

— THE AEO REPORT GUIDE · 2026

How to read your HubSpot AEO Grader report.

A plain-English walkthrough of every section of HubSpot's AEO Grader — what each score measures, why it matters, and how to interpret your own numbers to see how AI answer engines describe your brand.

FOR

**Marketing leaders
& RevOps teams**

READING TIME

~15 minutes

EDITION

June 2026

What you'll learn.

01	What AEO is, and the three engines in your report	P. 04
02	How the overall score is built	P. 07
03	Brand Recognition	P. 09
04	Market Competition	P. 12
05	Analysis Summary	P. 15
06	Contextual Analysis	P. 16
07	Sentiment Analysis	P. 17
08	Summary and what to keep in mind	P. 19
09	Where to go from here	P. 21

HOW TO USE THIS GUIDE

Before you **start.**

HubSpot's AEO Grader hands you a dense report full of scores, percentages, donut charts and AI-generated commentary. On its own, a lot of it is hard to act on — what does a "37/100" actually mean? Is "Niche" bad? Why is one engine more positive than another?

This guide walks through every section of the report, explains what each metric measures, why it matters, and how to interpret your own numbers. Where it helps, we use figures from an illustrative sample report as a worked example. Your numbers will differ — the point is to show you how to read them.

By the end you'll be able to look at your report and know three things: where you stand, what's driving the score, and which levers actually move the needle.

What AEO is, and the **three engines** in your report.

The shift you're **measuring**

For two decades, being "findable" meant ranking on Google. A buyer typed a query, scanned ten blue links, and clicked through to your site. You optimised for that with SEO.

Answer engines have changed the rules. When a prospect asks ChatGPT "who are the best HubSpot agencies in the Netherlands?" or "is [your brand] any good?", they don't get ten links. They get a single, synthesised answer — and often they never visit a website at all. This is the **zero-click** reality: the AI is the first, and sometimes only, touchpoint between your brand and a potential buyer.

Answer Engine Optimisation (AEO) is the practice of shaping how those engines describe you. It's the natural successor to SEO. The difference in one line: SEO decides whether you **appear** in search results; AEO decides how AI **represents** you when buyers skip search entirely.

The AEO Grader measures that representation. It asks three leading answer engines how they characterise your brand, then scores the answers across five dimensions.

— CHAPTER ONE · THE THREE ENGINES

Three engines, three views.

Your report runs the same analysis across three different models, because each one builds its view of your brand differently — and your buyers use all three.



OPENAI · CHATGPT

The most widely used answer engine. The Grader queries GPT-5.4 mini, which answers from its training data up to a fixed knowledge cutoff. If something about your brand happened after that cutoff, this engine won't know about it.



PERPLEXITY

Different by design. Perplexity searches the live web before answering, so it has no fixed knowledge cutoff. It reflects what's discoverable about you *right now*, which makes it the most responsive of the three to recent activity, fresh coverage and updated web content.



