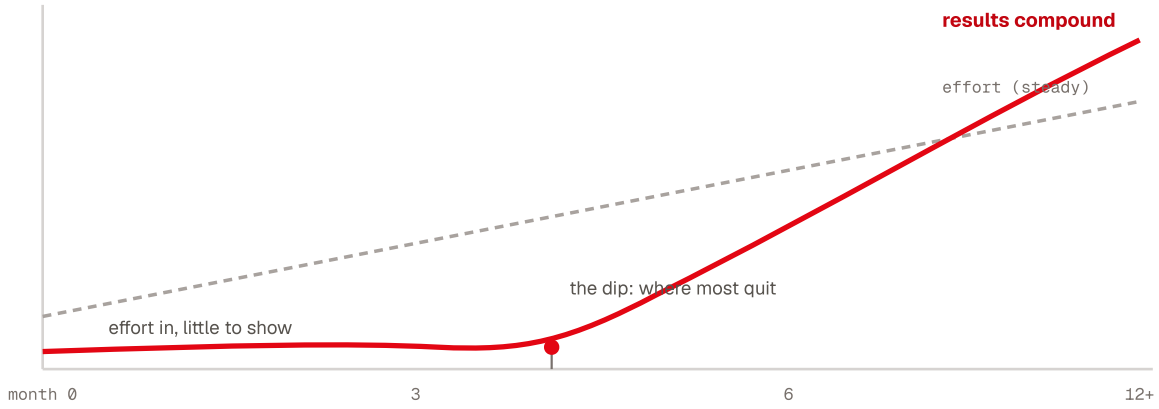
13 — THE HEAD GAME

EXPECT THE DIP

IN PLAIN WORDS

SEO barely shows results for the first few months, and that flat stretch is exactly where most people give up, right before it starts paying off. Knowing the slow start is normal is what lets you keep going. And when a ranking wobbles, do not panic and tear everything apart, wait it out.

The hardest part of SEO is not technical, it is temporal. Results lag effort by months, and that gap is where most people quit. Knowing the shape in advance is what lets you hold the line.



SEO's shape: near-flat months 1-3, long-tail movement 4-6, compounding 6-18. Over 72% of top-10 pages are 3+ years old (Ahrefs), so you compete against age, not just quality.

What to expect

Three to six months for meaningful movement, six to eighteen for a transformation. Only about 1.7% of new pages reach the top ten within a year (Ahrefs). Programs get cancelled between months three and five, almost always just before the compounding would have started. Early signals under six months are impatience talking, not failure.

Traps to avoid

1 **Quitting in the dip**

Months three to five are when the wheel feels stuck and the payoff is closest. Seth Godin's rule: push through the right dip, quit the wrong things fast.

2 **Panicking after a core update**

Do not delete content, rewrite titles, or disavow links in the first two to three weeks; rankings settle. A drop usually means another page now satisfies the intent better, not that you were penalized.

3 **Chasing vanity metrics**

Domain Rating and raw traffic give a dopamine hit while leads stay flat. That is optimization theater. Report the numbers tied to revenue.

THE CADENCE RULE

Watch for anomalies weekly (a traffic crash, a broken index); decide strategy monthly or quarterly. A keyword at #4 one week and #6 the next is normal noise, not a decline. Match how often you check to how often you can actually act.

14 — HONEST UNCERTAINTY

WHAT NOBODY KNOWS YET

IN PLAIN WORDS

An honest list of the things even the experts cannot say for sure, like exactly how much Google weighs clicks, or how big the AI answer effect really is. It is here so you can spot anyone selling you certainty they do not actually have. When studies disagree, trust the overall direction, not one dramatic number.

- **Click signals' exact weight.** Navboost's existence is court-confirmed; its current weighting is not public. Leaked fields are not live weights.
- **AI Overview prevalence and impact numbers disagree wildly** (16% vs 48% prevalence; CTR losses from 19% to 80% by methodology). Trust trends across studies, not any single stat.
- **Whether the early-2026 CTR recovery holds.** Seer's rebound may reflect a temporary UI change; Google iterates the AI Overview link treatment constantly.
- **Whether structured data influences AI citations.** Google says no special markup is needed; practitioner correlation studies are confounded by site quality. The AEO guide's schema section and the GEO guide's contested-tactics list carry this question.

- **Attribution blindness.** Search Console does not separate AI Overview clicks from classic organic, so nobody can precisely measure the effect on their own site.
- **CWV's ranking contribution** stays small and tie-breaker-like per Google, despite agencies marketing it as decisive.

Sources

- Google Search Central: How Google Search Works `developers.google.com/search/docs/fundamentals/how-search-works`
- Google Search Central: AI features and your website `developers.google.com/search/docs/appearance/ai-features`
- Google Search Central: core updates documentation `developers.google.com/search/docs/appearance/core-updates`
- Google Search Quality Evaluator Guidelines, Sept 2025 `guidelines.raterhub.com/searchqualityevaluatorguidelines.pdf`
- web.dev: Core Web Vitals thresholds `web.dev/articles/vitals`
- SparkToro: 2026 zero-click study `sparktoro.com/blog/in-2026-less-than-one-third-of-google-searches-still-send-a-click/`
- Ahrefs: AI Overviews reduce clicks by 58% `ahrefs.com/blog/ai-overviews-reduce-clicks-update/`
- Seer Interactive: AIO impact on Google CTR, 2026 update `seerinteractive.com/insights/aio-impact-on-google-ctr-2026-update`
- Semrush: AI Overviews study `semrush.com/blog/semrush-ai-overviews-study/`
- Ahrefs: 97% of llms.txt files never get read `ahrefs.com/blog/llmstxt-study/`
- Search Engine Land: links as a ranking factor (Illyes) `searchengineland.com/links-google-search-ranking-factor-gary-illyes-432422`
- Kopp Online Marketing: DOJ trial and leak analysis `kopp-online-marketing.com/what-we-can-learn-from-doj-trial-and-api-leak-for-seo`
- Aleyda Solis: query fan-out explained `aleydasolis.com/en/ai-search/google-query-fan-out/`
- Search Engine Journal: Google's guide calls AEO/GEO "still SEO" `searchenginejournal.com/googles-new-ai-search-guide-calls-aeo-and-geo-still-seo/575026/`
- Next.js Learn: rendering and ranking `nextjs.org/learn/seo/rendering-and-ranking`
