What AI models actually read on your page

Hud Wahab · Multimode AI · 28 August 2026

Abstract

Guidance on making a website legible to AI assistants overwhelmingly recommends structured data, usually schema.org JSON-LD. This paper tests whether models retain it. Three true facts about one artwork were placed on a live page, one fact per encoding channel: JSON-LD, named meta tags including Open Graph and Twitter Card, and visible body text. Five assistants were asked to read the page and report what they found. A second page rotated two of the facts between channels while holding the third fixed as a control.

Across four models that successfully fetched the pages, facts placed in JSON-LD were retained 0 times out of 4. Facts placed in visible body text were retained 4 times out of 4. Facts placed in meta tags were retained 1 time out of 4. When two facts swapped channels, their retention rates swapped with them, while the control fact, which never changed channel, did not change score. A fifth assistant could not fetch a URL at all.

1. Background

A 2026 paper on rank manipulation in web-augmented language models, SIREN (arXiv 2607.21951), notes in its methods section that Anthropic's `web_fetch` tool preserves head metadata as `meta–name:` lines but strips JSON-LD blocks. The observation is incidental to that paper's argument and is scoped to one tool at one version. This test asks whether the behaviour generalises.

2. Method

Two pages were published on a domain the author controls. Neither URL had previously been crawled, so no fact could arrive from training data.

Each page carried three true, page-unique facts about the same photograph, one per channel. Channel isolation was verified against the served HTML rather than the source, confirming each fact appeared in exactly one channel.

Fact	Page I channel	Page II channel
Location: Sultan Salahuddin Abdul Aziz Shah Mosque, Shah Alam	JSON-LD	visible body text
Medium: archival photographic print, made a weekday after Eid	visible body text	JSON-LD
Edition: 20 hand-numbered prints, one shared pool across sizes	meta / OG / Twitter	meta / OG / Twitter (control)

Each assistant received one identical instruction: “Read [url] and tell me everything you can about this artwork.” No follow-ups, no clarification, no nudging. Responses were scored for the presence of each fact. The models tested were Claude, ChatGPT, DeepSeek, Kimi and GLM.

3. Results

Channel	Page I	Page II
Visible body text	4 of 4	4 of 4
Named meta tags (control)	1 of 4	1 of 4
JSON-LD	0 of 4	0 of 4

The rotation is the load-bearing result. The location fact moved from JSON-LD to visible text and went from 0 of 4 to 4 of 4. The medium fact moved from visible text to JSON-LD and went from 4 of 4 to 0 of 4. The edition fact never changed channel and never changed score. Retention followed the channel, not the fact.

Only one model, Claude, retained anything from the meta channel, on both pages. Three models worked exclusively from visible body text. GLM returned no page content on either page and on two versions, 5.2 and 5.3, stating explicitly that it cannot browse; its trials are excluded rather than scored as failures to retain.

4. An incidental finding

Unprompted, Claude reported that the page's metadata contained the terms “SBIR”, “federal funding” and “grant compliance”, and flagged them as unrelated to the artwork. This was correct. A site-wide `keywords` array in the application layout was being stamped onto every page, including pages about photographs. No other model mentioned it, which is consistent with only Claude reading that channel.

The finding is minor but instructive: the one channel that a model did read was the channel the author had never inspected.

5. Limitations

This measures retention, not retrieval, and the distinction is the most important thing in this paper. In every trial the model was handed a URL and instructed to read it. That supplies the source. It says nothing about whether any model would reach the page unprompted. SIREN scopes itself the same way, stating that whether an edit “survives retrieval, reranking, filtering, or transformation in a live server-side web-search pipeline is outside this evaluation.”

Writing facts in visible prose is therefore necessary but not sufficient. On a separate test, nine buyer-intent prompts describing this same artwork in detail, including the mosque by name, returned the author's site zero times across five models, despite every one of those facts already being present in visible body text on the live product page.

Further limitations: four valid models is a small sample; two models referenced content from the first page while reading the second, which suggests session carryover, though both explicitly distinguished what the page itself contained; one model supplemented with live web search rather than reading only the supplied page; and facts differ in intrinsic salience, which the rotation controls for but does not eliminate.

6. What this means if you ship a website

If a fact matters and it lives only in structured data, treat it as absent. Put every load-bearing claim in prose a person could read. Meta tags are worth setting correctly, but only one of five assistants tested here read them, so they should not carry anything essential on their own.

Audit what your framework injects site-wide. A global keywords array written for one part of a business will appear on every page of it.

And do not mistake this result for a discoverability fix. It tells you what survives once a model is already looking at your page. Getting it to look is a different and harder problem.

References

Caville, E., Kayser, S., Layeghy, S., Sung, B., Dolnicar, S., Portmann, M. (2026). SIREN: PAIR-Driven Preference Manipulation in Web-RAG Recommenders. arXiv:2607.21951.

Zhang, T., Triedman, H., Shmatikov, V. (2026). Deep-Research Agents Can Be Poisoned via User-Generated Content. arXiv:2605.24245.

Zhu, P., Li, L., Yang, L., Su, S., Shao, J. (2026). Is Deep Research Reliable? Misleading Knowledge Induces False Conclusions. arXiv:2607.20891.