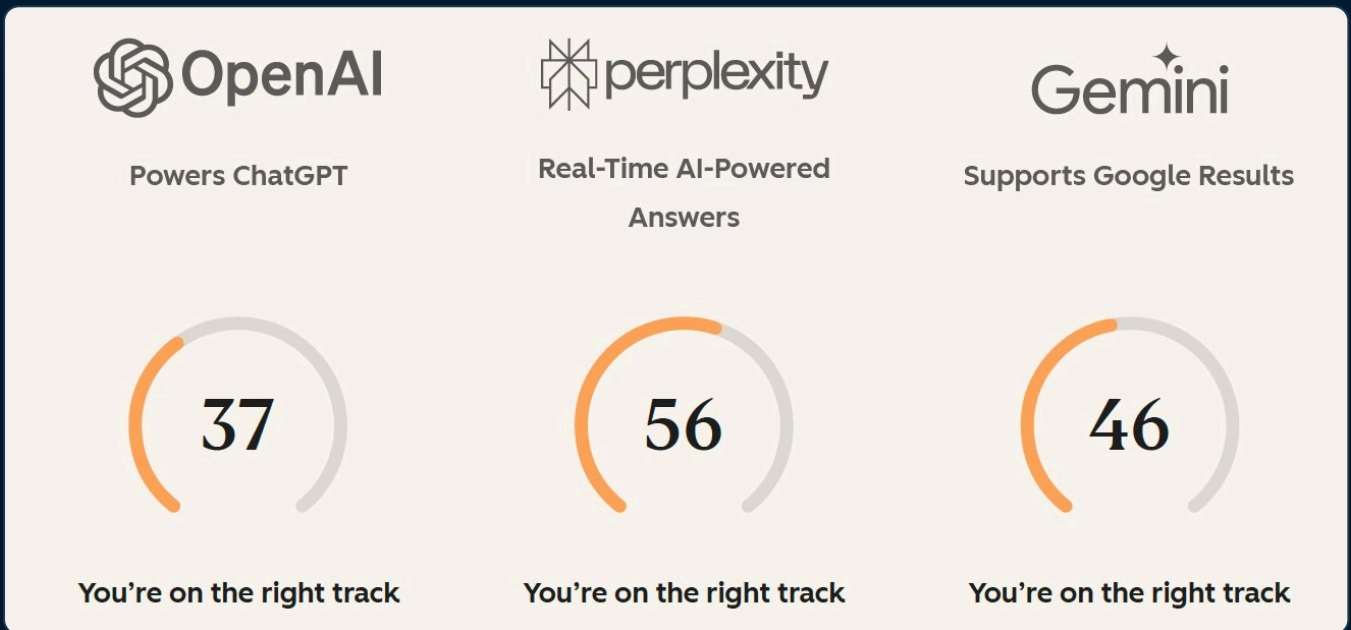
GEMINI · GOOGLE

Google's model, which also works from a training-data cutoff. Given Google's reach, this engine increasingly shapes how your brand surfaces inside the wider Google ecosystem.

CHAPTER ONE · READING THE GAP

Why the scores differ between engines.

It's normal — and informative — to see three different numbers. A higher Perplexity score usually means your **current, live** web presence is strong. A lower ChatGPT or Gemini score often means your brand wasn't deeply established in their training data at the time of their cutoff. Reading the gap between engines tells you whether your AEO problem is "we're invisible historically" or "we're not being picked up in real time" — two very different fixes.



FROM AN ILLUSTRATIVE SAMPLE REPORT

IN THE SAMPLE REPORT

the brand scored 37 on ChatGPT, 56 on Perplexity, and 46 on Gemini. The pattern — Perplexity highest, ChatGPT lowest — is typical of a brand whose live web presence is solid but which hasn't yet built deep, consistent coverage in the engines' historical training data.

— CHAPTER TWO

How the overall score is built.

Your headline AEO Score runs from **0 to 100** and is the sum of five weighted dimensions. The weighting is deliberate: it rewards not just **being known**, but being **described well**.

DIMENSION	MAX	WHAT IT REWARDS
Brand Sentiment	40	How positively the engines talk about you. The single biggest lever.
Brand Recognition	20	Whether the engines can discuss you with real depth, not just confirm you exist.
Presence Quality	20	How substantive, credible and varied the information behind those mentions is.
Market Score	10	How clearly the engines can place you competitively (your category role and brand archetype).
Share of Voice	10	Your slice of the conversation versus named competitors.
Total	100	

