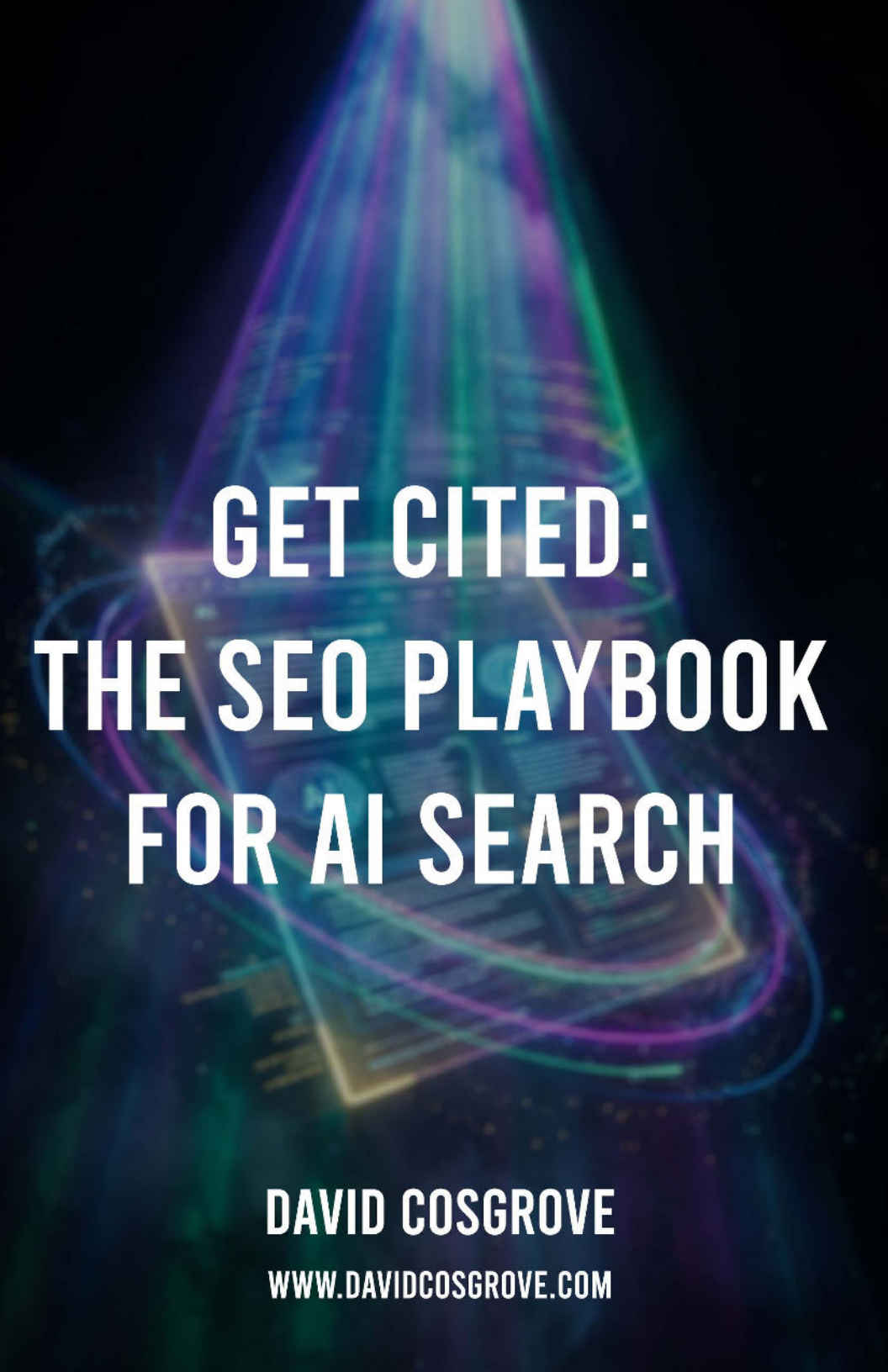

GET CITED:
THE SEO PLAYBOOK
FOR AI SEARCH
DAVID COSGROVE
WWW.DAVIDCOSGROVE.COM

Get Cited: The SEO Playbook for AI Search

How to Optimize Your Content for AI Search Systems

By David Cosgrove

30+ Years in Digital Marketing & SEO

Get Cited: The SEO Playbook for AI Search

Published by David Cosgrove Los Angeles Web Design

www.davidcosgrove.com

Table of Contents

Chapter 1: The Shift — Why Citability Matters Now

For over two decades, SEO has been about one thing: ranking. Get to page one. Climb to position one. Win the click.

That game is changing.

Google's AI Overviews now often appear at the top of search results (depending on the query, location, and user). ChatGPT, Perplexity, and other AI assistants are changing how people get answers. When someone asks a question, they increasingly get an answer—not a list of links.

This shift has profound implications for anyone who creates content online.

The End of Ten Blue Links

Traditional search showed you a list of websites and let you choose which to visit. You'd scan titles, read snippets, and click through to find your answer. The website that ranked highest got the most traffic.

AI-powered search works differently. The AI reads multiple sources, synthesizes an answer, and presents it directly to the user. Sometimes it cites sources. Sometimes it doesn't. But either way, the user often gets what they need without clicking through to any website.

This doesn't mean websites are irrelevant—far from it. The AI has to get its information from somewhere. But it means the competition has shifted. You're no longer just competing for ranking position. You're competing to be the source the AI cites when it generates its answer.

A Note on Different AI Systems

This playbook discusses Google's AI Overviews, ChatGPT, Perplexity, and other AI systems. While these platforms differ technically—different training data, different architectures, different citation mechanisms—they share a common constraint: extracting clean, well-structured answers from web content.

Google's AI Overviews pull from indexed web pages and cite sources inline. ChatGPT's browsing feature retrieves and synthesizes web content. Perplexity explicitly cites sources for its answers. The underlying need is the same: content that's easy to extract, easy to understand, and worth citing.

The principles in this playbook apply across these systems because they address that shared constraint.

Expert note (Mike King): *"SEO spent the past twenty-five years preparing content to be parsed and presented based on how it ranks for a single query. Now, we're engineering relevance to penetrate systems of reasoning*

across an array of queries." Source: https://ipullrank.com/how-ai-mode-works

The New Game: Ranking and Citation

Here's what the data shows:

Research from Ahrefs—analyzing 1.9 million citations across 1 million AI Overviews—found that 76% of citations come from pages ranking in the top 10 organic results. That's expected—authority still matters.

But here's what's interesting: 14% of citations come from pages ranking outside the top 100.

Read that again. Pages that don't even crack the first ten pages of traditional search results are getting cited in AI Overviews.

Why? Because they're structured in a way that makes them easy for AI systems to extract and cite. They answer questions directly. They define terms clearly. They present information in formats that AI can parse.

In other words: they're optimized for extraction.

Expert note: Rand Fishkin has noted that Google now answers almost two-thirds of queries without a click—meaning that if you're waiting to get traffic from Google, much of the time you've already lost. This shift means

influence has become more important than traffic alone. Source: https://sparktoro.com/blog/why-do-we-need-zero-click-marketing/

What's at Stake

If your content isn't optimized for AI extraction, you face two problems:

Problem 1: Invisibility in AI answers

When someone asks a question in ChatGPT, Perplexity, or Google's AI Overview, your content won't be among the sources cited—even if you rank well in traditional search. The AI will pull from competitors who've structured their content better.

Problem 2: Declining traditional traffic

As more users get answers directly from AI without clicking through, the overall click-through rates from search results decline. If you're not being cited in those AI answers, you're losing visibility on both fronts.

The opportunity, however, is significant. Sites that optimize for AI extraction now—while most competitors are still focused exclusively on traditional ranking—will establish themselves as go-to sources for AI systems. That advantage compounds over time.

Who This Playbook Is For

This playbook is for SEO professionals, content strategists, and website owners who want to adapt their practices for the AI search era.

You should have a working knowledge of SEO fundamentals—understanding of on-page optimization, content structure, and tools like Screaming Frog. This isn't an SEO basics course.

What this playbook provides is a systematic methodology for auditing your content's extraction-friendliness and fixing what you find. It's built on Google's own documented guidance, observable patterns from what actually gets cited in AI systems, and a decade of featured snippet optimization principles that now apply at scale.

