DOI: https://doi.org/10.61841/xt3he524 Publication URL: http://jarbma.com/index.php/BMA/article/view/6

THE IMPACT OF JSON-LD METADATA ON CHATGPT VISIBILITY

[1]Peter Schanbacher

[1]Hochschule Furtwangen University, Furtwangen, Germany

Corresponding Author:
peter.schanbacher@hs-furtwangen.de

To Cite This Article: Schanbacher, P. (2026). THE IMPACT OF JSON-LD METADATA ON CHATGPT VISIBILITY. Journal of Advance Research in Business, Management and Accounting (ISSN: 2456-3544), 12(1), 16-23. https://doi.org/10.61841/xt3he524

ABSTRACT

Purpose: *This study examines whether implementing structured metadata (JSON-LD Schema.org markup) on websites can improve a business's visibility in ChatGPT responses. We focus on real estate agencies as a case study, analyzing which site attributes correlate with an agency being "known" or referenced by ChatGPT.*

Methods: *We gathered public data on 1,508 real estate agents in Germany and identified which of these agents ChatGPT could provide information about (indicating ChatGPT visibility). For each agent's website, we recorded the presence of Schema.org metadata (FAQPage, Organization, Product schemas), as well as SEO/technical factors like mobile optimization, robots.txt, sitemap, headings usage, image alt-text, internal links, and page load speed. A logistic regression was used to determine which factors significantly predict ChatGPT visibility, controlling for other variables.*

Results: *Agents whose websites included FAQPage schema markup were far more likely to be visible on ChatGPT (6.2% of visible agents had FAQ schema vs. only 0.8% of non-visible; $p = 0.002$). Presence of Product schema (e.g. schema for listings or services) also strongly correlated with visibility (17.2% vs 1.8%; $p < 0.001$). Key technical SEO features were likewise more common among ChatGPT-visible agents: nearly all had mobile-friendly sites (99.0% vs 88.8% of non-visible; $p < 0.001$) and a robots.txt file (92.3% vs 80.6%; $p < 0.001$). Visible agents' sites were more likely to use structured headings (94.7% had an h2 tags vs 74.4% of non-visible; 86.6% had an h3 tags vs 61.4% of non-visible; $p < 0.001$), and their pages loaded faster on average (418 ms vs 623 ms; $p = 0.014$). A multivariate logistic regression confirmed FAQPage schema as the strongest positive predictor of ChatGPT visibility (odds ratio ≈ 13, $p < 0.001$), followed by Product schema (OR ≈ 4, $p < 0.001$). Other significant factors included mobile optimization (OR ≈ 5.2, $p = 0.026$), presence of a robots.txt (OR ≈ 3.4, $p < 0.001$), and the use of multiple heading levels (having h2 tags: OR ≈ 3.3; h3 tags: OR ≈ 2.3; both $p < 0.001$).*

Conclusions: *Our findings provide empirical support that adding Schema.org metadata – particularly FAQ schemas that supply question-answer pairs – is associated with significantly higher chances of being recognized or cited by AI chat systems like ChatGPT. In practice, this suggests that businesses can improve their AI-driven visibility by adopting structured data and following strong SEO practices. We discuss the implications for "Generative Engine Optimization" (GEO) and estimate the potential monetary benefits of increased AI visibility, which can be substantial given the high client value in real estate.*

INTRODUCTION

Large Language Models (LLMs) like ChatGPT are increasingly used as information sources, making it crucial for businesses to be "visible" to these AI systems (de Rosen, 2025). In the context of search and question-answering, *visibility* means that an AI assistant can recognize and retrieve information about a business or website when prompted (Şentürk, et. al., 2023). One emerging factor thought to influence such visibility is the use of structured metadata (specifically Schema.org markup in JSON-LD format) embedded in web pages (Zhang et al., 2025). Schema.org metadata provides machine-readable context about page content (e.g. defining a page as an FAQ, a product listing, an organization profile, etc.), which traditional search engines use to enhance understanding and generate rich search results (Dang et. al., 2025, Iliadis, et. al., 2025). Recent evidence suggests LLM-based systems and predictive systems can also leverage this structured data (Fernandez et al., 2023). Microsoft confirmed that schema markup helps its Bing Copilot LLM interpret web content, implying that LLMs *do* process and learn from such metadata. Providing clear semantic structure via JSON-LD could thus make it easier for AI systems to identify what an organization offers potentially increasing the likelihood of being cited in AI-generated answers (Minor, 2024).