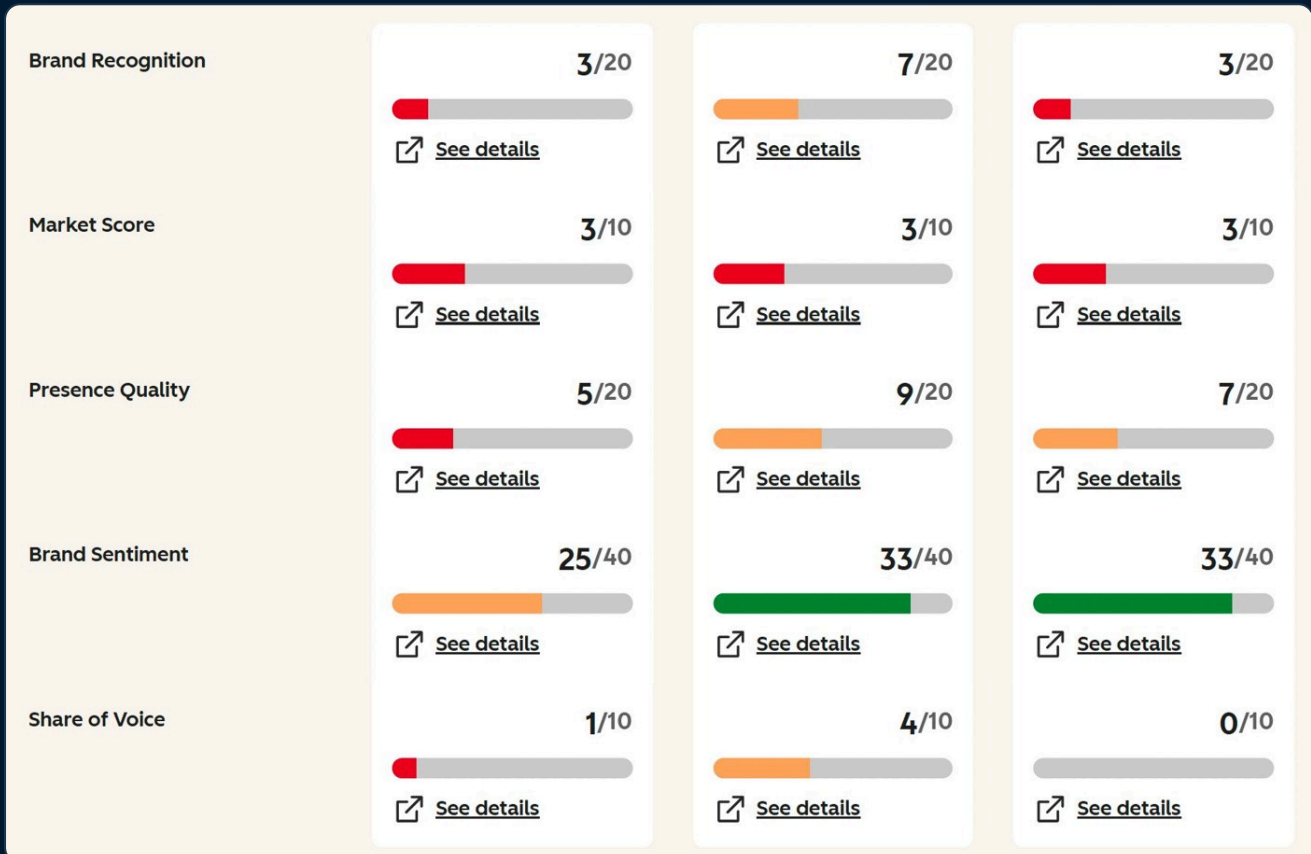
Why sentiment carries the most weight

Sentiment is worth nearly half the total because it answers the question that matters most commercially: **when AI mentions you, is it helping or hurting you?** A brand that's widely recognised but described with hesitation or criticism is worse off than a smaller brand that's described glowingly. Recognition gets you into the answer; sentiment determines whether that answer wins you the deal.

CHAPTER TWO · READING YOUR TOTAL

There is no universal "good score".

The Grader frames bands like "You're on the right track" rather than pass/fail, and for good reason — there is no universal "good score". A score is only meaningful relative to your category, your competitors, and your starting point. Treat the total as a baseline to improve from, not a grade to pass. The value is in the breakdown beneath it: that's where you find out **what** to fix.



FROM AN ILLUSTRATIVE SAMPLE REPORT

IN THE SAMPLE REPORT

every engine returned "You're on the right track." Sentiment was consistently the strongest contributor (25/40, 33/40, 33/40), while Share of Voice and Brand Recognition were the weakest. That's a clear signal: the brand is liked where it's known — the work is to be known more widely and more deeply.

