

Token Economics: Can Logprobs Help You Increase AI Search Visibility?

Metehan Yesilyurt

AI Search & Information Retrieval Researcher · metehan.ai

Published: August 01, 2025 ·

Canonical: <https://metehan.ai/blog/token-economics-log-probabilities-ai-search-visibility/>

AI search engines like ChatGPT, Perplexity, and Claude are starting to replace traditional Google-style search for many users. For SEOs and content writers, this raises a new question:

How do we optimize content for AI models?

One concept you might hear about in this context is **log probabilities** (or *logprobs*). In this post, I'll break down what logprobs are, how they work, and whether they can really help your content rank higher in AI-generated answers.

Note: It's an experimental idea that has caveats. Optimizing all your content/website with the logprobs approach can hurt your rankings.

Free script here: <https://github.com/metehan777/logprobs-for-ai-search>^[1]

What Are Logprobs?

Every large language model (LLM) predicts the next word or token in a sequence. It doesn't just pick one; it generates a **probability distribution** over all possible next tokens.

- **Log probability (logprob):** A negative number representing the likelihood of a token. Higher (less negative) logprob = more likely.
- Example: "dog" : logprob = -0.3 (≈ 74% chance)
- "cat" : logprob = -1.2 (≈ 30% chance)

Why use log values? Because probabilities get very small when predicting token sequences, and logs make math easier and more stable.

Why Can SEOs Care?

LLMs like ChatGPT don't "rank" pages like Google's PageRank. Instead, they **predict tokens** based on their training data and current context. If certain words, phrases, or entities are strongly associated with your brand or page, the model is **more likely** to predict them when answering queries.

In other words:

- If your brand/page name has **high logprobs** in relevant contexts, it may appear more frequently in AI answers.
- This is like building “token authority” — similar to keyword relevance, but in token space.

How Can You Measure Logprobs?

You can’t see logprobs inside ChatGPT itself, but OpenAI’s API and other model APIs allow developers to request them. Here’s a simplified explanation of the tool I built (see code above):

1. **Type a phrase** (e.g., “Best EV cars are”)
2. **Model predicts next tokens** with top 10 logprobs
3. **Convert logprobs to probabilities** to see how confident the model is

Example output:

TEXT
" Tesla" 50.25%

This shows which entities the model naturally associates with that query.

Using Logprobs for AI Search Optimization

Here’s how SEOs and content writers might use this insight:

1. **Entity Alignment** - Check if your brand/product appears in the top logprob predictions for relevant queries. - If not, create content or mentions that link your brand with the entity cluster.
2. **Query Expansion** - Logprobs can reveal synonyms or related entities that the model “expects.” - Use these in your content to increase semantic coverage.
3. **Content Auditing** - Compare logprob outputs before/after content changes to see if association strength improves.
4. **Competitive Analysis** - See which competitors the model predicts — useful for benchmarking “AI search share.”

Does This Really Improve Visibility?

Short answer: Indirectly, yes, but with caveats.

- Logprobs don’t directly boost rankings; they reflect the model’s current associations.
- By studying them, you can identify **gaps** (where your brand isn’t predicted) and create content to fill them.
- Over time, if your brand is widely mentioned in the right contexts, the model will likely predict it more often.

In other words: **Logprobs are diagnostic, not a ranking signal.** They show *what the model knows*.

Practical Example

•
Figure

•
Figure

•
Figure

"I tested your tool, and it shows different probabilities every time. Why?"

1. Stochastic Nature of LLMs

- Even at **temperature = 0.1**, models like GPT-4o don't produce 100% deterministic outputs.
- Tiny floating-point variations occur during **sampling** and **beam search** steps, especially when predicting top tokens.

2. Dynamic Context Effects

- Each time you send the prompt, the model recalculates probabilities based on: Hidden **attention weights**
- Subtle formatting (even invisible characters or whitespace)
- Session-level randomness in tokenization

If your input text changes slightly (e.g., one character), it shifts the entire probability distribution.

3. Floating-Point Precision

- Logprobs are computed in **log space** (natural log or base-2) and converted to probabilities.
- Converting back to percentages introduces **rounding noise** (e.g., -1.234 vs -1.238).

4. Server-Side Variability

- OpenAI's API runs across distributed GPUs. Different hardware batches can produce tiny numeric differences in softmax outputs, harmless but visible in logprob values.

5. Top-Logprobs Sampling

- When you request ``top_logprobs=10``, the model provides the **most likely tokens at that instant**.
- Small ranking shifts occur when two tokens are close in probability (e.g., 31% vs 30.8%).

Does This Affect AI Search Visibility Analysis?

Not really. The **relative ranking of tokens** (which token is #1 vs #2) stays stable even if percentages fluctuate by 1–2%. For your use case (checking if your brand/entity appears at all), those micro-variations don't matter.

Key Takeaways for SEOs & Writers

- **Logprobs reveal token likelihoods:** They're a window into how models "think."

- **Use for entity strategy:** Ensure your brand is strongly linked to target concepts.
- **Don't obsess over exact numbers:** Use trends, not absolutes.
- **Combine with traditional SEO:** Backlinks, structured data, and on-page optimization still matter.

Final Thought

AI search is still evolving. Unlike Google, which has decades of public SEO research, LLM optimization is new ground. Logprobs won't replace keyword research, but they offer a glimpse into the **hidden layer of token economics** shaping AI-generated answers.

Useful links:

https://cookbook.openai.com/examples/using_logprobs^[2]

<https://docs.together.ai/docs/logprobs>^[3]

<https://developers.googleblog.com/en/unlock-gemini-reasoning-with-logprobs-on-vertex-ai/>^[4]

https://python.langchain.com/docs/how_to/logprobs/^[5]

Logprobs and token economics are the deepest layer of AEO measurement. They reveal which phrases AI engines find natural to cite, which directly shapes AI citations in ChatGPT, Perplexity and Gemini. Answer engine optimization eventually converges on this work.

REFERENCES & SOURCES

[1] <https://github.com/metehan777/logprobs-for-ai-search> — <https://github.com/metehan777/logprobs-for-ai-search>

[2] https://cookbook.openai.com/examples/using_logprobs — https://cookbook.openai.com/examples/using_logprobs

[3] <https://docs.together.ai/docs/logprobs> — <https://docs.together.ai/docs/logprobs>

[4] <https://developers.googleblog.com/en/unlock-gemini-reasoning-with-logprobs-on-vertex-ai/> — <https://developers.googleblog.com/en/unlock-gemini-reasoning-with-logprobs-on-vertex-ai/>

[5] https://python.langchain.com/docs/how_to/logprobs/ — https://python.langchain.com/docs/how_to/logprobs/