By the end, you'll have a repeatable process you can apply to any website—yours or your clients'.

Let's get started.

Chapter 2: What Is Citability?

Citability is how easily AI systems can extract, understand, and cite your content.

> Important: Citability is not a Google metric or ranking factor. It's a practical shorthand for observable extraction-friendliness based on long-standing SEO and featured snippet optimization principles. Google has never used this term. This playbook uses it as a convenient way to describe a set of content characteristics that correlate with AI citation—not as a formal signal or score.

It's not a metric you'll find in Google Search Console or any analytics dashboard. It's a quality of your content—one that determines whether AI systems can easily use your work as source material when generating answers.

The Difference Between Good Content and Citable Content

Good content can take many forms. It might be a deeply researched longform article with nuanced analysis. It might be a personal essay with a distinctive voice. It might be a comprehensive guide that covers every angle of a topic.

Citable content is more specific. It's content that:

- Answers questions directly and immediately
- Defines terms clearly in extractable formats
- Structures information so AI can parse it
- Front-loads the most important information
- Uses specific entities (names, products, concepts) rather than vague language

Good content and citable content aren't mutually exclusive—the best content is both. But plenty of good content fails the citability test, and that's a problem in the AI search era.

Why the Same Page Can Rank But Never Get Cited

Imagine two pages about content marketing. Both rank on page one for competitive keywords. Both have strong backlink profiles and good domain authority.

Page A opens like this:

"In today's fast-paced digital landscape, businesses are increasingly recognizing the importance of connecting with their audiences in meaningful ways. As traditional advertising becomes less effective and

consumers grow more sophisticated in their media consumption habits, forward-thinking organizations are turning to innovative approaches that prioritize value creation over direct promotion..."

Five hundred words later, you finally learn what content marketing actually is.

Page B opens like this:

"Content marketing is a strategic marketing approach focused on creating and distributing valuable, relevant content to attract and retain a defined audience—and ultimately drive profitable action. Unlike traditional advertising, content marketing provides useful information that helps your audience solve problems."

Same topic. Same basic information. But Page B is citable. An AI system can extract that definition and use it directly. Page A requires the AI to wade through filler to find the substance—and often, it won't bother. It'll cite Page B instead.

The Featured Snippet Connection

If you've been doing SEO for a while, this might sound familiar. Featured snippet optimization has always rewarded content that answers questions directly and immediately.

That's not a coincidence.

Featured snippets were Google's first major step toward providing direct answers in search results. The same principles that win featured snippets—clear definitions, immediate answers, structured formats—are exactly what AI systems look for when selecting sources to cite.

Citability optimization is featured snippet optimization evolved for the AI era. The difference is scale: instead of optimizing for one snippet position, you're optimizing for how AI systems read and extract information across the board.

If you've spent years optimizing for featured snippets, you already understand most of this. The methodology in this playbook simply systematizes it for site-wide auditing.

The Four Patterns

Through testing, observation, and reverse-engineering what actually gets cited, we've identified four specific content characteristics that correlate with AI citation:

1. Immediate answers after subheadings
2. Definitional clarity with explicit definitions
3. Structured lists near relevant headers
4. Entity density in opening paragraphs

The next chapter breaks down each characteristic in detail.

Expert note: As Lily Ray reported from Jeff Coyle's presentation, citing sources, being authoritative, and offering statistics are factors that can help content appear in AI Overviews. Source: https://lilyray.nyc/seo-in-the-age-of-agents-takeaways-from-nyc-seo-week/

An Analogy: The Interview Subject

Think about what makes someone a good interview subject for a journalist.

The best sources give quotable soundbites. They speak in complete, clear statements. They provide specific facts and examples. They answer the question that was asked, directly and immediately.

The worst sources ramble. They bury their points in qualifications. They speak in abstractions. The journalist has to work hard to extract anything usable—and often, they'll just find a better source.

Writing for citability is like being a good interview subject. You're giving AI systems—and the users who rely on them—quotable, extractable, immediately useful information.

That doesn't mean dumbing down your content. It means structuring your expertise in formats that AI can work with.

Chapter 3: Four Citation-Ready Formatting Patterns

This chapter is the core of the methodology. These are the four specific content characteristics we audit for, why each matters, and how to evaluate your content against them.

Pattern 1: Immediate Answers After Subheadings

When you write a subheading like "What Is Content Marketing?" you're making a promise to the reader. They expect the next sentence to answer that question.

AI systems have the same expectation—but with zero patience for delay.

Why This Matters

AI systems scanning content for information to cite are essentially asking: "Does this page answer the question?" They look at your H2 subheadings (which often function as questions or topic statements) and check what immediately follows.

If the first paragraph after an H2 directly answers the implied question, the AI has what it needs. If the first paragraph is filler, context-setting, or "In today's world..." throat-clearing,

the AI may move on to a source that gets to the point faster.

The Buried Answer Problem

The most common citability killer is the buried answer. Writers often feel they need to provide context before delivering the main point. In traditional reading, that can work. In AI extraction, it fails.

Buried answer example:

"The question of how to optimize for search engines has evolved considerably over the past two decades. What once was a relatively straightforward matter of keyword placement has become a sophisticated discipline encompassing technical factors, content quality, user experience, and increasingly, artificial intelligence. Understanding this evolution is crucial for anyone hoping to succeed in today's competitive digital landscape. At its core, SEO is..."

The actual answer—"At its core, SEO is..."—is buried at the end. An AI system looking for a definition of SEO has to parse four sentences of preamble to find it.

Immediate answer example:

"SEO (Search Engine Optimization) is the practice of optimizing websites to rank higher in search engine results, driving more organic traffic. It encompasses technical factors,

content quality, and user experience. The field has evolved significantly as search engines have incorporated AI into their ranking systems."

Same information. But the answer comes first.

What We Extract

In the audit, we extract the first paragraph that follows each H2 subheading. This gives us a dataset of "H2 promises" and their "immediate fulfillments."

Pages where H2s are consistently followed by direct, substantive answers score well. Pages where H2s are followed by filler, questions, or vague statements score poorly.

Pattern 2: Definitional Clarity

AI systems love definitions. When they need to explain what something is, they look for content that explicitly defines terms.

Why This Matters

Definitions are the building blocks of AI-generated answers. When someone asks "What is X?" the AI needs a clear, citable definition. Pages that provide such definitions become source material. Pages that assume the reader already knows—or that define terms indirectly—get passed over.

The Trigger Phrases

Explicit definitions typically contain recognizable patterns:

- "X is a..."
- "X refers to..."
- "X means..."
- "X is defined as..."
- "X is the..."

These phrases signal to AI systems: "Here's a definition you can extract and use."

Good vs. Weak Definitions

Weak (indirect):

"When we talk about conversion rate optimization, we're really talking about making your website work harder. It's about understanding your visitors and giving them what they need. Many businesses struggle with this because they focus on traffic rather than what happens after someone arrives."

This discusses conversion rate optimization without ever defining it.

Strong (explicit):

"Conversion rate optimization (CRO) is the systematic process of increasing the percentage of website visitors who take a desired action—whether that's making a

purchase, filling out a form, or subscribing to a newsletter. CRO uses data analysis, user feedback, and testing to improve conversion performance."

This is extractable. An AI can cite it directly.

The 40-60 Word Sweet Spot

Featured snippet studies often suggest that concise definitions (commonly ~40–60 words) can perform well—long enough to be substantive, short enough to be a clean extraction. Treat this as a starting range and test by topic.

When writing definitions for citability, aim for this range. Lead with the noun being defined. Follow with a clear explanation. Include one or two specific details that add substance.

Pattern 3: Structured Lists Near Headers

Lists are inherently extractable. They present discrete points in scannable format—exactly what AI systems need when compiling information from multiple sources.

Why This Matters

AI Overviews frequently present information as bulleted or numbered lists. To generate these, AI systems look for source content that's already in list format—or that can be easily converted to list format.

Content that uses lists strategically, positioned near relevant H2 headers, is more likely to be extracted and cited than content that presents the same information in dense paragraphs.

Optimal Placement

The audit specifically looks for lists that immediately follow H2 subheadings. This positioning signals: "Here's a structured answer to the topic in this header."

