

# AI crawlers do not render JavaScript, so sign your texts!

Did you know that apart from Googlebot, none of the popular AI crawlers render JavaScript?

None of the following crawlers can render JavaScript:

- 🚫 OpenAI (OAI-SearchBot, ChatGPT-User, GPTBot)
- 🚫 Anthropic (ClaudeBot)
- 🚫 Perplexity (PerplexityBot)

That's right: most AI crawlers do not render JavaScript. Popular AI crawlers like those used by OpenAI and Anthropic do not even execute JavaScript. There's no renderer.

That means they won't see content that is rendered client-side through JavaScript.

This has implications for how websites must be optimized to ensure their content is accessible to these AI bots.

My simple conclusion? Those of us who want to see our first and last names included in future datasets used to train LLMs should sign our texts from now on.

This short text isn't about server-side rendering or other rendering strategies, it is about helping "AI search" with good old-fashioned signatures in your texts and articles.

All language models break down text into smaller units called tokens. That's called tokenization. These tokens can be words, subwords —and names. It all depends on the model's architecture and training data, but that's the concept simplified.

Each small unit of broken-down text, each token, is converted into a vector (embedding) in a high-dimensional space. This embedding captures semantic and syntactic information about the token. During training, the language model learns to associate these embeddings with each other based on semantic similarity and co-occurrence.

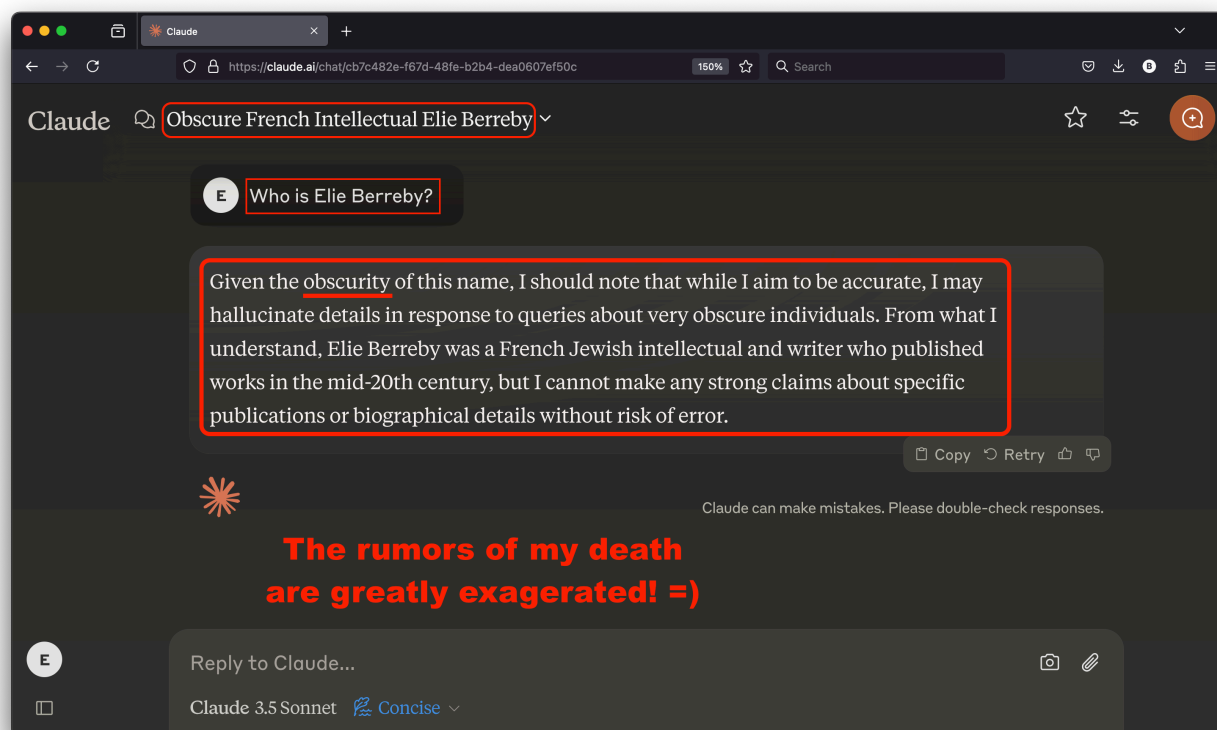
Words that often appear together in similar contexts will have embeddings that are close in the vector space. That's co-occurrence. And words with similar meanings will tend to cluster together in this space. That's semantic similarity.

