International Journal of Law and Society
2026, Vol. 9, No. 3, pp. 361–369
https://doi.org/10.11648/j.ijls.20260903.16

SciencePG
Science Publishing Group

Research Article

Schema Markup Adoption in Personal Injury Law Firm Websites: A Systematic Analysis of Structured Data Implementation Across North American Legal Services

Behzad Hussain*

Rank Brilliance, Semantic SEO Research, Lahore, Pakistan

Abstract

Structured data markup, particularly JSON-LD implementations of the Schema.org vocabulary, has become a critical factor in how search engines interpret, classify, and surface web content. Despite growing recognition of schema markup's role in search engine optimization (SEO), there has been limited empirical investigation into its adoption within specific professional service verticals. This study presents a systematic analysis of schema markup implementation across 500 personal injury (PI) law firm websites operating in the United States. Through automated crawling and programmatic code inspection, we examine the prevalence of key schema types, including LegalService, Attorney, Organization, FAQPage, BreadcrumbList, and WebPage, and assess the completeness, accuracy, and semantic richness of deployed structured data. Our findings reveal significant gaps: 67.6% of sampled firms implement some form of JSON-LD markup, yet only 40.0% deploy the LegalService schema type specifically designed for legal service providers. The mean Schema Completeness Index (SCI) across sites with structured data was 11.8 out of a possible 25. Entity disambiguation remains the weakest dimension: only 84.0% of sites with schema include @id properties and 81.4% include sameAs references. These findings have implications for legal services discoverability in both traditional search engine results pages (SERPs) and emerging AI-mediated answer engines. We propose a Structured Data Maturity Model for legal service websites and outline directions for future research.

Keywords

Schema Markup, JSON-LD, Structured Data, Personal Injury Law, Legal Services SEO, Schema.org, Search Engine Optimization, Entity Optimization

1. Introduction

The digital landscape for legal services has undergone a fundamental transformation over the past decade. Personal injury (PI) law firms, which once relied primarily on television advertising, referral networks, and print media, now compete in an increasingly complex digital ecosystem where organic search visibility directly influences client acquisition. Google processes an estimated 8.5 billion searches per day, and a substantial proportion of individuals seeking legal representation for personal injury matters begin their search online [5, 6]. The ability of a law firm's website to appear prominently in search engine results pages (SERPs) has become a primary competitive differentiator.

*Correspondence: Behzad Hussain (behzad@rankbrilliance.com)

Received: 12 June 2026; **Accepted:** 10 July 2026; **Published:** 10 August 2026

Search engines have evolved significantly beyond keyword-matching algorithms. Google's adoption of knowledge graph technology [17], the BERT language model [7], and more recently AI Overviews powered by large language models [1, 13, 15, 16], reflects a broader shift toward semantic understanding of web content. In this context, structured data markup, particularly JSON-LD implementations of the Schema.org vocabulary, serves as a machine-readable layer that helps search engines understand the meaning, relationships, and attributes of entities described on web pages [8, 10].

This paper presents a systematic empirical analysis of schema markup adoption across 500 personal injury law firm websites in the United States. Our objectives are threefold: (1) to establish a baseline measure of structured data adoption in the PI legal services vertical; (2) to identify common implementation patterns, errors, and omissions; and (3) to propose a Structured Data Maturity Model that can guide future research and practitioner efforts.

2. Literature Review

2.1. Structured Data and the Semantic Web

The concept of structured data on the web has its roots in Berners-Lee, Hendler, and Lassila's [3] vision of the Semantic Web. Schema.org, launched in 2011 as a collaborative initiative between Google, Microsoft, Yahoo, and Yandex, provided a standardized vocabulary for marking up web content with machine-readable metadata [4, 10, 12, 14]. JSON-LD (JavaScript Object Notation for Linked Data), now Google's preferred format for structured data, simplified implementation by decoupling structured data from HTML markup.

2.2. Semantic SEO and Topical Authority

Semantic SEO, as conceptualized by practitioners such as Gübür [11] extends traditional search engine optimization beyond keyword targeting toward entity-based optimization, topical authority construction, and contextual relevance signaling. Barnard [2] has similarly argued that establishing a clear, unambiguous digital identity through structured data is essential for entities seeking visibility in Google's Knowledge Graph [9]. In the context of personal injury law, where queries are characterized by high commercial intent and geographic specificity, the ability to provide comprehensive, well-structured information about specific practice areas directly influences ranking capacity.

