

Will JSON-LD injected via Google Tag Manager with no SSR be invisible to AI crawlers?



Mersudin Forbes

SEO Director | Founder @ Multiply

2h

How does this behave with traditional JSON structured data implemented using GTM not rendered server side?

Love · ❤️ 1 | Reply · 1 Reply



Elie Berreby

Senior SEO n00b

1s

Amazing question and great DMs! I'll try to answer this tomorrow in a post to help as many people as possible! 😊

[Elie Berreby Technical SEO January 18, 2025](#)

Will JSON-LD injected via Google Tag Manager with no SSR be invisible to AI crawlers?

Technical SEO: since many AI crawlers do not render JavaScript, what will happen to your Structured Data written in JSON-LD and implemented using Google Tag Manager without a server-side rendering strategy?

Googlebot can render JavaScript because it is a headless Chromium browser: it has a renderer by design! What about Bingbot? Applebot? Yes? No? Maybe? Let the games begin! LOL! 🍿

I'll mention "LLM crawlers" or "AI crawlers" but think of crawlers such as those used by ChatGPT (OAI-SearchBot, ChatGPT-User, GPTBot), Claude (ClaudeBot) or Perplexity (PerplexityBot).

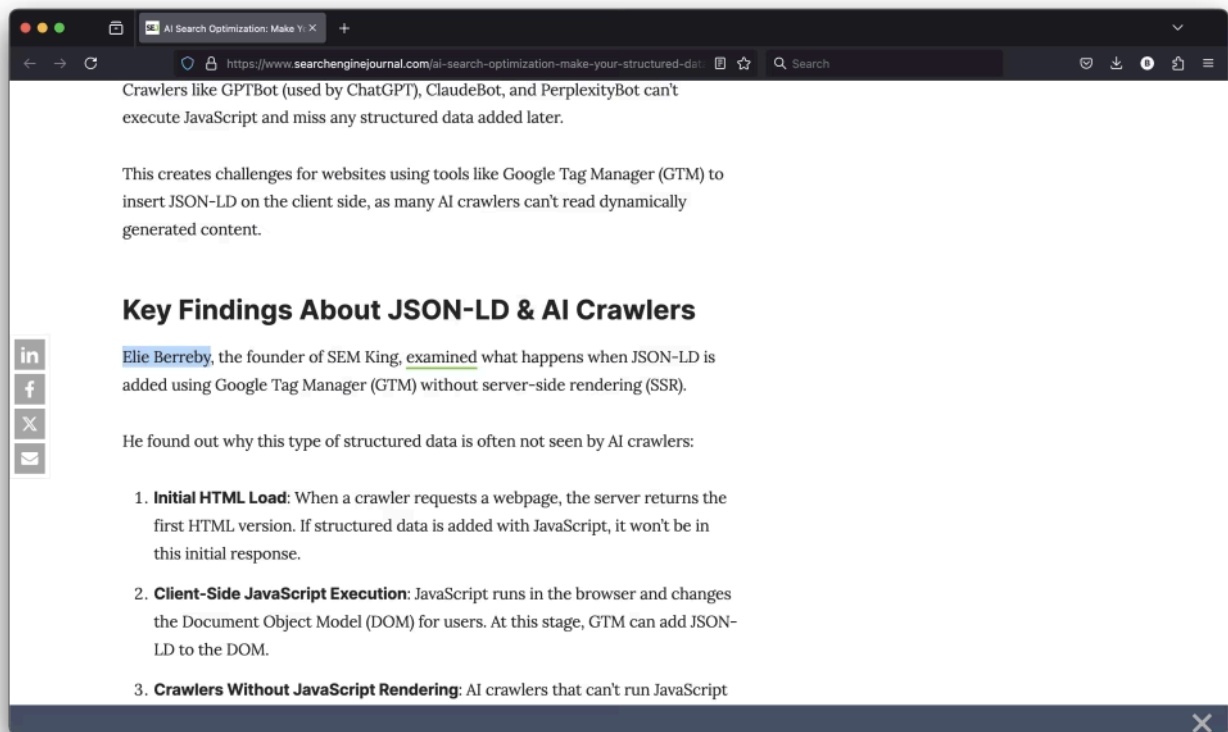
To simplify, I'll list the sequence of steps humans and AI crawlers go through when they try to interact with pages for which JSON-LD data is injected by Google Tag Manager (GTM) without SSR.