Example:

H2: Benefits of Content Marketing

Content marketing delivers several key benefits for businesses:

- Builds trust and authority with your target audience
- Generates organic traffic over the long term
- Supports other marketing channels including social and email
- Can be more cost-effective than many traditional campaigns while generating qualified leads

An AI system looking for "benefits of content marketing" can extract this list directly.

When Lists Help vs. Hurt

Lists aren't always appropriate. They work well for:

- Steps in a process
- Benefits or features
- Requirements or criteria
- Examples or types
- Comparisons

They work poorly for:

- Nuanced analysis that requires explanation
- Narrative content
- Arguments that build on each other

The goal isn't to convert everything to lists. It's to use lists where they serve the content—and to position them for maximum extractability when you do.

Pattern 4: Entity Density in Opening Paragraphs

Entities are specific, named things: people, products, companies, concepts, places, technologies. Entity density refers to how many specific entities appear in a given passage versus vague or generic language.

Why This Matters

AI systems understand content through entities. They map relationships between

named things and use those relationships to determine relevance. Content rich in specific entities signals substance and specificity. Content heavy on vague language signals filler.

> Important: Entity density supports comprehension and extraction—it does not replace authority, accuracy, or trust signals. You can name-drop brands and still be wrong. The goal is specificity where it adds clarity, not keyword stuffing or false authority.

The Opening Paragraph Test

First paragraphs matter disproportionately. They're what AI systems see first when evaluating a page. A first paragraph dense with relevant entities signals: "This page has specific, substantive information."

Low entity density:

"In today's world, businesses face many challenges when it comes to reaching their customers effectively. The landscape has changed dramatically, and what worked in the past may not work anymore. Organizations need to adapt their approaches to stay competitive."

Zero specific entities. This paragraph could be about anything.

High entity density:

"Google's AI Overviews have fundamentally changed how businesses approach SEO in 2024. Companies like HubSpot, Moz, and Ahrefs are adapting their content strategies to optimize for citation in AI-generated results, moving beyond traditional ranking metrics."

Multiple specific entities: Google, AI Overviews, SEO, 2024, HubSpot, Moz, Ahrefs. The AI immediately knows what this content is about and that it contains specific, current information.

Types of Entities to Include

- Product and service names
- Company and brand names
- People (especially recognized experts)
- Technologies and platforms
- Specific concepts and methodologies
- Dates and timeframes
- Locations (when relevant)
- Industry terms and jargon (appropriately)

The key is specificity. Replace vague nouns with named things wherever accuracy allows.

Expert note: Mike King has emphasized that the query fan-out technique is a critical concept for the SEO industry to understand. AI systems like Google's AI Mode break down queries into multiple sub-queries and pull relevant passages

from across the web. Your content needs to answer not just the main query, but the related questions the AI generates. Source: https://ipullrank.com/ai-mode-webinar-recap

Bringing It Together

These four characteristics work together. A page optimized for citability:

1. Opens with an entity-rich first paragraph that signals specificity
2. Uses H2 subheadings that function as questions or topic statements
3. Follows each H2 with an immediate, direct answer
4. Includes explicit definitions using trigger phrases
5. Positions structured lists near relevant headers

The audit we'll build in Chapter 5 extracts data for each of these characteristics, letting you score any page—or an entire site—on extraction-friendliness.

But first, let's establish something crucial: the core constraints are documented by Google. The formatting tactics are practical heuristics consistent with that guidance and long-standing featured snippet learnings.

Chapter 4: What Google Actually Says (With Sources)

The methodology in this playbook isn't invented. It's derived from Google's own published guidance about how AI Overviews select and cite sources, combined with a decade of documented featured snippet optimization principles.

This chapter provides the receipts. The foundational principles in this playbook are derived from official Google documentation. The URLs are included so you can verify them yourself—and so you can confidently reference them when discussing this approach with clients or colleagues.

Google Search Central: Succeeding in AI Search

In May 2025, Google published official guidance specifically addressing how to optimize content for AI experiences including AI Overviews. The post is titled "Top ways to ensure your content performs well in Google's AI experiences on Search."

Source URL:

https://developers.google.com/search/blog/2025/05/succeeding-in-ai-search

Key quote on content quality:

> *"Focus on your visitors and provide them with unique, satisfying content. Then you should be well positioned as Google Search evolves, as our core goal remains the same: to help people find outstanding, original content that adds unique value."*

This confirms that the fundamental principle—create content that directly serves user needs—applies to AI experiences just as it does to traditional search.

Key quote on click quality:

> *"We've seen that when people click to a website from search results pages with AI Overviews, these clicks are higher quality, where users are more likely to spend more time on the site."*

This addresses a common concern about AI Overviews reducing traffic. Google's data suggests that while overall clicks may decrease, the clicks that do occur are more valuable.

Google Search Central: AI Features and Your Website

Google maintains a dedicated documentation page explaining how AI features work and what site owners need to know. This is the authoritative source for understanding the mechanics of AI Overviews.

Source URL:

https://developers.google.com/search/docs/appearance/ai-features

Key quote on AI Overview purpose:

> *"AI Overviews help people get to the gist of a complicated topic or question more quickly, and provide a jumping off point to explore links to learn more."*

This tells us exactly what AI Overviews are optimizing for: quick, clear answers. Content that states the gist quickly and clearly may have an advantage.

Key quote on query fan-out:

> *"Both AI Overviews and AI Mode may use a 'query fan-out' technique—issuing multiple related searches across subtopics and data sources—to develop a response."*

This is crucial. AI Overviews don't just match your content to a single query. They're running multiple related searches to compile comprehensive answers. Content that addresses related subtopics—not just the primary query—has more opportunities to be cited.

Key quote on SEO fundamentals:

> *"You can apply the same foundational SEO best practices for AI features as you do for Google Search overall."*

Google is explicitly saying that existing SEO principles apply. This isn't a new game with new rules—it's an evolution of the same game.

Google Search Help: Featured Snippets

Featured snippets have been the primary way Google provides direct answers in search results since 2014. The principles that govern featured snippet selection are foundational to understanding AI Overview citations.

Source URL:

https://support.google.com/websearch/answer/9351707

Key quote on selection criteria:

> *"The featured snippets come from websites that Google finds. It picks them based on how well they answer your question and how helpful they are."*

Two criteria: (1) how well the content answers the question, and (2) how helpful it is. Not how long, not how comprehensive, not how many backlinks—how well it answers and how helpful it is.

Google Search Central: Featured Snippets Documentation

Source URL:

https://developers.google.com/search/docs/appearance/featured-snippets

This documentation explains how featured snippets work from a technical perspective and confirms that there's no special markup required—Google's systems automatically identify content suitable for featured snippet presentation.

A Note on Schema Markup

A common question: "Do I need special schema markup for AI Overviews?"

Google has been clear: no special schema is required for AI Overview inclusion. The standard SEO best practices apply.

That said, schema markup like FAQPage, HowTo, and Article can reinforce structure that already exists in your content. Schema doesn't create citability—it helps search engines understand structure you've already built.

If your content has clear FAQ sections, marking them up with FAQPage schema may help Google recognize that structure. But the schema itself isn't what gets you cited—the clear, extractable content is.

Don't add schema to poorly structured content expecting it to fix citability issues. Fix the content first.

The Featured Snippet → AI Overview Lineage

It's worth understanding why featured snippet optimization principles translate to AI Overview optimization.

Featured snippets were Google's first major experiment with providing direct answers in search results. They trained users to expect immediate answers. They trained content creators to structure content for extraction. They generated enormous amounts of data about what formats work best for direct-answer presentation.

AI Overviews are the evolution of this approach. They use more sophisticated AI to synthesize answers from multiple sources, but

they're still fundamentally looking for the same thing: content that clearly and directly answers questions.

The 40-60 word definition sweet spot? That came from featured snippet research. The importance of immediate answers after headers? Featured snippets. Structured lists as preferred formats? Featured snippets.

Citability optimization isn't reinventing the wheel. It's applying a decade of documented best practices to the current search landscape.

The Honest Limitations

To be clear about what Google hasn't said:

Google has not published a "citability algorithm" or a specific formula for AI Overview inclusion. They've provided principles and guidance, not a checklist.

Google has not confirmed that the specific extractions in this audit (H2 follow-ups, definitional patterns, etc.) are explicitly what their systems look for. These extractions are derived from observed patterns and logical inference based on how LLMs process text.