3. Methodology

3.1. Sample Selection

Our sample comprises 500 personal injury law firm websites drawn from Google Search results for 20 high-volume PI-related queries across major metropolitan areas in the United States. URLs were collected via a commercial Google SERP API and deduplicated by root domain. Non-law-firm domains (directories, aggregators, social media) were filtered using domain-pattern and content heuristics.

3.2. Data Collection

Data collection was automated using a custom Python pipeline. For each website, we: (a) crawled the homepage and up to 20 internal pages matching practice-area, attorney, FAQ, and contact page URL patterns; (b) extracted all JSON-LD script blocks; (c) parsed and validated extracted JSON-LD against Schema.org specifications; and (d) scored each implementation using the Schema Completeness Index (described in 3.3).

3.3. Schema Completeness Index (SCI)

We developed a Schema Completeness Index to quantify structured data implementation quality. The SCI evaluates five dimensions, each scored 0–5: (1) Schema Type Coverage, (2) Property Completeness, (3) Entity Disambiguation, (4) Hierarchical Nesting, and (5) Validation Status. Maximum SCI is 25.

4. Findings

4.1. Overall Schema Adoption Rates

Of the 500 websites analyzed, 338 (67.6%) deployed at least one form of JSON-LD structured data. Organization schema was the most commonly deployed type, present on 172 sites (34.4%). LocalBusiness schema was found on 45 sites (9.0%). Only 200 sites (40.0%) deployed the LegalService schema type specifically recommended by Schema.org for legal service providers.

Table 1. Schema Type Adoption Rates Across 500 PI Law Firm Websites.

Schema Type	Sites (n)	Adoption (%)	Avg Property Completeness (%)
Organization	172	34.4%	39.9%
LocalBusiness	45	9.0%	57.8%
WebPage	79	15.8%	41.0%
BreadcrumbList	38	7.6%	100.0%
FAQPage	128	25.6%	29.5%
Person/Attorney	206	41.2%	33.1%
LegalService	200	40.0%	47.9%
Review/AggregateRating	151	30.2%	—
HowTo	1	0.2%	—
Article/BlogPosting	61	12.2%	76.4%