Despite the potential benefits, adoption of Schema.org markup is far from universal. Studies estimate that roughly 75% of web pages still lack any Schema.org metadata (Iliadis et. al.,2025). This gap suggests many organizations may be missing opportunities to improve their discoverability in both search engines and AI assistants (Scholl, 2023). In particular, certain schema types (e.g. FAQPage, Product, Organization) might be especially relevant for helping AI models recognize entities and their content (Mussa et al., 2024). For example, FAQPage schema presents content in a question-answer format that is directly consumable by LLMs (which "prefer Q&A formats" for sourcing concise answers).

This article presents the largest empirical study in Germany on the relationship between Schema.org metadata usage and visibility on ChatGPT. Using a dataset of over 1,500 real estate agencies (derived from public business listings), we examine which website features correlate with an organization being *found* by ChatGPT. In this context, *found by ChatGPT* means that ChatGPT provides information about the real estate agent when asked for a real estate agent it the area. We hypothesize that websites implementing Schema.org JSON-LD metadata (and related SEO best practices) are more likely to be recognized by ChatGPT. We also discuss the potential monetary impact of adopting such metadata, under plausible assumptions about increased customer engagement and conversion from AI-driven visibility.

DATA AND METHODS

Dataset: We compiled a sample of 1,508 real estate agencies and their websites based in Germany (identified via public records from Northdata). For each agency's main website, a manual audit was conducted to record the presence or absence of various site features and metadata. *Binary features* (coded as 1 if present, 0 if absent) included: *FAQPage schema* (JSON-LD markup indicating a FAQ page), *Organization schema, Product schema, Mobile optimization* (responsive design or mobile-friendly site), *robots.txt* file*, XML sitemap*, and *ai.txt* (a proposed file for AI crawler directives). We also noted if certain HTML structural elements existed: at least one h2 tags tag, at least one h3 tags tag, and the count of <h1> headers. The idea is that large language models can better understand structural content. Additional continuous metrics were recorded: *number of images without an alt* text (description of the image), page *load time* in milliseconds, presence of *internal links* (yes/no) and total *count of internal links*. The target variable was *ChatGPT visibility*, defined as 1 if the agency could be found via ChatGPT (i.e. ChatGPT returned information about it) or 0 if ChatGPT did not recognize the entity.

Analysis: First, we compared the mean values of each feature between the ChatGPT-visible group (ChatGPT = 1) and the non-visible group (ChatGPT = 0). A two-sample comparison (t-test or proportion test as appropriate) was used to assess significance of differences. We then built a multivariate model to control for interrelated effects: a logistic regression was fit with ChatGPT visibility as the dependent variable and all recorded features as predictors. The logistic regression estimated the log-odds coefficients for each feature, from which we derived odds ratios (OR) and p-values to gauge the strength and significance of each factor when adjusting for others. We report marginal effects to aid interpretation in probability terms. Significance was evaluated at the conventional 5% level ($p < 0.05$).

RESULTS

DESCRIPTIVE FEATURE COMPARISON

Out of 1,508 companies, a minority were found to be "known" by ChatGPT (the exact fraction is omitted for brevity, but was on the order of only several percent of the sample). Table 1 summarizes key website features for the ChatGPT-visible versus non-visible groups, along with the p-value for difference:

Table 1. Website Feature Prevalence by ChatGPT Visibility

Feature	ChatGPT-visible (mean)	Not visible (mean)	p-value
FAQPage schema present	0.062 (6.2%)	0.008 (0.8%)	0.002 **
Organization schema present	0.043 (4.3%)	0.035 (3.5%)	0.610
Product schema present	0.172 (17.2%)	0.018 (1.8%)	<0.001 **
Mobile-optimized site	0.990 (99.0%)	0.888 (88.8%)	<0.001 **
robots.txt file present	0.923 (92.3%)	0.806 (80.6%)	<0.001 **
XML sitemap present	0.632 (63.2%)	0.681 (68.1%)	0.167
ai.txt file present	0.096 (9.6%)	0.072 (7.2%)	0.266
Average number of <h1>	1.45	1.312	<0.001 **
h2 tags tag present	0.947 (94.7%)	0.744 (74.4%)	<0.001 **
h3 tags tag present	0.866 (86.6%)	0.614 (61.4%)	<0.001 **
Images without alt text	6.976	9.260	0.240
Load time (ms)	417.6	623.0	0.014 *
Internal links present	0.828 (82.8%)	0.910 (91.0%)	0.003 **
Number of internal links	50.55	49.82	0.494