Brand Recognition.

Shows how visible your brand is across digital channels. Higher scores mean more people recognise your brand name.

This dimension is worth up to 20 points toward your total, but the report breaks it down further into a richer set of signals. Here's what each one means.

Recognition Score (0–100)

A composite of how well your brand stands out against competitors, based on online presence and industry relevance. A high score means the engines can talk about you with depth and specificity, not just acknowledge you exist. This is the headline number for the section.

Market Position

How the engines classify your standing in the category:

- ◇ **Leader** — treated as an industry standard or default reference point.
- ◇ **Challenger** — a rising, credible competitor.
- ◇ **Niche Player** — a specialist with a focused remit.

"Niche" is not a failing — for a specialist agency it can be accurate and even desirable. But if you're aiming to be considered alongside the biggest names, moving from Niche toward Challenger is the goal.

Archetype, confidence & depth.

Brand Archetype

The personality the engines read into your brand:

- ◇ **Innovator** — forward-thinking, ahead of the curve.
- ◇ **Disruptor** — challenges the established way of doing things.
- ◇ **Traditionalist** — reliable, established, dependable.

None is inherently better; what matters is whether the archetype the engines assign matches the one you're trying to project. A mismatch tells you your messaging isn't landing.

Confidence Level (%)

How certain the engine is about its own assessment, on a 0–100 scale. This is a meta-signal: it tells you how much data the engine had to work with. A low confidence level means thin or inconsistent information — the engine is guessing more than knowing. Raising confidence is often the first job, because everything else becomes more reliable once the engines have more to go on.

Mention Depth (0–10)

How thoroughly you're discussed, beyond a passing name-check. Deep mentions signal genuine engagement and understanding.

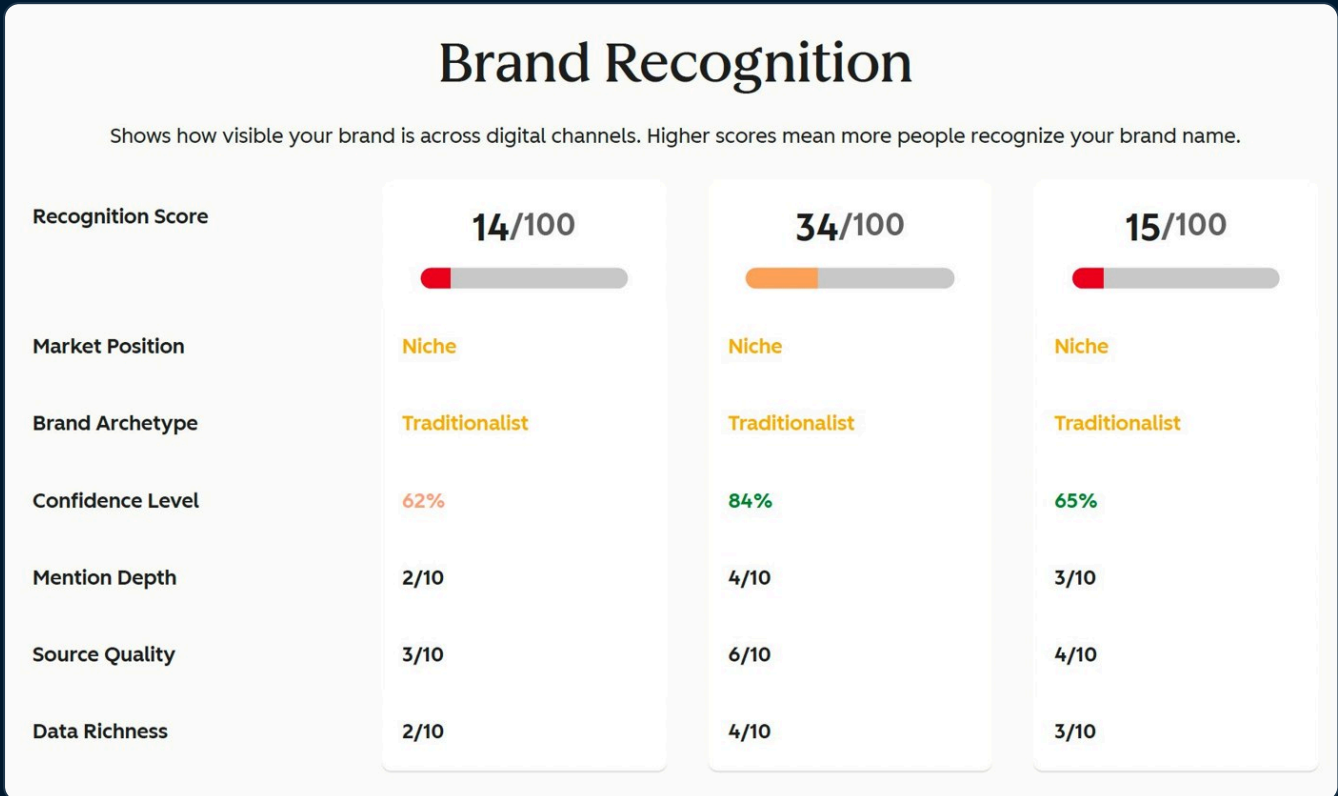
Source Quality (0–10)