Google has stated that no special schema or markup is required for AI Overview inclusion—the standard SEO best practices apply.

Why This Methodology Is Defensible

Given the above, here's why the citability audit methodology is defensible:

1. It's built on explicit Google guidance. The principles (direct answers, helpful content, good structure) come straight from Google's documentation.
2. It extends proven featured snippet practices. The specific techniques (immediate answers after headers, definitions in extractable formats) have years of documented effectiveness.
3. It's logically consistent with how LLMs work. Large language models extract and synthesize information from text. Content structured for easy extraction is more useful to these systems.
4. It systematizes observable patterns. The methodology isn't theoretical—it's based on analyzing what actually gets cited in AI Overviews.

This isn't guaranteed AI citation magic. It's systematic optimization based on everything Google has told us, combined with practical observation of what works.

Now let's get into the technical implementation.

Chapter 5: Setting Up the Audit in Screaming Frog

Screaming Frog SEO Spider is an industry-standard website crawler used by SEO professionals worldwide. One of its most powerful features is custom extraction—the ability to pull specific content from every page during a crawl.

This chapter walks through setting up the four extractions that power the citability audit.

Overview: Why Screaming Frog

Several tools can crawl websites, but Screaming Frog is ideal for this audit because:

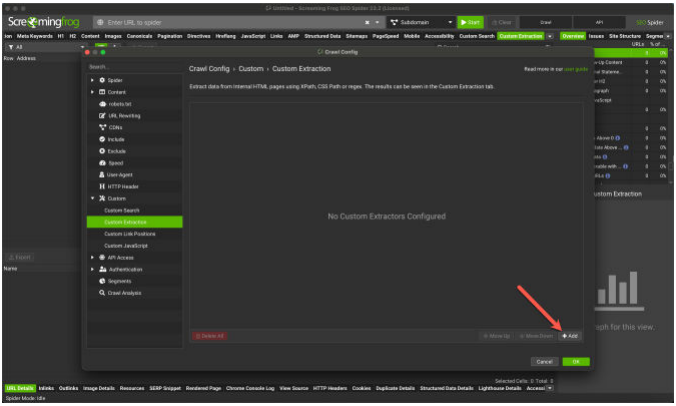
- Custom extraction supports both XPath and Regex patterns
- Results export cleanly to CSV for AI analysis
- It handles sites of virtually any size
- Most SEO professionals already have it

If you don't have Screaming Frog: the free version crawls up to 500 URLs, but Custom Extraction is a paid feature. A paid license is £199/year (plus VAT where applicable; currency varies).

Accessing Custom Extraction

In Screaming Frog, go to:

Configuration → Custom → Extraction



This opens the custom extraction configuration window. You can add up to 100 different extractions. We'll set up four.

For each extraction, you'll need to specify:

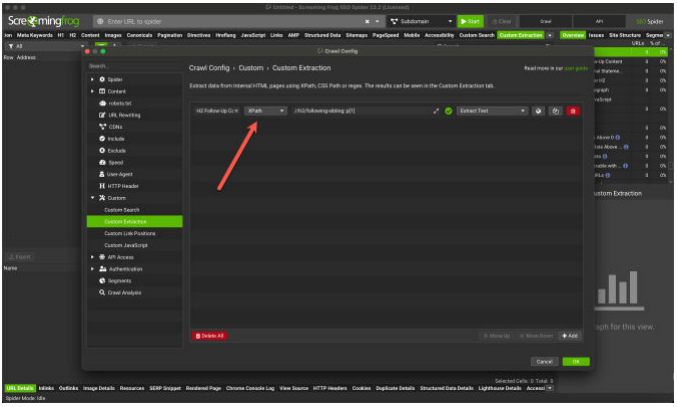
- A name (for identification in exports)
- The extraction type (XPath or Regex)
- The extraction pattern
- What to extract (we'll use "Extract Text" for all four)

Extraction 1: H2 Follow-Up Content

This extraction captures the first paragraph after each H2 subheading—letting us evaluate whether H2s are immediately followed by direct answers.

Configuration:

- Name: H2 Follow-Up Content
- Type: XPath
- XPath: //h2/following-sibling::*[1][self::p]
- Extract: Extract Text



What This Captures

For each crawled URL, this extracts the text of any paragraph that is the immediate next sibling after an H2. If the next element after an H2 is not a paragraph (e.g., a list, image, or

div), nothing is captured for that H2—that's intentional, because we're testing whether the answer appears immediately after the header.

What to Look For in Results

In your export, you'll see one row per URL. Each extraction cell can contain multiple matches (typically one per H2), often separated by line breaks. Review for:

- Filler intros ("In today's world...", "When it comes to...")
- Questions instead of answers
- Vague statements that don't deliver on the H2's promise
- Direct, substantive answers (these are your wins)

Extraction 2: Definitional Statements

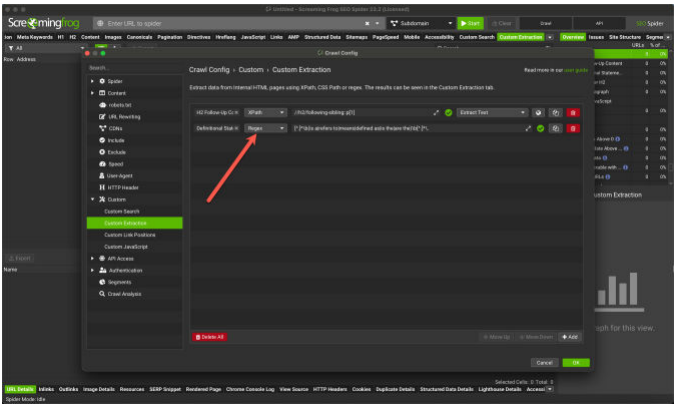
This extraction finds sentences that contain explicit definitional patterns—the kind AI systems can extract and cite directly.

Configuration:

- Name: Definitional Statements
- Type: Regex

```
[^.]*\b(is a|refers to|means|defined
as|is the|are the)\b[^.]*\.
```

- Extract: Extract Text



What This Captures

This regex finds any sentence containing the trigger phrases "is a," "refers to," "means," "defined as," "is the," or "are the." These phrases typically indicate explicit definitions.

> Note: This extraction doesn't capture all definitions—it identifies explicit, quotable definitional patterns that AI systems most easily extract. It's a heuristic, not a perfect definition detector. Some captured sentences won't be true definitions; some definitions won't be captured. The goal is identifying pages that lack explicit definitional language, not building a comprehensive definition database.

What to Look For in Results

Pages with zero definitional statements are potential problems—especially if they're about topics that should include definitions.

Pages with multiple clear, extractable definitions are strong on this signal.

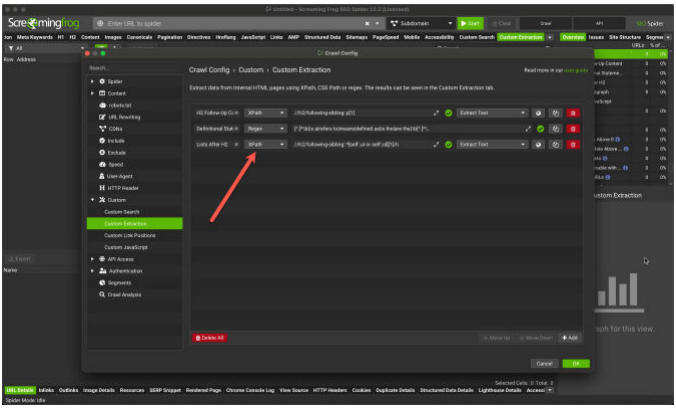
Review the definitions that are captured: Are they clear and citable? Or are they buried in longer sentences that dilute their impact?

Extraction 3: Lists After H2

This extraction captures list items that immediately follow H2 subheadings—the optimal placement for extractable structured content.

Configuration:

- Name: Lists After H2
- Type: XPath
- XPath: //h2/following-sibling::*[1][self::ul or self::ol]/li
- Extract: Extract Text



What This Captures

For each H2, this extracts the list items from a UL or OL that is the immediate next sibling after that H2. If the next element isn't a list, nothing is captured for that H2.

What to Look For in Results

Presence of captured lists indicates good structural practices.

Absence might indicate opportunities—are there H2 topics that would benefit from a list format?

