Bulletin of the Transilvania University of Braşov
Series V: Economic Sciences • Vol. 19(68) No. 1 - 2026
https://doi.org/10.31926/but.es.2026.19.68.1.9

FROM SEO TO GEO: A BIBLIOMETRIC ANALYSIS OF SEARCH ENGINE OPTIMIZATION AND GENERATIVE ENGINE OPTIMIZATION

Anca POPESCU[1]

Abstract: *Generative AI-based search platforms are redefining information discovery, challenging the dominance of traditional search engine optimization (SEO) and giving rise to the concept of Generative Engine Optimization (GEO). This research aims to provide a bibliometric mapping of the SEO–GEO intersection using the PRISMA 2020 methodology. Three thematic clusters and a fragmented co-authorship network were identified. A general model of the SEO-to-GEO transition was outlined, integrating the shift from ranking-based to citation-based visibility and the need to rethink content optimization for AI readability. In addition, these results can serve as a reference for researchers and digital marketing practitioners.*

Keywords: *search engine optimization, artificial intelligence, ChatGPT, large language models, generative artificial intelligence, generative engine optimization, digital marketing.*

1. Introduction

For around 30 years, SEO has been the primary way to improve online visibility. It affects how users find, rank, and access information (Killoran, 2013). The introduction of the PageRank algorithm by Brin and Page (1998) marked a significant shift in how search engines operate. Prior to that, ranking relied heavily on matching search terms to page content - a relatively simple mechanism. What PageRank brought to the table was an entirely different logic: it used the link structure between pages as a proxy for relevance and authority. This laid the groundwork for the practices we now associate with SEO, which grew over time to include keyword optimization, link acquisition, technical improvements to site architecture, and the adoption of structured data (Granka, 2010).

In recent years, however, this established framework has come under increasing pressure from generative AI-powered search platforms. The release of ChatGPT by OpenAI in November 2022 proved to be a pivotal moment - not so much because of the technology itself, but because of the speed at which it was adopted. The platform attracted 100 million users in just two months and, by mid-2025, it was processing more

[1] Transilvania University of Braşov, anca.radbata@unitbv.ro, ORCID ID 0000-0002-5247-9581

than 2.5 billion prompts per day. Shortly after, in February 2023, Microsoft integrated GPT-4 into Bing Chat, later rebranded as Copilot. Google followed in May 2023 with its Search Generative Experience, which it eventually restructured into AI Overviews. By early 2026, AI-generated responses appeared in nearly half of all monitored queries, reflecting a 58% year-over-year increase (BrightEdge, 2026). Perplexity AI, also launched in 2023, took a distinct approach by delivering citation-rich, synthesised answers from the outset. What these platforms share is a fundamental departure from the traditional model: rather than directing users to a ranked list of links, they construct answers from multiple sources, thereby reducing the need to visit individual pages (Aggarwal *et al.*, 2024).

Available empirical data also support the changes these technologies produce. According to BrightEdge (2025), sessions from AI-based sources increased by 527% year-on-year in the first half of 2025, even though, overall, they still represent less than 1% of total referral traffic. At the same time, Gartner (2024) estimates a 25% reduction in the volume of traditional searches between 2026 and 2028. Meanwhile, a study by SparkToro and Datos (2024) on zero-click searches shows that, out of 1,000 searches performed on Google in the United States, only 374 result in clicks to pages on the open web. For queries that include AI Overviews, the zero-click rate reaches approximately 83% (Barenholtz, 2025). Taken together, these data suggest that this is not just a change in tools, but a bigger change in how information is searched for and consumed online. In this context, a new optimization concept is emerging: Generative Engine Optimization (GEO).

GEO is a concept introduced by Aggarwal *et al*. (2024) and represents a clear departure from the classic SEO model. In SEO, the goal is to get the best ranking in search engine results. In GEO, the focus shifts to how visible the content is in answers generated by artificial intelligence and how likely it is for a source to be referenced or mentioned. The difference goes beyond terminology - it extends to the underlying logic. SEO was built around already well-known elements, such as keywords, backlinks, and domain authority. GEO, on the other hand, seems to depend more on the clarity of the information, the authority of the entity, the effectiveness of the content, and how it can be read and interpreted by automated systems, including through the use of schema markup (Aggarwal *et al*., 2024). At the same time, we should note that this field does not yet have fully stabilised terminology. In recent literature, several formulas are used almost interchangeably, such as GEO, Answer Engine Optimization (AEO), Large Language Model Optimization (LLMO), AI Optimization (AIO), or Generative Search Optimization (GSO).