The credibility of the sources shaping the engines' view of you. Authoritative, third-party sources lift your standing far more than self-published content.

Data Richness (0–10)

The variety and completeness of information available about you across channels. Breadth matters: a brand described from many angles reads as more substantial than one with a single thin profile.

The Brand Recognition panel.



FROM AN ILLUSTRATIVE SAMPLE REPORT

HOW TO READ IT IN PRACTICE

Low Mention Depth, Source Quality and Data Richness almost always point to the same root cause — not enough credible, third-party material exists about you for the engines to learn from. The fix isn't more self-promotion; it's earning substantive external coverage (case studies with outcomes, reputable directory listings, press, reviews) that the engines can draw on.

IN THE SAMPLE REPORT

Recognition Scores landed between 14 and 34 out of 100, all classified "Niche / Traditionalist", with Confidence between 62% and 84%. Mention Depth, Source Quality and Data Richness all sat low (2–6 out of 10). The consistent story: the brand is correctly understood as a credible specialist, but the engines simply don't have enough deep, third-party material to discuss it richly.

Market Competition.

Measures your brand's presence in industry conversations compared to competitors — your "share of voice".

This section shows whether the engines can place you in your competitive landscape, and how much of the category conversation you own.

Overall Market Score (0–10)

How clearly and consistently the engines can position you competitively. A low score usually means the engines have insufficient or conflicting signals about where you sit — not necessarily that you're losing, but that you're hard to place.

Market Position Breakdown (the donut chart)

Your estimated share of voice as a percentage, shown next to your named competitors. This is the most concrete competitive picture in the whole report: it tells you who the engines name in the same breath as you, and how the attention is divided. Pay close attention to which competitors appear — these are the brands the engines consider your peer set, which may or may not match who you think you compete with.

CHAPTER FOUR · THE CONVERSATION

Mentions & comparisons.

Mentions

The number of times your brand appears on its own in discussions. High standalone mentions indicate brand independence — you're discussed in your own right, not only as a comparison point.

Comparison Mentions

The number of times you're discussed alongside competitors. This reveals your competitive positioning: the topics and contexts where buyers (and engines) weigh you against alternatives.