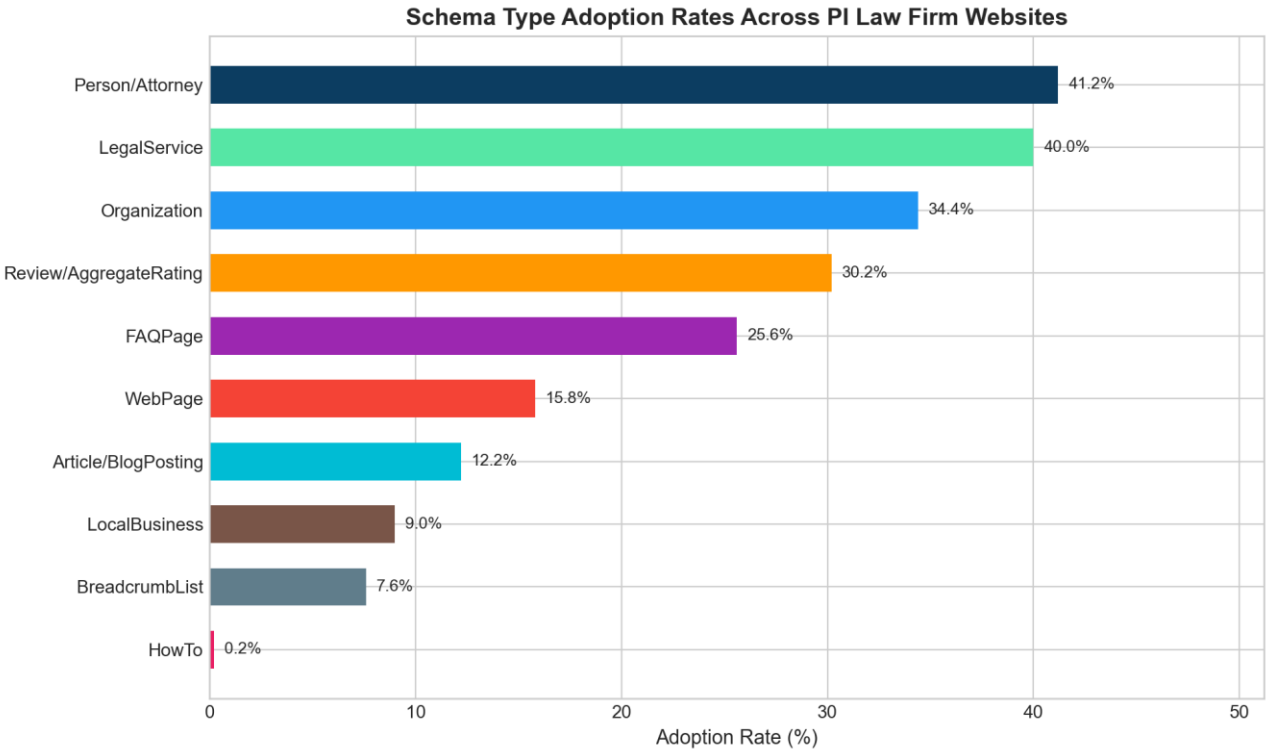

Figure 1. Schema type adoption rates across sampled PI law firm websites.

4.2. Schema Implementation Quality

The mean SCI across sites with structured data was 11.8 out of 25 (SD = 2.6). The median was 12.0 and scores ranged from 3 to 20. Entity disambiguation emerged as the weakest dimension: only 84.0% of sites with schema included @id properties and 81.4% included sameAs references.

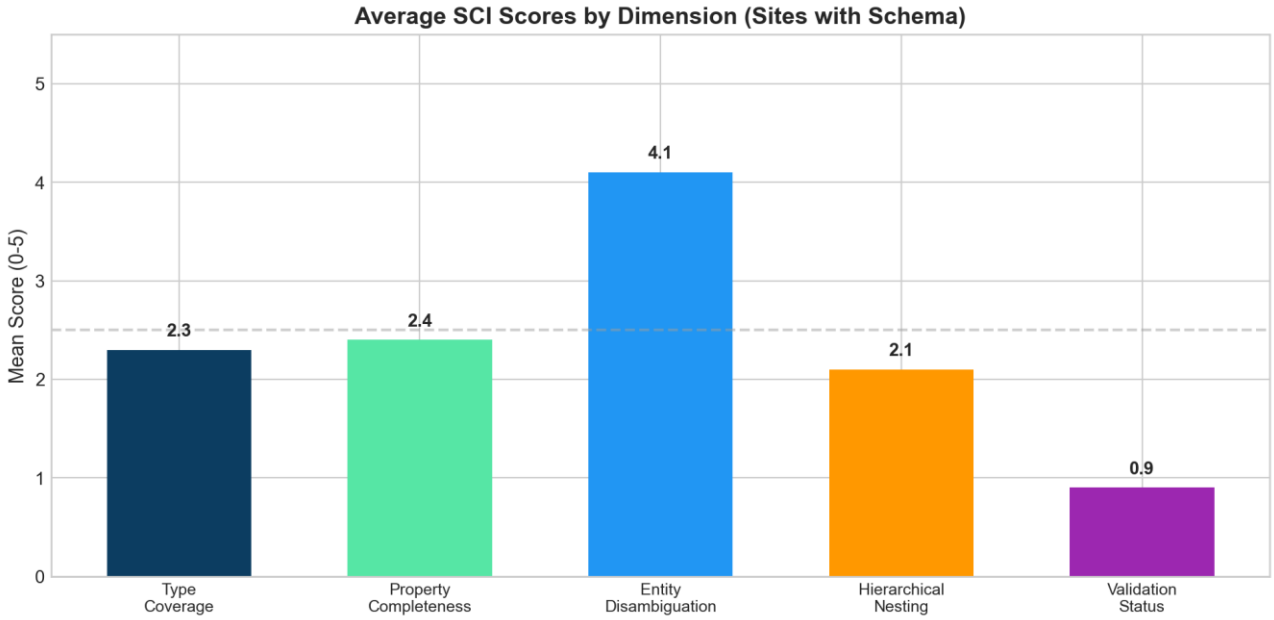

Figure 2. Mean SCI scores by dimension for sites with schema.