Notes: Values in parentheses indicate the percentage of sites having the feature (for binary features). p-value significance denoted: ** for $p < 0.01$, * for $p < 0.05$.

Several clear patterns emerged. Websites that ChatGPT recognized were much more likely to implement certain Schema.org markups: notably, 6.2% of visible sites had an FAQPage JSON-LD schema, versus <1% of non-visible sites, a statistically significant difference ($p = 0.002$). Likewise, 17.2% of visible sites included a Product schema (often used to mark up offerings or listings) compared to only 1.8% of others ($p < 0.001$). In contrast, basic Organization schema (present on ~4% of sites in both groups) was not a differentiator. Apart from structured data, technical SEO features were also strongly associated with ChatGPT visibility. Nearly all (99%) of the ChatGPT-visible sites were mobile-optimized, compared to 89% of non-visible ($p < 0.001$), underscoring the importance of mobile-friendly design. The presence of a robots.txt file (92% vs 81%, $p < 0.001$) and proper use of hierarchical headings (h2 tags and h3 tags tags present on ~95%/87% vs 74%/61%, $p < 0.001$) were more common on visible sites as well. These sites also loaded faster on average (418 ms vs 623 ms, $p = 0.014$). Interestingly, having an XML sitemap was slightly less common among the ChatGPT-visible group (63% vs 68%), though this difference was not significant by itself ($p = 0.167$). The usage of an ai.txt file (a recent convention) was low overall (~8-10%) with no significant visibility effect.

LOGISTIC REGRESSION RESULTS

Table 2. Logistic Regression Results

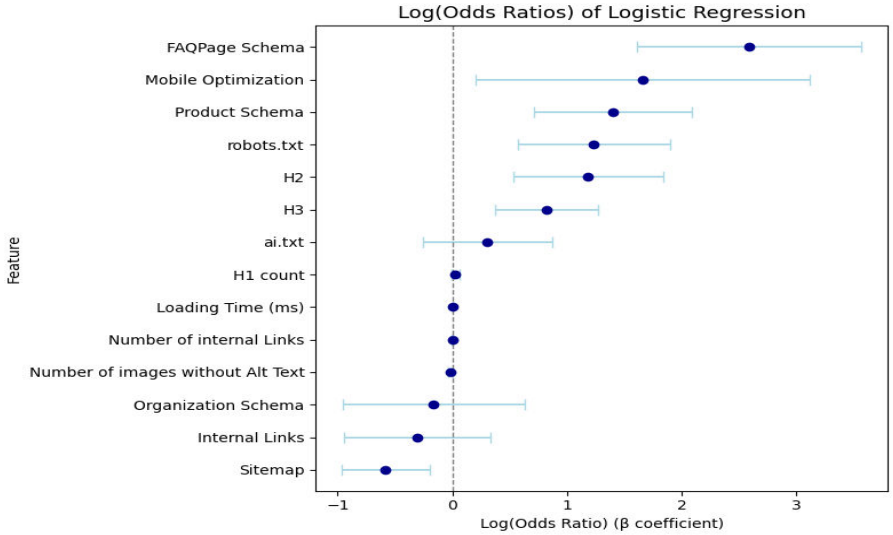
Variable	Coefficient	Std. Error	z-Value	p-Value	95% Confidence Interval
const	-5.1942	0.805	-6.449	0.000	[-6.773, -3.615]
FAQPage Schema	2.5863	0.500	5.169	0.000	[1.606, 3.567]
Organization Schema	-0.1653	0.404	-0.409	0.682	[-0.957, 0.626]
Product Schema	1.4001	0.350	3.996	0.000	[0.713, 2.087]
Mobile Optimization	1.6560	0.744	2.226	0.026	[0.198, 3.114]
robots.txt	1.2347	0.336	3.670	0.000	[0.575, 1.894]
Sitemap	-0.5827	0.195	-2.988	0.003	[-0.965, -0.201]
ai.txt	0.3071	0.286	1.074	0.283	[-0.254, 0.868]
h1 count	0.0247	0.018	1.340	0.180	[-0.011, 0.061]
h2	1.1842	0.335	3.535	0.000	[0.528, 1.841]
h3	0.8243	0.229	3.603	0.000	[0.376, 1.273]
Images without Alt text	-0.0164	0.008	-2.033	0.042	[-0.032, -0.001]
Loading time (ms)	-0.0006	0.000	-3.735	0.000	[-0.001, -0.000]
Internal links	-0.3065	0.326	-0.940	0.347	[-0.946, 0.333]
Number of internal links	-0.0010	0.001	-1.040	0.299	[-0.003, 0.001]