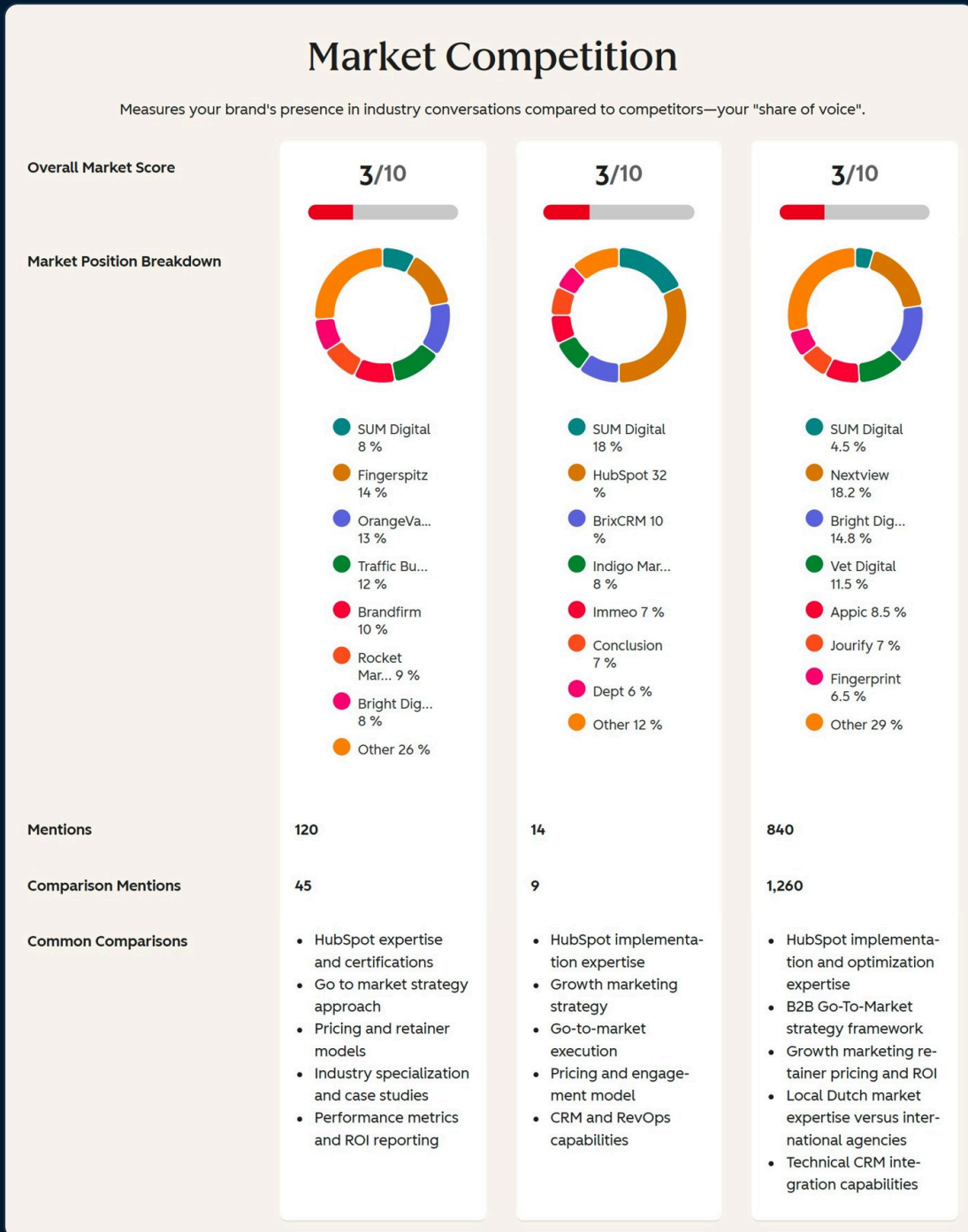
Common Comparisons

The specific themes on which you're most often compared — your competitive battlegrounds. These are the dimensions where deals are won and lost, so they're prime territory for proof points: if "pricing and engagement model" is a recurring comparison, that's exactly where clear, public evidence pays off.

HOW TO READ IT IN PRACTICE

Don't be discouraged by a small share-of-voice slice. The competitor mix and the common comparisons are the real intelligence here — they tell you which battlegrounds to invest in and which competitors the market genuinely benchmarks you against.

Your share of the conversation.



FROM AN ILLUSTRATIVE SAMPLE REPORT

IN THE SAMPLE REPORT

all three engines returned a Market Score of 3/10, and share of voice ranged from 4.5% to 18% depending on the engine and the competitor set it pulled. The recurring comparison themes — HubSpot expertise, go-to-market strategy, pricing models, and proof of results — point precisely to where stronger public evidence would move the needle.