QUICK CLARIFICATION: Since my text below has been quoted by the media multiple times and because I understand journalists do not have time to read the comments under my posts, allow me to clarify a few things directly here.

I wrote this short article to answer a strategic question regarding the future of Search. My text was written for the coming months/years, not for the present. As of January 2025, none of the major “AI crawlers” can directly make sense of Schema Markup (structured data). But do I expect this to last? No. I expect significant changes in 2025! And I help those who follow me to be future-proof!

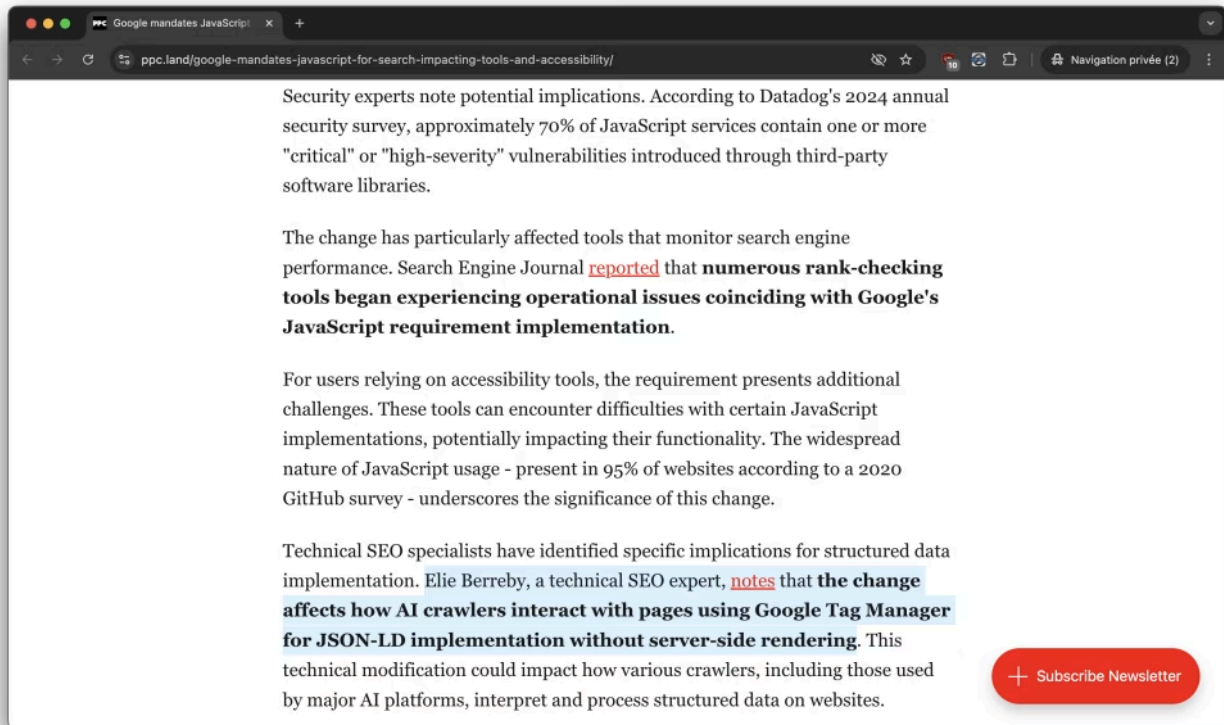
— Elie Berreby

Search Engine Journal



I also saw a few out-of-context mentions of my answer. That’s okay. I understand these are relatively complicated topics and I trust anyone who reads my text directly should have a clear understanding. If not, that’s also something I have explained in my comments, but if you are curious, feel free to reach out on social platforms: I always welcome constructive criticism when it is done respectfully =)

PPC Land



Okay, back to my text 😊

1. Initial HTML load

The server sends the initial HTML to the client (browser or crawler). Without server-side rendering, this initial HTML will NOT include your JSON-LD!

🪡 The JSON-LD will be injected by GTM later via client-side rendering.

2. The AI crawler reaches the page

✅ The AI crawler requests and receives the initial HTML of the page. At this point, it only sees what's in the initial HTML returned by the server.