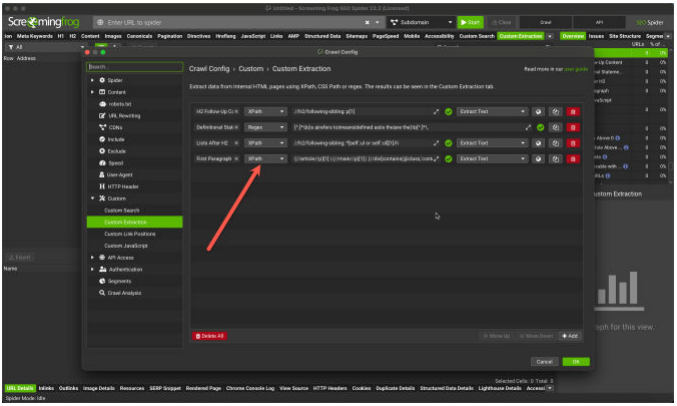
Review the list content: Are items substantive and specific? Or vague and filler?

Extraction 4: First Paragraph (Entity Density)

This extraction captures the opening paragraph of each page's main content—which we'll analyze for entity density.

Configuration:

- Name: First Paragraph
- Type: XPath
- XPath: ((//article|//main|//div[contains(@class,'content')])//p[normalize-space()!=''])[1]
- Extract: Extract Text



What This Captures

This XPath tries three common content container patterns and extracts the first

paragraph from whichever is present. It's designed to work across different site architectures.

Note: Some sites may require adjustment to this XPath based on their specific markup. If you're getting empty results, inspect your site's HTML structure and modify accordingly.

What to Look For in Results

We'll score entity density in the AI analysis phase. For now, review a sample of first paragraphs:

- Do they contain specific named entities?
- Or are they vague, generic intros?

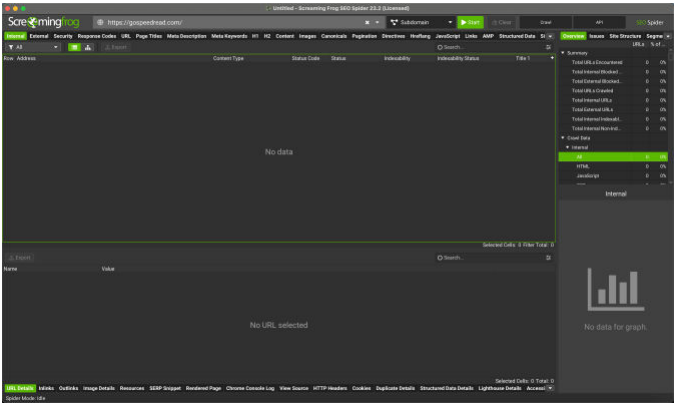
Running the Crawl

With extractions configured:

1. Enter your site URL in the main Screaming Frog window
2. Click Start to begin the crawl
3. Wait for completion (time varies by site size)

For this example, we will scan gospeedread.com. Speed Read AI is an artificial intelligence-powered screenplay and book coverage service that delivers studio-quality analysis in minutes rather than weeks.

The platform generates comprehensive coverage reports—including synopses, loglines, plot dynamics, character arcs, and budget considerations—helping production companies and studios quickly identify promising scripts from their submission piles. (Full disclosure: I work with the founder.)



Crawl Settings to Consider

Spider mode vs List mode:

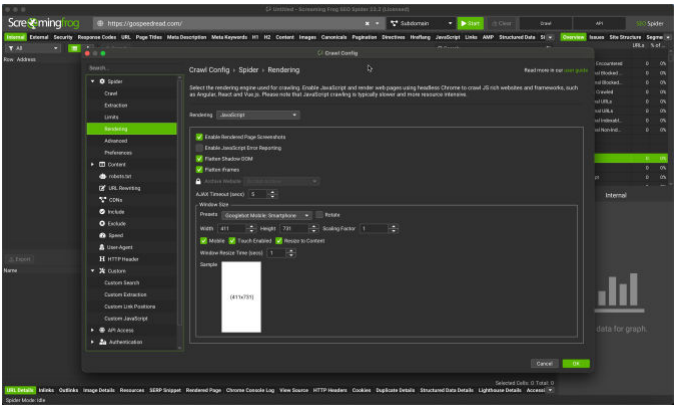
Use Spider mode (the default) to crawl your entire site following links. Use List mode if you want to audit a specific set of URLs.

Crawl limits:

Under Configuration → Spider → Limits, you can set maximum URLs and crawl depth. For a first audit, you might limit to 500 or 1,000 URLs to manage analysis time.

JavaScript rendering:

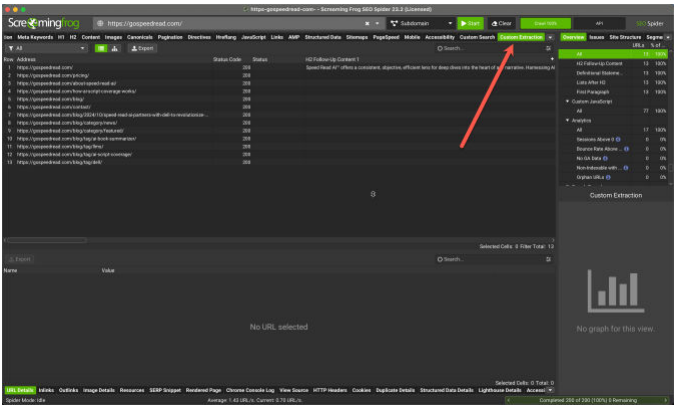
If your site uses JavaScript frameworks, you may need to enable JavaScript rendering under Configuration → Spider → Rendering. This is slower but necessary for some sites.

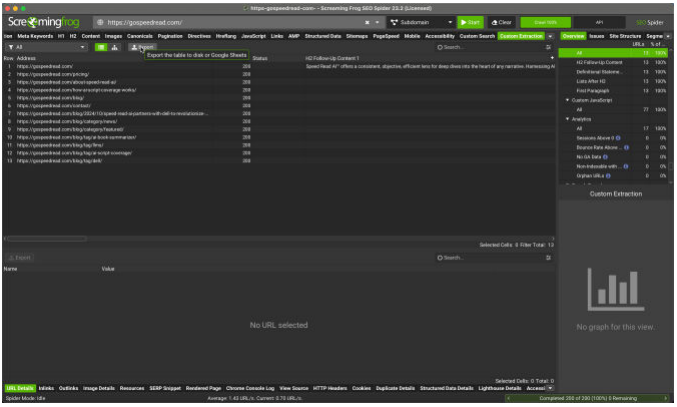


Exporting Results

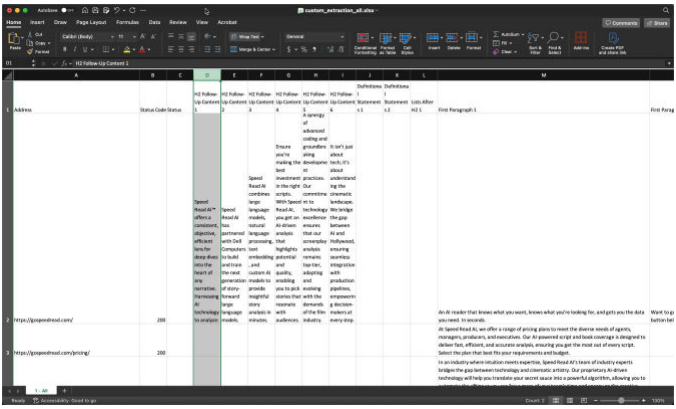
Once the crawl completes:

4. Go to the Custom Extraction tab in the results pane
5. Click Bulk Export → Custom Extraction → All
6. Save as CSV





You'll get a spreadsheet with columns for each extraction plus the source URL. This is the raw data for the AI analysis layer.



Cleaning the Export

Before AI analysis, do some basic cleanup:

- Remove rows where all extraction columns are empty
- Remove non-content pages (privacy policy, terms, etc.) unless you want to audit those too
- Ensure URLs are complete (not relative)

Now you're ready for the AI analysis.

Chapter 6: The AI Analysis Layer

Raw extraction data tells you what content exists. AI analysis tells you whether that content is extraction-friendly.

This chapter provides the prompts and workflows for scoring your extracted data using Claude or GPT.