Analysis Summary.

Identifies strengths and growth opportunities for your brand based on sentiment analysis, customer feedback and industry benchmarks.

Key Strengths

What the engines consistently associate with you in a positive light. Read these as your earned positioning: the messages that have successfully landed. They're worth protecting and amplifying, and they make good anchors for new content because you already have momentum behind them.

Growth Areas

Opportunities for improvement, derived from competitive gaps, common criticisms, and emerging market trends. These are the most directly actionable lines in the whole report. Each one is effectively a hypothesis about what would raise your scores — treat them as a shortlist, then prioritise by impact and effort.

Market Trajectory

Whether the engines read your brand as trending up, holding steady, or declining, based on year-over-year sentiment and data patterns. A positive trajectory is reassuring, but it's a tailwind, not a result — it tells you the category is moving your way, which makes now the right time to invest in visibility rather than a reason to wait.

HOW TO READ IT IN PRACTICE

*Cross-reference your Growth Areas against your weakest scored dimensions from sections 3 and 4. Where they overlap, you've found your priorities — the things that are both flagged in the narrative **and** dragging your number down.*

Contextual Analysis.

Assesses how your brand is perceived within the industry by analysing recurring narratives, perceptions and topic associations.

Narrative Themes

The key storylines the engines have formed about you, extracted from digital media and conversations. Where Brand Recognition tells you **how well** you're known and Sentiment tells you **how positively**, this section tells you **what you're known for** — the actual associations the engines make.

This matters because narrative is the part of your AEO you can most directly influence through content and messaging. If the themes the engines repeat are the ones you want to own, your positioning is working. If they're vague, outdated, or off-target, you've found a messaging gap: the engines are characterising you in a way you didn't choose.

HOW TO READ IT IN PRACTICE

*Put this list next to your own positioning statement. Three questions: Are the themes you want to own **present**? Are there themes here you'd rather **not** lead with? Is anything important **missing**? The gaps and surprises are where your content strategy should focus next.*

Sentiment Analysis.

Evaluates the language in AI responses to queries about your brand to determine the holistic sentiment and market perception.

This is the heavyweight section — worth up to 40 of your 100 points — so it gets the most detail in the report. It measures not just **whether** you're mentioned, but **how warmly**.

Sentiment Analysis Total Score (0–100)

Your aggregated sentiment across digital channels and historical patterns. Higher means more positive. This is the number that most directly reflects whether AI is acting as an advocate or a sceptic when buyers ask about you.

The total is built from four lenses, each reported with both a number and the engines' written reasoning:

General

Your overall perception across audiences, drawn from public reviews, digital discourse and general mentions. The broad "what does the market think of you" reading.

Contextual

Perception within your specific sector and use cases. This is sentiment filtered through the realities of your category — for a professional-services brand, that means how you're judged against the standards buyers hold agencies to (measurable ROI, proof of results, specialist credibility).

Sources, polarization & evidence.

Source-Based

Sentiment broken down by type of source: news, social, industry directories, owned media. This reveals where your positive (or negative) perception is coming from. Heavy reliance on owned media (your own site and materials) is a yellow flag — engines weight independent sources more heavily, so sentiment built mostly on self-published content is less durable.

Polarization (lower is better)

How divided opinion is about you. A low polarization score means consistent, predictable perception. A high score means opinion is split, which makes your brand harder for engines to characterise confidently. For most specialist B2B brands, low polarization is normal and healthy.

Source Analysis

An assessment of the reliability, authority and bias of the sources feeding your perception. The report flags whether "reliable data" was available, lists the specific sources it leaned on, and rates each. This is your evidence base laid bare — it shows exactly which sources are shaping the AI's view of you and how credible each one is.