3. JavaScript is executed in your browser (Chrome, Safari, Firefox, etc)

✅ JavaScript and GTM scripts start to execute. GTM begins injecting and manipulating DOM elements. But unlike traditional web browsers, LLM/AI crawlers that cannot execute JavaScript can't see how the Document Object Model changes!

4. GTM manipulates the DOM to inject JSON-LD using your browser

After the page loads, GTM injects JSON-LD into the of the document (ideally). This action happens after the initial HTML load and requires JavaScript execution. An execution that our AI crawlers can NOT perform!

5. Most AI crawlers are left behind

🚫 ChatGPT, Claude and Perplexity do NOT notice the DOM manipulations: they have no idea GTM added your JSON-LD data!

🚫 The crawlers only read the static HTML as it was sent by the server, without any client-side changes. They can't see your structured data!

6. But can AI crawlers read your JSON-LD?

The JSON-LD data, now part of the DOM after GTM's actions, is NOT accessible to the AI crawlers without JS because:

🔴 The JSON-LD wasn't in the initial HTML (no server-side rendering).

🔴 The poor LLM crawlers can't execute JavaScript and see the DOM after modifications.

Serve your Structured Data using different methods!

📌 Unless your Schema markup data is part of the initial HTML response or served through a method that doesn't require JavaScript execution, AI crawlers (except those like Googlebot that can execute JavaScript) will NOT see any JSON data managed by Google Tag Manager on the client side!

Now, what can

🧠 My pro tip? You can serve JSON data via STATIC files that do NOT depend on JavaScript execution, which makes them accessible to all crawlers! Mine is super original: `schema.json` 🤪

Rendering methods to ensure AI crawlers see your Schema markup

For enterprise SEO projects, you could consider 3 main approaches.

- **Static HTML:** simply include your Schema markup directly in the HTML document that is served from the server! But is Static Site Generation (SSG) always an option on enterprise websites?
- **Server-Side Rendering (SSR):** render pages on the server where the Schema markup is part of the initial HTML load. Think of the best JavaScript frameworks: Next.js on top of React or Nuxt.js (Vue.js)

- **Prerendering:** use prerendering techniques to serve a version of the page where JavaScript has already been executed, including the Schema data. Do this well to ensure there's no cloaking!

Still hard to understand? I'll explain differently!

Because Google Tag Manager injects/modifies JavaScript on the client side, LLM crawlers that don't run JavaScript can't see the JSON data that GTM pushes into the DOM. When your JSON data is not rendered SERVER SIDE but instead generated and manipulated by Google Tag Manager on the CLIENT SIDE, crawlers that cannot render JavaScript will miss this structured data!

Your JSON-LD will be invisible to them because it was not part of the initial server response.

Google Tag Manager operates client-side and modifies the page AFTER it has initially loaded.

Our little AI crawlers won't process dynamic JavaScript content that runs AFTER the page load!

If your JSON data is generated or manipulated by JavaScript (through GTM), crawlers unable to render JavaScript won't see the final state of the data as it appears to humans like you and me! Non-JS AI Crawlers only see the initial HTML content BEFORE any client-side JavaScript execution.

Simplistic TLDR

Google Tag Manager injects and modifies JavaScript on the client-side. Since LLM crawlers that don't run JavaScript cannot execute this code, they won't be able to see the JSON-LD that GTM pushes into the Document Object Model (DOM). If your structured data is generated or manipulated by GTM on the client-side, it will be invisible to these AI crawlers. They'll only see the initial HTML content served by the server, not the dynamically generated content that appears after JavaScript execution.

Is server-side rendering perfect?

Server-side rendering means both the browsers and crawlers receive HTML that already includes the results of the JavaScript execution, including the structured data generated by GTM. With SSR, LLM crawlers can see the complete, rendered HTML, including the JSON-LD data, as it would appear in your browser. But some complex GTM configurations will still be challenging. Sure, server-side rendering is MUCH better than client-side rendering. But SSR isn't perfect!

AI crawlers & JavaScript: expect change soon!

♥ That was a GREAT question that came at a special time in the history of web crawling: agents unable to execute JavaScript crawl aggressively but can only interact with static snapshots of pages! They only see the initial version before any client-side JavaScript changes occur! As of January 2025, that's true for most AI crawlers!

Do I expect this to last? No! Expect change soon! ⌚

© Elie Berreby January 18, 2025 – semking.com