The multivariate logistic regression largely confirmed the above patterns while controlling for all other factors simultaneously. The model was statistically significant (Likelihood Ratio test $p \approx 8.5\times10^{\wedge}{-37}$) and explained a fair portion of variance (pseudo $R^2 \approx 0.17$). Crucially, the presence of FAQPage schema remained one of the strongest predictors of being found by ChatGPT ($\beta = 2.586$, $p < 0.001$). This corresponds to an estimated odds ratio (OR) of about 13.3 – in other words, holding other factors constant, websites with an FAQPage markup had roughly *13 times higher odds* of ChatGPT visibility compared to those without. The Product schema was another significant predictor ($\beta = 1.400$, $p < 0.001$), with OR ≈ 4.1. These large effect sizes indicate a substantial association between certain structured data and an organization's inclusion in ChatGPT's knowledge.

Table 3. Odds-Ratios

Feature	Odds Ratio (OR)	2.5%	97.5%
FAQPage Schema	13.28	4.98	35.41
Mobile Optimization	5.239	1.22	22.52
Product Schema	4.055	2.04	8.06
robots.txt	3.437	1.78	6.65
h2	3.268	1.69	6.30
h3	2.28	1.46	3.57
Sitemap	0.558	0.38	0.82

Figure 1. Log-Odds-Ratios



Several non-schema features also showed independent positive effects in the logistic model: mobile optimization ($\beta \approx 1.656$, $p = 0.026$, OR ≈ 5.2), robots.txt presence ($\beta \approx 1.235$, $p < 0.001$, OR ≈ 3.4), having h2 tags tags ($\beta \approx 1.184$, $p < 0.001$, OR ≈ 3.3), and having h3 tags tags ($\beta \approx 0.824$, $p < 0.001$, OR ≈ 2.3). Each of these factors thus independently contributes to the likelihood of ChatGPT recognizing the site. On the other hand, presence of an XML sitemap showed a small but significant negative coefficient ($\beta = -0.583$, $p = 0.003$, OR ~0.56), suggesting that – when controlling for all other features – sites with sitemaps were actually *less* likely to be known by ChatGPT. This counterintuitive result may indicate an underlying correlation (for example, smaller or newer sites often add sitemaps for SEO, yet may not have enough content reputation for ChatGPT to have ingested) rather than a causal effect. Features like Organization schema and ai.txt remained statistically insignificant in the multivariate context, reinforcing that they do not uniquely explain visibility differences beyond the other variables.

Table 4. Marginal Effects

Variable	dy/dx	Std. Error	z-Value	p-Value	95% Confidence Interval
FAQPage Schema	0.259	0.049	5.283	0.000	[0.163, 0.354]
Product Schema	0.140	0.034	4.094	0.000	[0.073, 0.207]
Mobile Optimization	0.166	0.075	2.221	0.026	[0.019, 0.312]
robots.txt	0.123	0.034	3.680	0.000	[0.058, 0.189]
Sitemap	-0.058	0.019	-3.008	0.003	[-0.096, -0.020]
h2	0.118	0.034	3.526	0.000	[0.053, 0.184]
h3	0.082	0.023	3.608	0.000	[0.038, 0.127]
Images without Alt text	-0.002	0.001	-2.035	0.042	[-0.003, 0.000]