4.3. Common Implementation Errors

Validation testing revealed that 263 of 338 sites with schema (77.8%) contained at least one validation error. The most frequent errors were: missing required properties, unstructured address fields, and missing image/logo properties.

4.4. Firm Size and Schema Sophistication

Table 2. Schema Completeness Index by Firm Size.

Firm Size	n	Mean SCI	SD	With Schema (%)
Solo/Small (1-3 attorneys)	143	10.6	3.4	79.7%
Mid-size (4-20 attorneys)	58	10.8	4.2	77.6%
Large (21+ attorneys)	52	11.3	3.7	82.7%

Table 3. Structured Data Maturity Model, Distribution Across Sample.

Level	Description	Sites (n)	Percentage
Level 0	No Implementation	0	0.0%
Level 1	Basic Identity	150	30.0%
Level 2	Service Declaration	152	30.4%
Level 3	Entity Network	191	38.2%
Level 4	Semantic Authority	7	1.4%
Level 5	Full Semantic Integration	0	0.0%

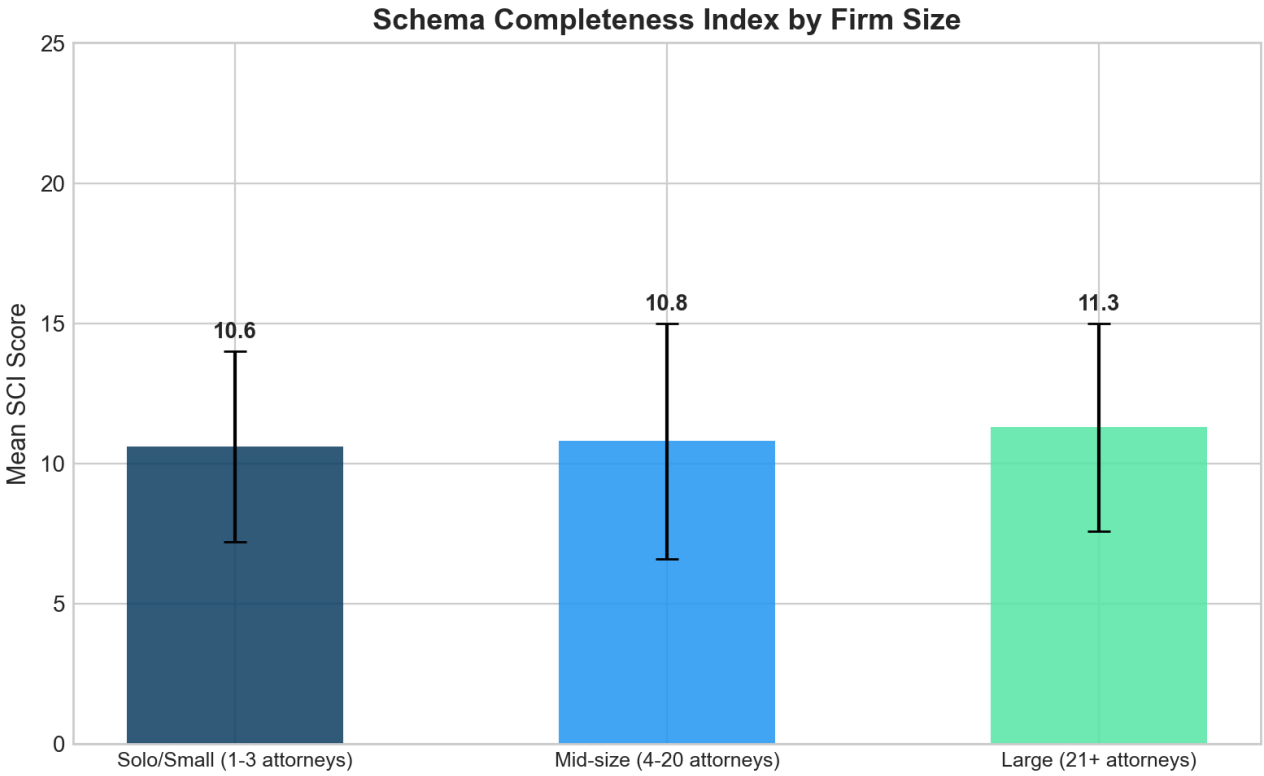

Figure 3. SCI scores by firm size category.

