

Get cited in Google AI Overviews — a checklist (2026)

This is the checklist I run when a UK small business asks me to "get them into the AI answers". Let me be honest with you up front, because most guides will not be: Google does not publish the algorithm that decides which pages an AI Overview cites. Nobody outside Google does. So this is not a magic formula — it is informed best practice, built on the things Google *does* tell us in its public documentation, plus what I have seen actually move pages on real client sites.

A GUIDE FOR UK OPERATORS

Author

Jordan Gilbert, CTO · UK Web Marketing

Edition

v1.0 · June 2026

Audience

UK small businesses that want to be found in AI search.

TABLE OF CONTENTS

1. The checklist at a glance
 2. What AI Overviews are and how citation works
 3. Answer-first writing
 4. Authority + E-E-A-T
 5. Structure that machines parse well
 6. Technical foundations
 7. What I do NOT do — the anti-checklist
 8. Sources + further reading
-

01 The checklist at a glance

#	ITEM	LIVES UNDER
1	Each page answers one real question a person would ask	Content
2	The answer is near the top, not buried below the fold	Content
3	Copy is written the way people actually ask things	Content
4	Claims are specific — names, numbers, dates, not vagueness	Content
5	Content is genuinely useful, not written for a robot	Content
6	A named, real author is on the page	Authority (E-E-A-T)
7	The author's relevant credentials are stated	Authority (E-E-A-T)
8	First-hand experience is shown, not just claimed	Authority (E-E-A-T)
9	Primary sources are cited and linked	Authority (E-E-A-T)
10	Facts are accurate and kept up to date	Authority (E-E-A-T)
11	An "About" page makes the business and people real	Authority (E-E-A-T)
12	Headings are descriptive and in a logical order	Structure
13	Paragraphs are short; one idea each	Structure
14	Lists and tables are used where they fit the content	Structure
15	Key terms are defined plainly	Structure
16	A real FAQ section answers real questions	Structure
17	The site is crawlable — not blocked in robots.txt by mistake	Technical
18	HTML is clean and the main content is in the HTML, not only JS	Technical
19	Relevant schema.org markup is present and valid	Technical
20	The site is fast — Core Web Vitals in good shape	Technical
21	Pages are mobile-friendly and render without errors	Technical
22	Crawlers are not blocked unless that is a deliberate choice	Technical

02 What AI Overviews are and how citation works

AI Overviews are the AI-generated summaries Google shows at the top of some search results. Instead of (or above) the usual list of blue links, Google generates a short answer to the query and links out to a handful of web pages it drew on. Being one of those linked pages is the goal of this guide.

I need to repeat the honest part, because everything below depends on it.

Google does not publish the ranking or selection algorithm behind AI Overviews. Anyone who tells you exactly which factors get a page cited, with confidence, is guessing or selling. What follows is best-practice inference from Google's own published guidance plus observed behaviour — not a documented, guaranteed formula.

2.1 The fundamentals are the same as classic SEO

Here is the reassuring bit. Google's own guidance has been consistent for years: create helpful, reliable, people-first content; make your site easy to crawl and understand; demonstrate experience and expertise. Those instructions did not change when AI Overviews arrived. A page that is well-written, trustworthy, structured, and technically sound is a page that does well in classic search *and* is a sensible candidate for an AI answer. You are not learning a new game. You are playing the same one well.

2.2 What "citation" practically means

When the AI summary appears, it shows a small set of source links. To be among them, your page generally needs to: (a) be in Google's index at all, (b) clearly and directly address the question being asked, and (c) be a source Google judges trustworthy enough to put its name next to. None of those are exotic. All three are things you control through the work in the rest of this checklist.

2.3 What you cannot control

You cannot control whether a given query even triggers an AI Overview — Google decides that per query, and it changes over time. You cannot control which competitors get cited alongside you, and you cannot guarantee inclusion for any specific search. Treat AI Overview citation as an *outcome* of good fundamentals, not a target you can hit directly. Anyone promising guaranteed placement is promising something they cannot deliver.

03 Answer-first writing

The single biggest shift I make on client pages is structural: answer the question first, explain afterwards. AI systems and impatient humans both reward a page that gets to the point.

3.1 Lead with a clear, concise answer

If the page exists to answer "how much does a small business website cost in the UK?", then the first thing under that heading should be a direct answer — a sentence or two — before any preamble. A page that makes the reader scroll through three paragraphs of throat-clearing to reach the answer is a page that is harder to summarise and easier to skip.

3.2 Put the summary near the top

For longer pages, a short summary or key-takeaways block near the top does double duty: it helps a human decide they are in the right place, and it gives any summarising system a clean, self-contained statement of what the page says. Put the conclusion first; let the detail follow.