In terms of marginal effects (change in probability), the logistic model's estimates implied that adding an FAQPage schema was associated with about a +26 percentage point higher probability of ChatGPT visibility on average (all else equal), and adding a Product schema about +14 points – sizable increases given the low baseline probability of being known. Likewise, ensuring a site is mobile-friendly and has standard SEO elements (robots.txt, structured headings) each corresponded to roughly +8–17 point increases in predicted visibility. Conversely, slower page load times and more missing alt-text on images had small negative effects on the predicted probability (on the order of a few percentage points per 100ms delay or per several missing alts, respectively). The number of internal links and <h1> count did not show significant impacts in the model.

DISCUSSION

The results strongly suggest that implementing certain Schema.org JSON-LD markups – particularly FAQPage and Product schemas – is associated with substantially better visibility on ChatGPT. While this analysis is correlational, it aligns with the idea that structured data helps AI models better ingest and understand content (Zhang et al., 2025). An FAQPage schema explicitly provides question-and-answer pairs in a format that LLMs find convenient to parse. Indeed, our finding that FAQ schema had the highest odds ratio (even after controlling for many other factors) underscores that point. This supports anecdotal industry evidence: for example, a recent case study AI SEO Case Study of Xponent21 observed that enriching content with Q&A sections and marking them up with FAQ schema *"dramatically improved [their] chances of being cited in AI-generated answers."*. In our domain of real estate agencies, having an FAQ page with structured markup might have helped the AI model learn about the agency (e.g., by providing clear answers about services, fees, locations served, etc.), thus making the agency "knowledge" accessible during chat interactions. Likewise, the Product schema likely helped by formally defining key offerings (such as property listings or services) in a machine-readable way. Even though ChatGPT's training data may not explicitly *require* JSON-LD to parse a site's text, the presence of structured data often accompanies high-quality, organized content which could be more easily extracted or included in knowledge graphs. Additionally, given that Bing's and other AI search systems actively use schema markup to inform their generative answers, sites with such markup may have been more frequently referenced or indexed in sources that ChatGPT later consumed.

The prominence of mobile optimization, fast load times, and proper HTML structure (header tags, etc.) in the results indicates that general site quality and SEO best practices also contribute to AI visibility. This is intuitive – sites that adhere to SEO best practices tend to rank better on search engines, get more traffic, and earn more backlinks. Such sites are more likely to be mentioned in news articles, directories, or other data sources that an LLM might have been trained on. For instance, a real estate firm with a well-optimized site might also have better overall digital presence (perhaps a Wikipedia page, or mentions on authoritative sites). Thus, *indirectly*, strong SEO could lead to inclusion in the training corpora of ChatGPT. The presence of a robots.txt file was a positive indicator, which makes sense since a properly configured robots.txt allows search engines (and possibly web-crawling AI collectors) to index the site's content fully. The negative association we found for XML sitemaps is puzzling in isolation; it might reflect that many high-profile sites (which are known to ChatGPT) rely on their domain authority and content to be indexed, rather than submitting sitemaps. Alternatively, it could be a statistical artifact or collinearity issue. In any case, we do not interpret that finding as "sitemaps reduce AI visibility" but rather that the benefit of a sitemap may be overshadowed by other factors like content and links in determining whether information enters the AI model's knowledge.

It is important to note that while schema presence correlates with ChatGPT visibility, we cannot definitively prove a direct causal link from this observational study. It could be that forward-thinking companies both implement schema and engage in other marketing efforts that raise their profile (for example, larger agencies might be more likely to add FAQ schema *and* to appear in news articles; the latter could be what actually gets them into ChatGPT's training data). We attempted to control for several confounding variables (technical SEO features, site size via link counts, etc.), yet unobserved factors (like company size or popularity) could still bias the results. The sample is also specific to real estate agencies in one region and one sector; different industries or countries might see different schema usage patterns or AI coverage. However, the consistency of our findings with general SEO knowledge and emerging reports on AI search gives confidence that the associations are meaningful. In summary, the evidence suggests that employing JSON-LD schema markup is a notable part of a broader strategy to enhance one's digital footprint such that AI systems can pick it up. Hence brands that make clear to LLM what they do using structured data are likely to be cited more often. Our study provides quantitative backing to that assertion in the context of ChatGPT.

