# Debate on the impact of Schema.org structured data on AI search engine output



A recent discussion on LinkedIn has sparked considerable debate within the SEO community regarding the influence of Schema.org structured data on large language models (LLMs), such as those utilised by AI search engines. Patrick Stox, a noted figure in the SEO space, raised questions regarding the validity of claims that optimising content with structured data would improve rankings or outputs from systems like ChatGPT Search.

In his post, Stox queried, “Did I miss something? Why do SEOs think schema markup will impact LLM output?” His comments appear to focus on the context of AI search engines, prompting questions about how they interact with structured data. It has been noted that LLMs are primarily trained on a diverse range of text sources, including web content, books, and various documents, which are synthesised to generate responses without direct replication of source material. This indicates that merely enhancing web content may not lead to direct referrals from LLMs.

AI search engines operate through mechanisms such as Retrieval Augmented Generation (RAG), relying on search indexes and knowledge graphs that are built from crawled data rather than structured input from Schema.org. For instance, Perplexity AI employs an adapted version of Google's PageRank algorithm to rank content that is derived from its search index, underscoring the critical role of crawled text data. Furthermore, major players in the search industry, like Google and Bing, extensively process HTML content to derive essential text for ranking, indicating that structured data may not be the sole or primary factor in their assessment processes.

The limited role of Schema.org in major search engines is reiterated by the fact that Google only utilises a small fraction of this data for specific search experiences and rich results; thus, the extensive reliance on structured data as a ranking mechanism appears to lack robust evidence. Stox expresses concern that the proposition of Schema.org as a basis for improved performance in AI search engines may stem from miscommunication, noting that ideas can often become distorted over time within the SEO community. He points to Jono Alderson's original suggestion that structured data could assist AI search engines in understanding web content, highlighting how this may have evolved into a broader belief not necessarily grounded in current capabilities.

These discussions indicate a prevailing issue within the SEO field, where a lack of concrete understanding may lead to the propagation of inaccurate information. Stox recounts instances where misleading claims, such as those regarding the non-use of IP addresses in local search results, demonstrate the potential for confusion within the profession.

Contributing to the discourse, SEO expert Christopher Shin reinforced Stox’s assertions, clarifying that while structured data plays a role in Search Engine Results Pages (SERPs), LLMs predominantly rely on data interpretation rather than direct engagement with the output of search engines. His insights align with the view that there is value in maintaining a pragmatic approach to SEO, encouraging professionals to base their strategies on factual information.

The prevailing sentiment among participants in this discussion is a push for more grounded and realistic practices in SEO, counteracting tendencies to engage in speculation devoid of substantial evidence. The ongoing dialogue suggests a need for clarity and precision in the field, as industry experts advocate for a more informed and pragmatic approach to leveraging technology in SEO practices.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

* <https://www.positional.com/blog/structured-data-seo> - This article explains that structured data, while not a direct ranking factor, helps search engines better understand and display content, leading to rich snippets and improved click-through rates.
* <https://www.searchpilot.com/resources/case-studies/testing-schema-markup> - This resource details how schema markup enhances search visibility, improves click-through rates, and increases organic traffic, but does not directly impact rankings.
* <https://www.searchenginejournal.com/ranking-factors/schema-markup-ranking/> - This article clarifies that schema markup improves the appearance of search results and helps search engines understand content better, but it is not a ranking factor according to Google's John Mueller.
* <https://www.positional.com/blog/structured-data-seo> - This article supports the idea that Google uses structured data to provide a 'roadmap' for interpreting and displaying content, but it does not directly influence rankings.
* <https://www.searchpilot.com/resources/case-studies/testing-schema-markup> - This case study highlights that schema markup helps content appear in knowledge graphs and increases visibility, but it does not directly affect search engine rankings.
* <https://www.searchenginejournal.com/ranking-factors/schema-markup-ranking/> - This article quotes Google's John Mueller, stating that structured data helps search engines understand entities on the page but does not make the page rank better.
* <https://www.positional.com/blog/structured-data-seo> - This article explains how schema markup like Schema.org's 'Product' markup helps Google identify product details, which can be displayed in rich results.
* <https://www.searchenginejournal.com/ranking-factors/schema-markup-ranking/> - This article mentions a case study by Google showing that structured data can increase traffic and time on page, but it does not directly impact rankings.
* <https://www.searchpilot.com/resources/case-studies/testing-schema-markup> - This resource emphasizes that incorporating schema markup ensures websites remain competitive as search engines evolve, even though it is not a direct ranking factor.
* <https://www.searchenginejournal.com/ranking-factors/schema-markup-ranking/> - This article reinforces the notion that while schema markup is useful for search engines, it is not used as a ranking factor by Google.
* <https://www.searchenginejournal.com/seos-are-recommending-structured-data-for-ai-search-why/537959/> - Please view link - unable to able to access data