4.5. Schema Type Selection: Specific vs Generic

A central question in this study concerns whether PI law firms adopt the semantically precise LegalService schema type (along with related types such as Attorney and LawFirm) or rely on more generic business type schemas like Organization, LocalBusiness, and ProfessionalService. Our analysis reveals that only 293 of 500 sites (58.6%) employ at least one legal-specific schema type. In contrast, 84 sites (16.8%) rely exclusively on generic business type schemas, leaving 123 sites with no business type schema at all. This finding indicates that the vast majority of PI law firms are not signaling their legal practice classification to search engines through the most appropriate vocabulary, missing an opportunity for entity disambiguation in legal services verticals.

Table 4. Most Common Business Type Schema Combinations Used by PI Law Firms.

Schema Type Combination	Sites (n)
LegalService, Organization	137
LegalService	105
Organization	63
Attorney, LegalService, Organization	12
Attorney	11
LocalBusiness	10
LegalService, LocalBusiness, Organization	9
LocalBusiness, Organization	8

4.6. Property Coverage Within Business Type Schemas

Beyond schema type selection, the depth of property population reveals how completely each business is described to search engines. We examined property coverage across the most common business type schemas used by PI law firms.

Table 5. Property Coverage Across Business Type Schemas (% of sites using each type).

Property	Org (%)	LocalBiz (%)	LegalSrv (%)
name	98.8%	100.0%	100.0%
url	97.5%	97.1%	83.8%
logo	91.3%	42.9%	43.8%
image	59.5%	80.0%	90.4%
description	20.2%	51.4%	58.1%
telephone	12.0%	88.6%	69.9%
email	5.0%	31.4%	20.2%
address	12.8%	100.0%	93.8%
sameAs	69.8%	60.0%	52.6%
areaServed	3.7%	22.9%	32.4%
@id	86.8%	31.4%	51.5%
geo	0.8%	62.9%	43.8%
priceRange	1.2%	65.7%	51.5%
openingHoursSpecification	2.9%	28.6%	22.8%

4.7. @graph Usage and Nesting Quality

The @graph construct in JSON-LD allows multiple related entities to be combined into a single, semantically connected structure. Proper use of @graph (or hierarchical nesting) enables search engines to understand entity relationships, such as the connection between an Organization and its constituent Attorneys. Our analysis found that 321 sites (78.7% of those with any schema) employ the @graph wrapper to combine schema markup. Conversely, 3 sites (0.7%) deploy multiple top-level schema blocks that remain flat and disconnected, with no nesting or @graph relationships to signal entity associations to search engines. This fragmentation limits the semantic value of structured data, as search engines must infer relationships that could be made explicit.

4.8. @id Coverage on Key Schema Types

The @id property serves as a unique identifier that enables cross-referencing between schema entities, both within a single document and across the broader web. Without @id values, search engines cannot reliably link related entities or build coherent knowledge graph representations. We examined @id coverage across key schema types deployed by PI law firms.

Table 6. @id Property Coverage by Schema Type.

Schema Type	Sites Using Type	With @id	Coverage (%)
WebSite	336	314	93.5%
WebPage	305	292	95.7%
BreadcrumbList	284	267	94.0%

Schema Type	Sites Using Type	With @id	Coverage (%)
LegalService	272	140	51.5%
Organization	242	210	86.8%
Person	148	124	83.8%
Attorney	39	13	33.3%
LocalBusiness	35	11	31.4%
ProfessionalService	4	3	75.0%
LawFirm	1	1	100.0%

4.9. Person Schema for Attorney Representation

The Person schema type (along with the more specific Attorney type) provides a machine-readable mechanism for representing individual attorneys, their credentials, areas of expertise, and organizational affiliations. Our analysis shows that only 167 sites (33.4%) implement Person schema for attorney representation. Across these sites, we identified 2480 individual Person entities, averaging 14.9 per site. Of these, only 805 entities include the worksFor property linking the attorney to their law firm — the most basic form of entity relationship modeling. This pattern suggests that even firms employing Person schema rarely use it to construct meaningful entity networks.

Table 7. Property Coverage Within Person/Attorney Schema Entities.