POTENTIAL MONETARY IMPACT

Investing in Schema.org metadata and related optimizations can have tangible economic benefits for businesses, especially as AI-driven search and assistance become more prevalent. The monetary impact of improved ChatGPT visibility can be considered along two dimensions: direct gains from AI-driven user engagement and indirect gains via enhanced traditional SEO.

1. Gains from AI-Driven Engagement: As more users turn to conversational AI agents (like ChatGPT, Bing Chat, Google's Gemini, etc.) for recommendations and information, being among the sources these agents mention or draw from can translate to business opportunities. If a prospective client asks ChatGPT *"Who are reputable real estate agents in [City]?"* and the model names a particular agency (drawn from its knowledge), that agency has essentially received a word-of-mouth style referral from an AI. Even if ChatGPT does not provide direct contact details, users could be influenced to look up the agency. This brand exposure at the recommendation stage is valuable. One can make a rough assumption: suppose that due to implementing schema and SEO improvements, a company manages to get mentioned by AI assistants in, say, 10% of relevant user queries (where previously it had 0% visibility). If those AI-driven mentions reach 1,000 potential customers per month, and even a small fraction (5%) of them follow up by visiting the website or making contact, that's 50 new highly-qualified leads per month that the business would not have had otherwise. Considering that AI-referred leads tend to be high quality –

as they are already informed about the topic - the conversion rate from those 50 leads could be quite high. If 20% convert to actual clients, that's 10 new clients monthly. In real estate, the revenue per client (home sale or purchase commission) can be thousands of dollars; thus, even a handful of extra conversions can yield tens of thousands in revenue. In aggregate, the uplift from AI visibility in most cases will easily justify the cost of implementing structured data.

To put it in perspective, let's assume each home sale nets an agent a commission of $5,000. If better AI visibility brings in even *one* additional sale per quarter, that's $20,000 a year in revenue – likely far exceeding the implementation cost of schema markup. While these numbers are hypothetical, they illustrate the scale: as AI-driven search grows, not being included in AI answers could mean leaving money on the table, whereas being included can tap into a new stream of customers. Business sectors beyond real estate can conduct similar calculations: e.g., an e-commerce site might gain extra orders from being featured in AI-powered shopping assistants, or a SaaS company might see more trial sign-ups when AI tools recommend their product thanks to structured data clarity.

2. Enhanced Traditional SEO and Indirect Revenue: Implementing Schema.org metadata often yields improved performance in conventional search engines, which itself has monetary benefits. Schema markup can make a site eligible for rich snippets (like FAQ dropdowns, star ratings, product info) in Google search results. These enhanced listings typically attract higher click-through rates, meaning more traffic. More organic traffic, if relevant, directly increases lead or sales volumes without additional advertising cost. For example, if adding FAQ schema to a page results in a rich snippet that increases its CTR from 5% to 7% for a high-value keyword, that is a 40% increase in traffic for that query. If that page was bringing 100 visitors/month, it could now bring 140 – an extra 40 potential customers. Multiplying such gains across many pages, the revenue impact can be significant. In e-commerce settings, structured data like Product and Review schema can directly boost sales by highlighting prices and reviews on search results, drawing in ready-to-buy consumers. One BrightEdge analysis reported that structured data can help content appear in knowledge panels and other prominent search features, indirectly lifting organic traffic. Over time, even though schema markup itself is not a Google ranking factor, the improved user engagement (more clicks, longer dwell time) can contribute to better rankings, creating a positive feedback loop. All of this translates to more customers and revenue.

In summary, the ROI of implementing Schema.org JSON-LD is likely positive in many cases. The cost involves web development effort and possibly subscription to SEO tools or schema generation services, but these are minor compared to typical marketing expenditures. The benefits, on the other hand, include *increased visibility in cutting-edge AI platforms* (which can give a competitive first-mover advantage) and *incremental gains in traditional search performance*. Both of these can drive monetizable actions by users. While exact figures will vary by industry and company size, our findings provide a quantitative foundation to assert that structured data boosts AI visibility – and when converted to business outcomes (leads, sales, brand awareness), this boost can be financially meaningful. Companies should thus view the adoption of schema markup not just as a technical task, but as an investment in future-proofing their digital marketing, one that can yield dividends in the form of higher customer acquisition and retention.