3.3 Write the way people actually ask

People do not search in keyword soup any more — they ask questions in plain language, often out loud. Write headings and copy that match real questions: "Do I need planning permission for a shop sign?" reads like a real query because it is one. Mirroring the natural phrasing of the question makes it obvious to both reader and machine that your page is about exactly that.

3.4 One question, one page

A page that tries to answer ten questions answers none of them well. Where a topic has several distinct questions behind it, consider a page (or a clearly separated section) for each. Focused pages are easier to write, easier to rank, and easier to cite as the answer to one specific thing.

04 Authority + E-E-A-T

E-E-A-T stands for Experience, Expertise, Authoritativeness, and Trust. It is not a score you can see, and it is not a direct ranking factor you can switch on — Google describes it as a set of qualities its systems try to reward. But it is the most useful lens I know for "would a reasonable person, or a careful algorithm, trust this page?"

4.1 Put a real, named author on the page

A page with no author is a page with nobody standing behind it. Name the person who wrote it, link to a short bio, and make it a real human with a real history — not "admin" or "the team". For a small business, this is often the owner, and that is exactly right: the owner is the expert.

4.2 State relevant credentials

If the author is a qualified electrician writing about consumer units, say so. Relevant qualifications, professional memberships, and years in the trade are the difference between an opinion and an authority. Do not pad it with irrelevant titles — state what makes this person credible *on this topic*.

4.3 Show first-hand experience, not just expertise

The first "E" — Experience — is about having actually done the thing. A page about fitting a wood-burning stove written by someone who fits them for a living should *read* like it: real specifics, the bit that catches people out, a photo of actual work. First-hand detail is hard to fake and easy to recognise, and it is exactly what generic AI-spun content lacks.

4.4 Cite primary sources

When you state a fact — a regulation, a standard, a figure — link to where it comes from. For UK businesses that often means gov.uk, an industry body, or a regulator. Citing primary sources makes your page more useful, more checkable, and more trustworthy, and it signals that you have done the work rather than paraphrased someone else's blog.

4.5 Keep facts accurate and current

An out-of-date price, a superseded regulation, a "2023" in the copy — each one chips away at trust. Put a visible "last updated" date on pages where facts change, and actually update them. Accuracy is the cheapest form of authority there is, and it is the one most sites neglect.

4.6 Make the business real with an About page

An About page that names the people, the location, the history, and how to get in touch turns an anonymous website into a real business. Trust is partly about being a known quantity. A site where you cannot tell who is behind it, or where they are, starts from a deficit.

05 Structure that machines parse well

A summarising system has to *understand* your page before it can quote it. Clear structure is how you make a page easy to understand — for a screen reader, for a person skimming on a phone, and for an AI pulling out the answer.

5.1 Descriptive headings in a logical order

Headings are the skeleton of the page. Use a sensible hierarchy — one H1, then H2s and H3s under it — and make each heading describe what its section actually contains. "Our process" tells a machine nothing; "How long does a website take to build?" tells it exactly what is below.

5.2 Short paragraphs, one idea each

A wall of text is hard to parse and harder to quote. Keep paragraphs short and built around a single idea. A summarising system can lift a tight, self-contained paragraph cleanly; it struggles with a 300-word block that wanders across four points.

5.3 Lists and tables where they fit

When content is genuinely a list — steps, options, features — format it as a list. When it is genuinely tabular — comparisons, specs, prices — use a table. Structured content is easier for a machine to

read accurately and is exactly the kind of clean, liftable chunk that ends up in an answer. Do not force prose into a table where it does not belong, though — fit the format to the content.

5.4 Define your terms

If a page uses jargon or a term someone might not know, define it plainly the first time. A clear, standalone definition ("PECR is the UK regulation covering cookies and electronic marketing") is useful to the reader and is precisely the kind of crisp statement a summary can quote with confidence.

5.5 A real FAQ section

A genuine FAQ — real questions people ask, answered directly — maps neatly onto how AI search works, because AI search is mostly answering questions. Keep them real: questions your customers actually ask, answered honestly. A fake FAQ stuffed with keywords helps nobody and reads as spam.

5.6 A clear question-and-answer shape overall

Across the whole page, aim for a question-and-answer shape: here is what you wanted to know, here is the answer, here is the detail. Pages built this way are easier to read, easier to mark up, and easier to cite than pages built as one long argumentative essay.

06 Technical foundations

None of the writing matters if Google cannot crawl, render, and understand the page. The technical layer is unglamorous and it is non-negotiable.

6.1 The site must be crawlable

Google has to be able to reach the page to index it, and it has to be indexed to be cited. The commonest own-goal I see is a page accidentally blocked in `robots.txt`, or left with a stray `noindex` tag from a staging build. Check both. A page Google cannot see cannot be quoted.