Person Property	Coverage (%)
name	99.4%
image	92.5%
url	78.1%
@id	45.7%
sameAs	43.0%
telephone	41.3%
worksFor	32.5%
description	27.8%
alumniOf	19.8%
jobTitle	17.5%
knowsAbout	11.0%
award	7.5%

5. Discussion

5.1. The Schema Adoption Gap

Our findings reveal a significant structured data adoption gap in the PI legal services sector. While 67.6% of firms deploy some JSON-LD, the adoption of semantically rich, industry-specific schema types remains strikingly low. The 40.0% adoption rate for LegalService schema suggests that the vast majority of PI law firms are missing an opportunity to explicitly declare their practice areas, jurisdictions, and service attributes in a machine-readable format.

5.2. Proposed Structured Data Maturity Model

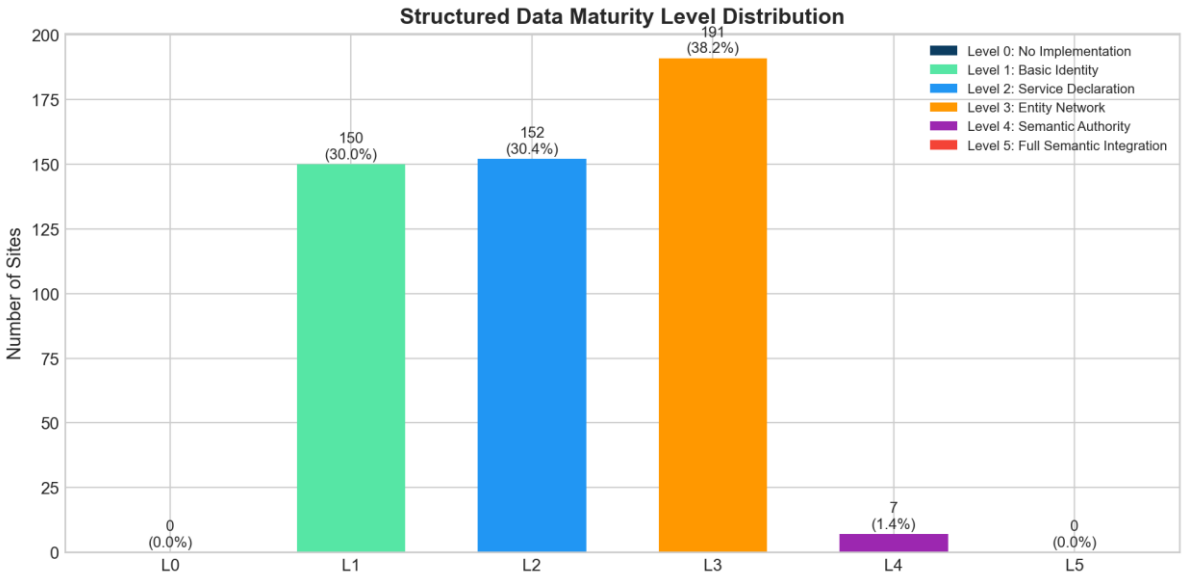

Figure 4. Structured Data Maturity Level distribution.

6. Implications

These findings offer actionable guidance for PI law firms and their digital marketing teams. First, firms should move beyond basic Organization schema toward LegalService and Attorney-specific implementations. Second, entity disambiguation through @id and sameAs properties should be treated as a priority, particularly as AI-driven search systems increasingly depend on entity resolution. Third, schema implementation should be approached as a site-wide strategy rather than a homepage-only effort.

7. Limitations and Future Research

This study has several limitations. Our sample of 500 websites, while substantial, was drawn from search results and may introduce selection bias toward firms with existing digital presence. Attorney count estimation relied on heuristics and may be inaccurate for some firms. We did not measure the causal relationship between schema implementation and organic search rankings. Future research should include longitudinal tracking of schema adoption and ranking changes, experimental measurement of LegalService schema impact on SERP features, and comparative analysis across legal service verticals.

8. Conclusion

This study provides the first systematic analysis of structured data adoption among personal injury law firm websites in North America. Analyzing 500 websites, we find that while 67.6% deploy some JSON-LD markup, only 40.0% use LegalService schema and fewer than 81.4% implement entity disambiguation via sameAs properties. The mean SCI of 11.8/25 indicates predominantly superficial implementations. The Structured Data Maturity Model proposed herein offers a framework for benchmarking and improving structured data strategy. As search engines and AI systems increasingly rely on structured data for entity understanding, the competitive implications of this adoption gap will intensify.

Abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
BERT	Bidirectional Encoder Representations from Transformers
DOI	Digital Object Identifier
E-E-A-T	Experience, Expertise, Authoritativeness, Trustworthiness
FAQ	Frequently Asked Questions
HTML	HyperText Markup Language
JSON-LD	JavaScript Object Notation for Linked Data
LLM	Large Language Model
ORCID	Open Researcher and Contributor ID
PI	Personal Injury
SCI	Schema Completeness Index
SEO	Search Engine Optimization
SERP	Search Engine Results Page
URL	Uniform Resource Locator

Author Contributions

Behzad Hussain: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., & Deshpande, A. (2024). GEO: Generative Engine Optimization. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24), 5654–5665. https://orcid.org/10.1145/3637528.3671900

[2] Barnard, J. (2022). The Fundamentals of Brand SERPs for Business. Kalicube Pro Publishing.

[3] Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. Scientific American, 284(5), 34–43. https://orcid.org/10.1038/scientificamerican0501-34

[4] Bizer, C., Meusel, R., & Primpeli, A. (2024). Web Data Commons: Extraction of Structured Data from the Common Crawl (JSON-LD, Microdata, RDFa Corpus, October 2024). University of Mannheim. https://webdatacommons.org/structureddata/

[5] Clio. (2023). Legal Trends Report 2023. Themis Solutions Inc. https://www.clio.com/resources/legal-trends/2023-report/

[6] Clio. (2024). Legal Trends Report 2024. Themis Solutions Inc. https://www.clio.com/resources/legal-trends/2024-report/

[7] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT 2019, 4171–4186. https://orcid.org/10.18653/v1/N19-1423

[8] Google Developers. (2024). Introduction to Structured Data Markup in Google Search. Google Search Central Documentation. https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data

[9] Google LLC. (2025). Search Quality Rater Guidelines (September 2025 edition). Google LLC. https://services.google.com/fh/files/misc/hsw-sqrg.pdf

[10] Guha, R. V., Brickley, D., & Macbeth, S. (2016). Schema.org: Evolution of Structured Data on the Web. Communications of the ACM, 59(2), 44–51. https://orcid.org/10.1145/2844544

[11] Gübür, K. T. (2023). Holistic SEO & Digital Marketing. Holistic SEO Digital. https://www.holisticseo.digital/

[12] HTTP Archive. (2024). Structured Data. The 2024 Web Almanac. https://almanac.httparchive.org/en/2024/structured-data

[13] Hu, X., Li, X., Chen, J., Li, Y., Wang, Y., Liu, Q., Wen, L., & Yu, P. S. (2024). Evaluating Robustness of Generative Search Engines on Adversarial Factual Questions. arXiv preprint arXiv: 403.12077. https://orcid.org/10.48550/arXiv.2403.12077

[14] Mika, P. (2015). On Schema.org and Why It Matters for the Web. IEEE Internet Computing, 19(4), 52–55. https://orcid.org/10.1109/MIC.2015.81

[15] Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2024). Unifying Large Language Models and Knowledge Graphs: A Roadmap. IEEE Transactions on Knowledge and Data Engineering, 36(7), 3580–3599. https://orcid.org/10.1109/TKDE.2024.3352100

[16] Sharma, N., Liao, Q. V., & Xiao, Z. (2024). Generative Echo Chamber? Effects of LLM-Powered Search Systems on Diverse Information Seeking. Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24), 1–17. https://orcid.org/10.1145/3613904.3642459

[17] Singhal, A. (2012). Introducing the Knowledge Graph: Things, Not Strings. Google Official Blog. https://blog.google/products/search/introducing-knowledge-graph-things-not/

Biography

Behzad Hussain is an Independent Researcher and Certified Semantic SEO Expert, serving as CEO and Founder of Rank Brilliance, a Semantic SEO agency based in Lahore, Pakistan. His research investigates the intersection of entity-based search optimization, structured data adoption, and legal services discoverability, with a particular focus on personal injury law firm websites across North American markets. Trained in Chartered Accountancy under the Institute of Chartered Accountants of Pakistan (ICAP), Hussain brings quantitative rigor to applied SEO research, bridging practitioner expertise with empirical methodology. He holds certification in Semantic SEO under the framework developed by Koray Tugberk Gübür and has led structured data implementations for legal, healthcare, SaaS, and FinTech clients across the United States and Canada.