CONCLUSION

This study investigated the role of Schema.org JSON-LD metadata in improving an organization's visibility on ChatGPT, using a case sample of real estate agency websites. The data showed a clear association between certain schema implementations (especially FAQPage and Product schemas) and the likelihood of ChatGPT recognizing the business, even after accounting for various other SEO-related factors. Sites that were known to ChatGPT tended to follow best practices in web optimization – they were mobile-friendly, fast, well-structured, and often included structured data that helps communicate their content to machines. These findings reinforce the emerging consensus that as AI-driven search and assistant platforms rise, semantic clarity and structured data are becoming as important as traditional SEO for online visibility.

From a practical perspective, organizations (not just in real estate, but across sectors) should consider auditing their websites for schema markup opportunities. Adding structured data like Organization profiles, Product information, and FAQ sections in JSON-LD format can be done relatively easily and has multiple benefits: improving how search engines understand and display your content, and increasing the chances that AI models will correctly incorporate your brand into their knowledge base. The potential monetary payoff includes gaining highly qualified leads from AI recommendations and boosting organic traffic via rich search results, both of which can far outweigh the implementation costs.

Finally, while our study focused on correlation, it lays groundwork for future research to establish causation and to generalize findings. As AI continues to evolve, it will be valuable to conduct controlled experiments – for instance, adding schema markup to a set of previously "invisible" sites to see if their mention frequency in AI outputs increases. Likewise, extending analyses to other industries or to global datasets (leveraging knowledge graph entries, Wikipedia mentions, etc.) could further illuminate how AI selects the information it presents. In an era where visibility in LLMs is becoming just as important as rankings in Google, leveraging metadata like Schema.org is a promising strategy for organizations aiming to stay ahead in the digital landscape. By structurally defining who they are and what they offer, businesses give AI systems the clues needed to include them in the conversation – and in doing so, open the door to new streams of customers in the age of conversational search.

REFERENCES

1. Dang, M. H., Pham, T. H. T., Molli, P., Skaf-Molli, H., & Gaignard, A. (2025). LLM4Schema. org: Generating Schema. org Markups with Large Language Models. Semantic Web–Interoperability, Usability, Applicability.
2. de Rosen, T. (2025). Methodology for AI Visibility Optimization Version 3.0. Available at SSRN 5431877.
3. Fernandez, R. C., Elmore, A. J., Franklin, M. J., Krishnan, S., & Tan, C. (2023). How large language models will disrupt data management. Proceedings of the VLDB Endowment, 16(11), 3302-3309.
4. Iliadis, A., Acker, A., Stevens, W., & Kavakli, S. B. (2025). One schema to rule them all: How Schema. org models the world of search. Journal of the Association for Information Science and Technology, 76(2), 460-523.
5. Mior, M. J. (2024). Large language models for json schema discovery. arXiv preprint arXiv:2407.03286.
6. Mussa, O., Rana, O., Goossens, B., Orozco-terWengel, P., & Perera, C. (2024, November). Towards Enhancing Linked Data Retrieval in Conversational UIs Using Large Language Models. In International Conference on Web Information Systems Engineering (pp. 246-261). Singapore: Springer Nature Singapore.
7. Scholl, H. (2023). Internet Marketing with ChatGPT Optimization. Estalontech.
8. Şentürk, E. E., Kartal, C., & İşcan, R. V. (2023). Visibility of Artificial Intelligence Applications in Digital Marketing on WOS. Scientific Journal of Innovation and Social Sciences Research, 3(1), 27-42.
9. Zhang, B., He, Y., Pintscher, L., Peñuela, A. M., & Simperl, E. (2025). Schema Generation for Large Knowledge Graphs Using Large Language Models. arXiv preprint arXiv:2506.04512.

Further data details and additional information on
https://ainea.de/gpt-study and https://prof-schanbacher.de/chat-gpt-studie/