Most first and last names do not carry semantic meanings. That's not true in every language, but we will ignore semantic similarity and focus on co-occurrence.

If your author's FirstName + LastName is common enough in the training data of a language model, your entire name might be tokenized as a single token. That happens for popular

writers or political leaders: their “FirstName LastName” appears frequently enough together in the training corpus to deserve a single token.

Unknown, unpopular or “obscure” figures require separate tokens. Their first and last names are tokenized separately. For example, I’ve been a nobody online for years because I love privacy. As a result, my name “Elie Berreby” would be tokenized into two tokens: [‘Elie’, ‘Berreby’].



When names are used in articles/texts/citations, the exact combination of first and last names (co-occurrence) helps correctly attribute works to authors.

What’s worse than not existing as an author? Seeing popular LLMs confuse you with similarly named people, especially if they aren’t known for doing fantastic things. I searched for several unpopular names and I got the craziest results. When the LLMs are “confused” they tend to hallucinate.


If you have a common name, you should opt for a disambiguation strategy: use a middle name or anything that will differentiate you. Co-occurrence data helps because it ensures that the correct combination of names is matched. This reduces errors like confusing different authors with similar names.

When you search for information about or by an author, LLMs typically rely on exact or near-exact matches of names.

I believe we should sign our future articles with a signature in your body text, such as

“© *FirstName LastName.*”

Doing this systematically can transform how conversational AI models understand and answer questions about YOU.




**Kim Kuhlman, PhD** · 1st  
B2B Content Marketing Strategist | SEO | AI Expert | Data-driven busine...

2d ...

Then forgive the naive question, why do they bother with long robots.txt files?

**Insightful** · 💡 1 | Reply · 3 Replies

[See previous replies](#)



**Elie Berreby** **Author**  
Senior SEO n00b

2d ...

**Kim** none of the popular "AI crawlers" render JavaScript:

- 🚫 OpenAI (OAI-SearchBot, ChatGPT-User, GPTBot)
- 🚫 Anthropic (ClaudeBot)
- 🚫 Meta (Meta-ExternalAgent)
- 🚫 ByteDance (Bytespider)
- 🚫 Perplexity (PerplexityBot)

The only one that fully renders JavaScript is Googlebot, which has implications regarding AI overviews.

But it seems we have different topics in mind?

Bad bots are aggressive web crawlers that ignore the directives in the robots.txt file. You are right, robots.txt directives are suggestions, not rules. We assume most respectable bots would abide by those directives.

That said, I often bypass and ignore those directives myself when conducting technical SEO audits because this can help identify issues. Bad bots do it for different reasons (usually to access confidential content).

**Insightful** · 🧠👍 2 | Reply

As of today, most authors/writers/creators have their names outside of their texts. The goal here is simple: include your name directly in your text. If you repeatedly include your first and last names in the body of your texts, your name could become a single token.

Only Googlebot fully renders JavaScript through Google's powerful renderer. And yet, Google Search still has a hard time identifying authorship. This became clear to me after auditing multiple websites in 2024.

Signing your content will make your name part of the core text. Over time, you might end up in the datasets of LLMs such as the ones used by Google's AI overviews, ChatGPT, Gemini, and Claude.

Those crawlers aren't designed to ignore core/body text. If being mentioned is essential to you, sign your creations!

This is an extract from yesterday's 9,300+ word [case study about one of my previous Hacker News post and how Google and Bing shadow-banned my content while giving the top position to copycats!](#)

© Elie Berreby - January 3, 2025 - Content first published on [semking.com](#) at 3:30 PM UTC / 10:30 AM EST (Eastern Standard Time).