The Full Scoring Prompt

Here's the complete prompt to use with your exported data. Copy this, paste your CSV data below it, and run it through Claude or ChatGPT:

```
I'm running a content extraction audit on
my website to assess how easily AI
systems can extract and cite my content.
Below is data exported from Screaming
Frog with four custom extractions per
URL:

1. H2 Follow-Up Content — The first
paragraph after each H2

2. Definitional Statements — Sentences
containing "is a," "refers to," "means,"
"defined as"
```

```
3. Lists After H2 — Bullet/numbered list
items immediately following H2s

4. First Paragraph — The opening
paragraph of each page
```

```
For each URL, analyze all four data
points and score the page 1-10 on
extraction-friendliness — how easily
could an AI extract a clean, direct,
factual answer to use in a response or
citation?
```

```
FLAG THESE PROBLEMS:

- Vague or filler intros ("In today's
world," "When it comes to," "It's no
secret that")

- H2s not immediately followed by a
direct answer

- Missing definitional statements on
pages that should define a concept

- Buried answers where key info appears
late instead of upfront

- Weak entity density (no specific names,
products, or concepts named)
```

```
DELIVER:

1. A ranked list of the 10 weakest pages
with specific issues noted
```

```
2. A ranked list of the 10 strongest
pages and what makes them work

3. Three site-wide patterns you're seeing
that hurt extraction-friendliness

4. Three quick-win recommendations I can
apply across multiple pages



Here's the data:

[PASTE CSV DATA HERE]
```

Or, an even simpler prompt like this works:

```
Score each page 1-10 on 'LLM citability'
— how easily could an AI extract a clean,
factual answer from this content? Flag
pages with buried answers, vague intros,
or missing definitional statements.

[PASTE CSV DATA HERE]
```

Understanding the Output

The AI will return a structured analysis with four components:

Weakest Pages

These are your priorities for improvement. Each entry should include the URL and specific issues identified. Common problems include:

- Filler-heavy first paragraphs

- Missing definitions on concept pages
- H2s that ask questions but don't immediately answer them
- Generic language with no specific entities

Strongest Pages

These are your models. Study what makes them work. Often you'll find:

- Clear definitions early in the content
- H2s immediately followed by direct answers
- Strong entity density from the first sentence
- Well-placed lists that structure key points

Site-Wide Patterns

These are systemic issues to address at scale. Common patterns include:

- All blog posts start with the same filler intro template
- Service pages never include explicit definitions
- Content writers consistently bury the answer in paragraph 3 or 4

Quick Wins

These are changes you can implement across multiple pages quickly. Examples:

- Add a definitional sentence to the first paragraph of every service page
- Convert the opening paragraph of each blog post from contextual setup to direct answer
- Add a bulleted "key benefits" list immediately after the main H2 on product pages

Batching Large Sites

If your site has thousands of pages, you'll need to batch the analysis:

7. Split your CSV into groups of 50-100 pages
8. Run the prompt on each batch
9. Compile the weakest pages from each batch for final prioritization
10. Look for patterns that appear across multiple batches

Alternatively, start with your highest-traffic pages or most important content categories.

Individual Signal Prompts

Sometimes you want to focus on just one signal. Here are targeted prompts for each:

First Paragraph Analysis

```
Here are opening paragraphs from my
website pages. For each, score 1-10 on
extraction-friendliness — how easily
could an AI pull a clean, factual answer
from this? Flag any that are vague, use
filler intros like "In today's world," or
bury the point. Include the URL in your
response.
```

Definitional Statement Analysis

```
These are definitional sentences
extracted from my site. Which pages have
strong, quotable definitions an AI would
cite? Which are weak or missing entirely?
List URLs that need definitional content
added.
```

H2 Follow-Up Analysis

```
Score whether each H2 is immediately
followed by a direct answer or
explanation. Flag any where the first
sentence is fluff or fails to deliver on
the H2's promise.
```

Iterating on Results

The first analysis gives you a baseline. After implementing fixes:

11. Re-run the crawl and extraction
12. Re-run the AI analysis
13. Compare scores before and after
14. Identify any new issues introduced

Citability optimization isn't a one-time project. It's an ongoing practice.

Chapter 7: Fixing What You Find

The audit identifies problems. This chapter shows you how to fix them.

Each section addresses a specific citability issue with before/after examples and concrete techniques.

Fixing Vague Intros

The most common citability killer is the vague intro—paragraphs that set context, acknowledge complexity, or establish tone before getting to substance.

The Pattern to Avoid

Vague intros typically start with:

- "In today's [adjective] [landscape/world/era]..."
- "When it comes to [topic]..."
- "It's no secret that..."
- "[Topic] has become increasingly important..."
- "Many businesses struggle with..."

These phrases signal to an AI: "Skip this paragraph—the substance comes later."

Before and After

Before (vague intro):

"In today's competitive digital marketplace, businesses are constantly looking for ways to stand out from the crowd and connect with their target audiences more effectively. The landscape has evolved significantly over the past decade, and what worked just a few years ago may not be as effective today. This is particularly true when it comes to email marketing, which remains one of the most powerful tools in a marketer's arsenal despite the rise of social media and other channels."

After (direct opening):

"Email marketing generates $42 for every $1 spent—the highest ROI of any digital marketing channel. It remains the most effective way to nurture leads, retain customers, and drive direct conversions. This guide covers the fundamentals of email marketing strategy, from list building to automation."

The after version leads with a specific, citable claim (the ROI stat), immediately establishes what email marketing is (lead nurturing, retention, conversions), and tells the reader what to expect.

The Rewrite Technique

15. Find your first substantive sentence (often buried in paragraph 2 or 3)
16. Move it to position one
17. Add one specific fact or stat if available
18. Delete the original context-setting paragraphs

Don't worry about being "too abrupt." Readers (and AI systems) prefer immediate value.

Adding Definitional Statements

If your page is about a concept and doesn't include an explicit definition, add one.

The Definition Formula

Effective definitions follow a pattern:

[Term] is [category] that [distinguishing characteristics].

For example:

- "Content marketing is a strategic marketing approach that focuses on creating valuable content to attract a defined audience."
- "A sales funnel is a visual representation that illustrates the customer journey from initial awareness to final purchase."

- "Technical SEO refers to the optimization of website infrastructure to help search engines crawl and index pages effectively."

Placement

Definitions should appear:

- In the first or second paragraph of the page
- Immediately after an H2 that names the concept (e.g., "What Is Technical SEO?")
- In a visually distinct callout box if your design supports it

Don't bury definitions in the middle of the page. AI systems may not find them there.

Length

Aim for 40-60 words. Long enough to be substantive, short enough to be cleanly extracted.

Restructuring for Immediate Answers

When an H2 asks a question (explicitly or implicitly), the content that follows must answer it immediately.

The Inverted Pyramid

Journalism uses the "inverted pyramid"—most important information first, supporting details after. Apply this to every H2-to-content block:

19. Direct answer to the H2's implied question
20. Key supporting point(s)
21. Additional context, examples, or nuance

Before and After

Before (buried answer):

H2: How Long Does SEO Take?

"This is one of the most common questions clients ask when they begin working with an SEO agency. The answer, as with many things in marketing, is that it depends. There are numerous factors that influence how quickly you'll see results, including your current domain authority, the competitiveness of your industry, and the quality of your existing content. Generally speaking, most businesses begin to see meaningful improvements in their organic traffic within 4-6 months."

After (immediate answer):

H2: How Long Does SEO Take?

"Most businesses see meaningful organic traffic improvements within 4-6 months of implementing an SEO strategy. The exact timeline depends on your starting domain

authority, industry competitiveness, and content quality. Highly competitive industries may require 8-12 months. New domains with minimal existing authority often need 6-9 months to gain traction."

The answer (4-6 months) is now in sentence one, where an AI can extract it immediately.

Optimizing List Placement

Lists are extraction-friendly, but only if they're properly positioned and substantive.

When to Convert Paragraphs to Lists

Convert to lists when content is:

- A set of distinct items (benefits, features, requirements)
- Steps in a process
- Parallel options or alternatives
- Examples or types of something

Positioning

For maximum citability, position lists:

- Immediately after the relevant H2 subheading
- After a single introductory sentence that sets up what follows

Avoid: Lists buried in the middle of long paragraphs, or lists that appear before the relevant H2.