HOW TO READ IT IN PRACTICE

*A strong sentiment score built mostly on **owned media** is more fragile than a slightly lower score built on **independent** sources. The path to durable, defensible sentiment is more credible third-party coverage — reviews, directories, press, and case studies others can verify — not more polished self-description.*

Summary and what to keep in mind.

Pulling the whole report together, here's how to think about your results:

- ◆ **Your total is a baseline, not a verdict.** Its value is in the breakdown — that's where the actions are.
- ◆ **Read the gaps between engines.** A high Perplexity score with lower ChatGPT/Gemini scores points to a strong *live* presence but thin *historical* depth — and tells you whether to focus on real-time discoverability or long-term coverage.
- ◆ **Sentiment is your biggest lever (40 points), but it's strongest when it rests on independent sources.** Recognition and Presence Quality (20 points each) are usually capped by the same thing: not enough credible third-party material for the engines to learn from.
- ◆ **Narrative themes and common comparisons tell you *where* to focus** — they reveal the battlegrounds and the associations worth investing in.

What they know, and why now.

The brands that win in AI search won't be the ones that checked their score and waited for the next training run. They'll be the ones who started building the evidence today — so that when the engines next update, the story they tell has already changed.

The three engines don't all work from the same information, and two of them work from a fixed point in time. Anything about your brand that's newer than an engine's knowledge cutoff simply won't be reflected in that engine's view.

- ◆ **OpenAI (GPT-5.4 mini)** has a knowledge cutoff of **31 August 2025** (this may vary slightly depending on your time zone). Reference: developers.openai.com/api/docs/models/gpt-5.4-mini
- ◆ **Perplexity** has **no knowledge cutoff** — it searches the web in real time, so it reflects your current, live presence. Reference: perplexity.ai/hub/faq/what-is-perplexity
- ◆ **Gemini (3 Flash Preview)** has a knowledge cutoff of **January 2025**. Reference: ai.google.dev/gemini-api/docs/models/gemini-3-flash-preview

Why an "old" cutoff makes this more urgent, not less

A fair question: if two of the three engines work from data that's 10 to 18 months old, why does today's snapshot matter? Turn the question around. That lag is exactly **why** you should act now.

The cutoff freezes how AI describes you — and that description is live to buyers today. Whatever ChatGPT and Gemini "learned" about you before their cutoff is what they're telling your prospects right now, every time someone asks. The data is dated; the conversation is not. A thin or unflattering picture isn't a historical footnote — it's the answer a buyer hears this afternoon.

These engines change slowly, and they have momentum. Training data isn't refreshed on demand; it's rebuilt on the model-maker's schedule, and depth compounds over many months. A weak ChatGPT or Gemini score today is therefore a **leading indicator** — the most reliable preview of how those engines will keep representing you for the next year unless new, credible material starts accumulating now.

Perplexity proves the point. Because it reads the live web, the gap between your Perplexity score and your ChatGPT/Gemini scores is a direct measure of the ground you stand to gain as the slower engines catch up. The wider that gap, the more upside is waiting to be captured by the brand that moves first.

These results are generated by AI tools and are not a substitute for professional advice. A snapshot tells you where you stand today; the real value comes from acting on it and tracking how perception shifts over time.

— WHERE TO GO FROM HERE

Don't just read it. Act on it.

Your AEO Grader report is a strong starting point — but a snapshot only tells you so much. The hard part is turning these scores into a prioritised plan: which dimensions to attack first, which sources to build, which narratives to reinforce, and how to track whether any of it is working as the engines update. That's exactly the conversation we'd like to have with you.

Book a discovery call. We'll walk through your results together, separate the signal from the noise, and map out a concrete approach to improving where it counts. You'll leave with a clear sense of your top priorities and what moving the needle actually involves — whether that's earning credible third-party coverage, sharpening your narrative, or setting up continuous monitoring across all three engines.

100+

B2B companies in SaaS, manufacturing & professional services trust us with their HubSpot.

10y

Helping B2B SMEs grow with proven frameworks and a getting-it-done mindset.

1

Accredited HubSpot Partner team that plugs in as an extension of yours.

BOOK YOUR DISCOVERY CALL →