6.2 Clean HTML with content in the HTML

The main content of the page should be in the HTML that the server sends, not assembled entirely by JavaScript after the fact. Google can render JavaScript, but server-rendered or statically generated content is the safer bet — it is there to be read immediately, with nothing left to chance. Clean, semantic HTML (proper headings, lists, paragraphs) helps too.

6.3 Structured data with schema.org

Schema.org markup is machine-readable code that tells Google what a page *is* and what is on it. It does not guarantee anything, but it removes ambiguity. Use the types that genuinely fit:

SCHEMA TYPE	USE IT FOR
Article	Blog posts, guides, news, how-tos
FAQPage	A page (or section) of genuine questions and answers
Organization	The business as an entity — name, logo, contact, social profiles
LocalBusiness	A business with a physical location and a service area

Only mark up content that is actually present and visible on the page. Marking up content that is not there — or faking FAQ markup — is against Google's structured-data guidelines and can earn a manual penalty. Validate every block before you ship it.

6.4 Good Core Web Vitals

Core Web Vitals are Google's published page-experience metrics: loading (Largest Contentful Paint), interactivity (Interaction to Next Paint), and visual stability (Cumulative Layout Shift). A fast, stable page is better for users and is part of how Google assesses pages. Compress images, keep scripts lean, and test on a real phone on mobile data — not just your office fibre.

6.5 Mobile-friendly and error-free rendering

Google indexes the mobile version of your site. If a page is broken, cramped, or unusable on a phone, that is the version Google judges. Test pages on an actual device, fix anything that overflows or overlaps, and make sure the page renders without console errors that could block content from loading.

6.6 Do not block crawlers unless you mean to

There are legitimate reasons to block certain bots, and that is a real business decision — but make it deliberately. An overly aggressive `robots.txt`, a firewall rule that blocks Google's crawler, or a blanket block on AI crawlers will quietly remove you from the very results you are trying to appear in. Know what you are blocking and why.

07 What I do NOT do — the anti-checklist

A few opinionated negatives. These are positions I take when I am the one building a client's site — and most of them are also things Google's own guidance warns against.

- **I do not publish AI-spun spam content.** Generating a hundred thin articles with a language model to "feed the algorithm" produces exactly the unhelpful, mass-produced content Google's guidance singles out. One genuinely useful page beats fifty hollow ones.
- **I do not keyword-stuff.** Repeating "best plumber Leeds plumber Leeds cheap plumber Leeds" does not help and reads as spam to people and to Google. Write for the human; the keywords take

care of themselves when the writing is honest.

- **I do not invent author bios.** A fake "Dr. Sarah Mitchell, 15 years' experience" with a stock-photo headshot is a lie, and it is exactly the kind of fabricated authority that undermines trust. The author is whoever actually wrote it. Usually that is the business owner, and that is fine.
- **I do not fabricate "studies" or statistics.** I will not put "73% of customers prefer..." on a page unless there is a real, citable source behind it. Made-up numbers are the fastest way to destroy credibility, and they are increasingly easy to check.
- **I do not build thin doorway pages.** A pile of near-identical pages — "plumber in Leeds", "plumber in Bradford", "plumber in Wakefield" — with the town name swapped and nothing else of value is a doorway pattern Google explicitly discourages. If you genuinely serve multiple towns, give each a page with real, distinct content about that area.

08 Sources + further reading

These are real, primary sources. Where I have stated something as best-practice inference rather than documented fact, I have said so in the text — the documents below are the documented part.

- Google Search Central — documentation home (crawling, indexing, structured data, page experience) · developers.google.com/search
- Google Search Central — "Creating helpful, reliable, people-first content" · developers.google.com/search/docs/fundamentals/creating-helpful-content
- Google Search Central — how AI features in Google Search work · developers.google.com/search/docs/appearance/ai-features
- Google Search Central — E-E-A-T and the quality of your content · developers.google.com/search/docs/fundamentals/creating-helpful-content#eeat
- Google Search Central — intro to structured data · developers.google.com/search/docs/appearance/structured-data/intro-structured-data
- Google Search Central — Core Web Vitals and page experience · developers.google.com/search/docs/appearance/page-experience
- Google Search Central — spam policies (includes scaled content, doorways, keyword stuffing) · developers.google.com/search/docs/essentials/spam-policies
- schema.org — vocabulary reference (Article, FAQPage, Organization, LocalBusiness) · schema.org

A note on the long-form version

This is the v1.0 edition. The long-form version (planned for late 2026) will include worked examples — a real client page rewritten answer-first, the Article and FAQPage schema blocks I actually drop onto

small-business sites, and a before-and-after of a page that started getting cited. If you want to be told when it ships, message me at ukwebmarketing.com — same person who wrote this.