Making List Items Substantive

Each list item should be a complete, meaningful statement—not just a word or short phrase.

Weak:

- Traffic
- Conversions
- Brand awareness

Strong:

- Increases organic traffic through improved search visibility
- Drives conversions by attracting qualified prospects
- Builds brand awareness by positioning you as an industry authority

The strong version is citable. The weak version requires context to understand.

Strengthening Entity Density

Replace vague nouns with specific entities wherever accuracy allows.

The Entity Replacement Technique

Go through your opening paragraph and identify every vague noun. Ask: "Can I name something specific here?"

Vague → Specific:

- "the industry" → "the SaaS industry" or "B2B software"
- "businesses" → "enterprise companies" or "SMBs" or "e-commerce brands"
- "tools" → "platforms like Salesforce and HubSpot"
- "experts" → "Google's John Mueller" or "Rand Fishkin"
- "recent changes" → "Google's March 2024 core update"

Dates and Timeframes

Including specific dates signals currency and relevance:

- "In 2024" instead of "recently"
- "As of Q3 2024" instead of "currently"
- "The 2024 industry report" instead of "a recent study"

Balance

Don't over-stuff entities to the point of unreadability. The goal is specificity where it

adds value, not keyword stuffing. One or two strong entities per sentence is usually optimal.

Chapter 8: Competitive Analysis

Running the citability audit on competitor sites reveals what's working in your space and where you can differentiate.

Selecting Competitors to Audit

Choose three types of competitors:

Direct Competitors

Sites competing for the same customers and keywords. You probably know who these are.

Sites Getting Cited

For your target queries, check what's currently appearing in AI Overviews and note the cited sources. These may not be direct competitors, but they're winning the citability game in your space.

Industry Leaders

Sites with strong authority in your industry, even if they don't compete with you directly. These often set the content quality bar.

Running the Competitor Audit

For each competitor:

22. Run the same Screaming Frog crawl with all four extractions
23. Export and analyze using the same AI prompts
24. Note their scores and patterns

Comparative Analysis Prompt

Use this prompt to compare your results to competitors:

```
I've run citability audits on my site and two competitors. Here's the data for each:  [MY SITE DATA] [COMPETITOR A DATA] [COMPETITOR B DATA]  Compare citability across all three sites. Identify: (1) Where competitors outperform me and why, (2) Where I outperform competitors, (3) Patterns competitors use that I should adopt, (4) Gaps in competitor content I could exploit.
```

What to Look For

Structural Patterns

Do competitors consistently use certain structures? Common patterns in high-citability sites include:

- Definition boxes or callouts at the top of concept pages
- "Key Takeaways" or "Quick Answer" summaries before detailed content

- FAQ sections with clear question/answer formatting
- Lists positioned immediately after topic H2s

Content Depth

Are competitors providing more comprehensive definitions? More specific examples? More current data?

Entity Usage

Do competitors name more specific entities in their opening paragraphs? Are they referencing industry sources, specific tools, or recognized experts?

Turning Insights Into Actions

The competitive analysis should produce a specific action list:

- Patterns to adopt (structural elements that work)
- Gaps to fill (topics competitors cover poorly or not at all)
- Differentiation opportunities (ways to provide more value than competitors)

Prioritize actions by impact (how much will this improve citability?) and effort (how long will this take?). Start with high-impact, low-effort improvements.

Chapter 9: Scaling and Automation

The manual workflow (Screaming Frog extraction → CSV export → Claude analysis) works well for initial audits and smaller sites. For ongoing optimization or large sites, you'll want more automation.

Screaming Frog's OpenAI Integration

Screaming Frog includes built-in OpenAI integration that lets you run AI prompts against page content during the crawl itself.

Setup

25. Go to Configuration → API Access → OpenAI
26. Enter your OpenAI API key
27. Configure prompts to run on extracted content

When to Use API Integration

Use the built-in API when:

- You're crawling thousands of pages
- You want scores as columns directly in your crawl data

- You need to run the same analysis repeatedly across sites

When to Stick With Manual

Manual analysis (export to CSV, paste into Claude) works better when:

- You're auditing fewer than 500 pages
- You want more nuanced, conversational analysis
- You're iterating on what questions to ask
- You want to avoid API costs during exploratory work

Building Repeatable Workflows

Audit Frequency

For most sites:

- Full audit: Quarterly
- Priority pages audit: Monthly
- New content check: Before publication

Tracking Scores Over Time

Create a simple tracking spreadsheet:

- URL
- Initial citability score
- Issues identified
- Fixes implemented

- Post-fix score
- Date of last audit

This lets you measure improvement and catch regressions.

Team Implementation

If you're working with content teams or clients:

Writer Training

Teach content writers the four signals. Provide before/after examples. Make "citability" part of the content brief.

Editorial Checklists

Before content is published, verify:

- First paragraph contains specific entities and gets to the point
- Each H2 is immediately followed by a direct answer
- Key concepts have explicit definitions
- Lists are positioned near relevant headers

QA Process

For important content, run the single-page extractions and analysis before publishing. Catch citability issues before they go live.

Chapter 10: What This Doesn't Guarantee

This playbook provides a systematic approach to optimizing content for AI extraction and citation. It doesn't guarantee AI citation.

Let's be clear about limitations.

Reality check: Lily Ray has highlighted that Google's AI Overviews can surface spam and misinformation, which makes accuracy, sourcing, and trust signals non-negotiable alongside structure. Sources: Lily Ray (LinkedIn) — https://www.linkedin.com/pulse/google-ai-overviews-have-major-spam-problem-lily-ray-go74f and Search Engine Journal — https://www.searchenginejournal.com/does-googles-ai-overviews-violate-its-own-spam-policies/547020/.

Citability Is One Factor

AI systems consider many factors when selecting sources to cite:

- Domain authority and trust signals
- Backlink profile and external citations
- Overall content quality and accuracy
- Recency and freshness

- E-E-A-T signals (Experience, Expertise, Authoritativeness, Trustworthiness)

Citability optimization helps with extraction and citation probability, but it doesn't override weak performance on these other factors. A perfectly structured page on a low-authority domain may still lose to a less-structured page on a highly trusted site.

E-E-A-T: Citability ≠ Credibility

Optimizing for extraction makes your content easier for AI to cite. It doesn't make your content more trustworthy.

Google's E-E-A-T framework (Experience, Expertise, Authoritativeness, Trustworthiness) still applies. AI systems prefer extractable content from credible sources.

Credibility signals include:

- Author bios with verifiable credentials
- First-hand experience with the topic
- Original data, research, or screenshots
- Source attribution for claims
- Consistent track record on the topic
- External citations from other authoritative sources

The best approach combines structural optimization (this playbook) with genuine expertise. Structure without substance won't

get you cited; substance without structure might get overlooked.

YMYL Topics: Higher Standards Apply

For Your Money or Your Life (YMYL) topics—medical, financial, legal, safety-related content—AI systems apply stricter standards.

> YMYL Note: Citability thresholds are higher for YMYL topics. Expect heavier weighting on authority, consensus, and institutional sources. A well-structured article from an unknown source will rarely beat WebMD, Mayo Clinic, or Investopedia for health or financial queries.

If you're creating YMYL content, structural optimization helps but won't overcome authority deficits. Focus on building genuine expertise and trust signals first.

AI Systems Evolve

What works today may shift tomorrow. AI models improve. Citation algorithms update. New signals emerge.

The principles in this playbook (direct answers, clear definitions, good structure) are based on fundamentals that have been stable for years. But specific patterns may change. Stay current with Google's documentation and adjust accordingly.

This Doesn't Replace Good Content

Citability optimization is a formatting and structure layer. It helps AI systems extract value from your content. But if your content lacks substance, expertise, or accuracy, no amount of structural optimization will help.

The foundation is still: create content that genuinely serves your audience's needs with accurate, expert-level information.

Expert note (Dr. Marie Haynes, SEO and E-E-A-T expert): *"Language models really like quoting statistics."* Original research, proprietary data, and first-hand experience add unique value that AI systems actively seek to cite. Ask yourself: "If this content disappeared from the web, would anyone miss it?" Source: https://97thfloor.com/articles/podcasts/ai-search-what-seos-need-to-know-and-do-right-now/

Measurement Is Limited

Unlike traditional SEO where you can track rankings and traffic, measuring AI citation performance is harder:

- No Google Search Console report for AI Overview citations
- Third-party tools are emerging but still limited

- AI Overviews appear inconsistently across queries and users

You can manually check your target queries, but comprehensive measurement requires tools that are still developing.

Expert note: As Rand Fishkin has argued, traffic is a vanity metric—influence has always been more valuable. In the zero-click era, being cited and mentioned builds brand awareness even without direct clicks. Source: https://sparktoro.com/blog/why-do-we-need-zero-click-marketing/

The Right Framing

Think of citability optimization as table stakes for AI-era SEO. It won't guarantee results, but failing to optimize for citability increasingly means failing to compete.

It's like mobile optimization in 2015. You couldn't guarantee mobile rankings, but if your site wasn't mobile-friendly, you were at a significant disadvantage. The same logic now applies to content structure and citability.

Do the work. Build the practice. Accept that outcomes aren't guaranteed, but odds improve significantly with systematic optimization.

Need a Hand?

This playbook gives you the complete methodology. I'm always happy to chat.

Email: david@davidcosgrove.com

That's it. Tell me what you're working on and I'll let you know if I can help.

Appendix A: Complete Prompt Library

Full Citability Audit Prompt

```
I'm running a content extraction audit on my website to assess how easily AI systems can extract and cite my content. Below is data exported from Screaming Frog with four custom extractions per URL:  1. H2 Follow-Up Content — The first paragraph after each H2 2. Definitional Statements — Sentences containing "is a," "refers to," "means," "defined as" 3. Lists After H2 — Bullet/numbered list items immediately following H2s 4. First Paragraph — The opening paragraph of each page  For each URL, analyze all four data points and score the page 1-10 on extraction-friendliness.  FLAG THESE PROBLEMS: - Vague or filler intros - H2s not immediately followed by a direct answer - Missing definitional statements - Buried answers - Weak entity density DELIVER: 1. Ranked list of 10 weakest pages with issues 2. Ranked list of 10 strongest pages 3. Three site-wide patterns hurting extraction-friendliness 4. Three quick-win recommendations [PASTE CSV DATA HERE]
```

First Paragraph Analysis

```
Here are opening paragraphs from my website pages. For each, score 1-10 on extraction-friendliness — how easily
```

```
could an AI pull a clean, factual answer
from this? Flag any that are vague, use
filler intros, or bury the point. Include
the URL in your response.
```

Definitional Statement Analysis

```
These are definitional sentences
extracted from my site. Which pages have
strong, quotable definitions an AI would
cite? Which are weak or missing entirely?
List URLs that need definitional content
added.
```

H2 Follow-Up Analysis

```
Score whether each H2 is immediately
followed by a direct answer or
explanation. Flag any where the first
sentence is fluff or fails to deliver on
the H2's promise.
```

Competitive Analysis Prompt

```
I've run citability audits on my site and
two competitors. Compare extraction-
friendliness across all three. Identify:
(1) Where competitors outperform me, (2)
Where I outperform them, (3) Patterns to
adopt, (4) Gaps to exploit.  [MY SITE
DATA] [COMPETITOR A DATA] [COMPETITOR B
DATA]
```

Appendix B: Quick Reference Checklists

Pre-Publication Citability Checklist

Before publishing content, verify:

- First paragraph contains at least 2-3 specific entities
- First paragraph gets to the main point within 2-3 sentences
- Each H2 is immediately followed by a direct answer
- Key concepts have explicit definitions using "is a" or similar
- Definitions are 40-60 words
- Lists are positioned immediately after relevant H2s
- List items are complete, substantive statements
- Author bio with credentials is present (for E-E-A-T)
- Claims are attributed to sources where appropriate

Audit Setup Checklist (Screaming Frog)

Configure these four extractions:

- H2 Follow-Up Content: XPath //h2/following-sibling::*[1][self::p]
- Definitional Statements: Regex [^.]*\b(is a|refers to|means|defined as|is the|are the)\b[^.]*\.
- Lists After H2: XPath //h2/following-sibling::*[1][self::ul or self::ol]/li
- First Paragraph: XPath ((//article|//main|//div[contains(@class,'content')])//p[normalize-space()!=''])[1]

Set all to Extract Text (not Inner HTML)

Content Rewrite Checklist

For each page identified as weak:

- Identify the first substantive sentence
- Move it to position one
- Delete or move context-setting filler
- Add a 40-60 word definition if needed
- Replace vague nouns with specific entities
- Convert appropriate content to lists

- Verify each H2 is immediately followed by its answer

Appendix C: Resources and Links

Google Official Documentation

Top ways to ensure your content performs well in Google's AI experiences on Search:

https://developers.google.com/search/blog/2025/05/succeeding-in-ai-search

AI Features and Your Website:

https://developers.google.com/search/docs/appearance/ai-features

How Google's Featured Snippets Work:

https://support.google.com/websearch/answer/9351707

Featured Snippets and Your Website:

https://developers.google.com/search/docs/appearance/featured-snippets

Ahrefs Research on AI Overviews

76% of AI Overview Citations Pull From Top 10 Pages (1.9M citations analyzed):

https://ahrefs.com/blog/search-rankings-ai-citations/

Insights From 55.8M AI Overviews Across 590M Searches:

https://ahrefs.com/blog/insights-from-56-million-ai-overviews/

Do AI Overviews Reduce Clicks?:

https://ahrefs.com/blog/ai-overviews-reduce-clicks/

AI Overview Brand Correlation Study:

https://ahrefs.com/blog/ai-overview-brand-correlation/

AI Overviews Change Every 2 Days:

https://ahrefs.com/blog/ai-overview-change/

What Triggers AI Overviews? 86 Factors Analyzed:

https://ahrefs.com/blog/ai-overview-triggers/

Ahrefs AI Overview Video Presentation:

https://www.youtube.com/watch?v=j356-w13h8Q

Industry / Expert Commentary
Lily Ray — "SEO in the Age of Agents: Takeaways from NYC SEO Week" (May 8, 2025): https://lilyray.nyc/seo-in-the-age-of-agents-takeaways-from-nyc-seo-week/

Search Engine Journal — “Does Google’s AI Overviews Violate Its Own Spam Policies?” (May 19, 2025):

https://www.searchenginejournal.com/does-googles-ai-overviews-violate-its-own-spam-policies/547020/

Lily Ray (LinkedIn) — “Google AI Overviews Have a Major Spam Problem” (May 2025):

https://www.linkedin.com/pulse/google-ai-overviews-have-major-spam-problem-lily-ray-g074f

Lily Ray (Search Engine Land) — “AI search is booming, but SEO is still not dead” (July 18, 2025):

https://searchengineland.com/ai-search-booming-seo-still-not-dead-458935

Dr. Marie Haynes — “AI & Search: What SEOs Need to Know (and Do) Right Now” (June 2025):

https://97thfloor.com/articles/podcasts/ai-search-what-seos-need-to-know-and-do-right-now/

Rand Fishkin (SparkToro) — "Why Do We Need Zero Click Marketing?" (January 2025):

https://sparktoro.com/blog/why-do-we-need-zero-click-marketing/

Mike King (iPullRank) — "How AI Mode Works" (May 2025):

https://ipullrank.com/how-ai-mode-works

and AI Mode Webinar Recap (July 2025):

https://ipullrank.com/ai-mode-webinar-recap

Screaming Frog Documentation

SEO Spider Pricing:

https://www.screamingfrog.co.uk/seo-spider/pricing/

Web Scraping & Custom Extraction Tutorial:

https://www.screamingfrog.co.uk/seo-spider/tutorials/web-scraping/

Tools

Screaming Frog SEO Spider:

https://www.screamingfrog.co.uk/seo-spider/

Claude AI:

https://claude.ai

ChatGPT:

https://chat.openai.com

About the Author

David Cosgrove is a digital marketing consultant and with over 30 years of experience. He runs David Cosgrove Los Angeles Web Design from East Granby, Connecticut, serving Fortune 20 companies, Emmy-winning actors, Pulitzer Prize-winning authors, and scaling startups.

His work specializes in web design, SEO, digital strategy, and AI integration services, helping clients position themselves at the forefront of how search and AI are evolving.

www.davidcosgrove.com

Get Cited: The SEO Playbook for AI Search

www.davidcosgrove.com