Even though interest in this topic is growing in both academia and the professional world, the available literature offers limited bibliometric insights into the transition from SEO to GEO. To date, most contributions have focused either on concrete optimization strategies (Aggarwal *et al.*, 2024), on the analysis of specific platforms or contexts (Cutler, 2023), or on the broader relationship between artificial intelligence and marketing (Tafesse and Wien, 2024). What is missing, however, is precisely an overview of this emerging field. Such a perspective can be provided by bibliometric methods, which allow for tracking publication dynamics, identifying authors or works with greater

influence, and outlining the main research themes with the help of tools such as keyword co-occurrence analysis, co-authorship networks, or bibliographic linkage (Donthu *et al*., 2021; Zupic and Čater, 2015).

Moreover, the methodological literature shows that bibliometric analyses are not reserved only for already consolidated fields. They can also be applied to smaller corpora when the research targets an emerging field, at an early stage of development, as long as the source selection and analytical steps are carried out rigorously (Donthu *et al*., 2021; Aria and Cuccurullo, 2017).

The present study aims to present four main contributions. First, it proposes a bibliometric mapping of the intersection between SEO and GEO in the academic literature, with the aim of outlining the intellectual landscape of this emerging field more clearly. Second, it aims to identify the main thematic clusters and how they have evolved over time, using specific science mapping techniques. A third contribution aims to highlight the current fragmentation of the field and to capture an early trend of terminological convergence around the concept of GEO. Finally, the study aims to formulate a research agenda for the coming period, starting from the empirical gaps identified in the analysed literature.

Starting from this, the research is guided by four main questions:

RQ1: What are the publication patterns and growth dynamics of the SEO–GEO field (2023–2026)?

RQ2: Who are the most influential authors, institutions, and countries?

RQ3: What are the thematic clusters and intellectual structure of the field?

RQ4: What are the research gaps and emerging future directions?

The following sections address these questions through a theoretical framing, a detailed methodology, and a combined presentation of results and discussion.

2. Theoretical background

The evolution of web search can be traced in four technological stages, and this division is useful for better understanding the thematic clusters discussed in Section 4. The first stage, corresponding to the period 1994–1997, was marked by the hand-made web directories. The second began in 1998, with the advent of the PageRank algorithm proposed by Brin and Page (1998), a moment that clearly changed the functioning of search engines. From that point on, Google gradually established itself as the dominant search engine, and SEO began to take shape more clearly as a distinct practice and then as a professional field in its own right (Killoran, 2013). The third stage, spanning from 2013 to 2022, is linked to the development of semantic search, supported by updates and models such as Hummingbird, RankBrain, and BERT (Devlin *et al*., 2019). In the same period, formats such as Featured Snippets and Knowledge Panels have become more prevalent, and zero-click searches have become more common. Starting in 2023, a new stage is emerging, associated with generative engines built around Retrieval-Augmented Generation (RAG) mechanisms that combine the internal knowledge of linguistic models with real-time document retrieval.

The transition from SEO to GEO also implies a fundamental shift in how digital visibility

is measured and evaluated. In SEO, indicators like click-through rate (CTR), organic traffic, and search result positions are important. However, when it comes to GEO, we focus on different factors. We look at how often a source is cited, how well it appears in answers, and its overall influence on the information shared. Aggarwal *et al*. (2024) tried to capture this difference by including some specific metrics, such as Position-Adjusted Word Count and Subjective Impression. Their results show that classic SEO techniques no longer have the same effect when we talk about generative engines. Instead, the inclusion of relevant statistics and citations seems to improve visibility by approximately 30–40%. In this context, we can observe the emergence of another logic of online visibility, close to what some authors call a "citation economy". In other words, the value of content no longer depends only on the number of users who actually enter a site, but increasingly on the extent to which that content is retrieved and cited by artificial intelligence systems - extending the logic of academic citation analysis into the digital ecosystem (Aggarwal *et al.*, 2024). Industry data goes in the same direction. Queries that activate AI Overviews have higher zero-click rates than traditional search queries (Barenholtz, 2025), suggesting that AI Overviews disrupt the classic SEO funnel at its entry point.

3. Methodology

The present study is based on quantitative bibliometric research that combines performance analysis with science mapping techniques, in line with the recommendations of Donthu et al. (2021) and Zupic and Čater (2015). More specifically, performance analysis was used to capture publication trends and the field's main actors. At the same time, the science mapping component focused on co-authorship relationships, keyword co-occurrence, and bibliographic coupling. Even though the corpus size is relatively small (n = 32), this is in line with the practice encountered in bibliometric research applied to fields still in an early stage of development, where the rigour of the selection process matters more than the very large volume of documents included (Aria and Cuccurullo, 2017).

The Scopus database was chosen as the main data source following a comparative evaluation of both Scopus and Web of Science (WoS). Scopus provides broader coverage of conference proceedings and emerging interdisciplinary fields, particularly in Computer Science (Mongeon and Paul-Hus, 2016), and offers superior phrase-level matching - both essential given the conference-dominated nature of the SEO–GEO literature (65.6% of the final corpus).

The Web of Science query TS = (generative engine optimization) returned 138 results. However, because the Topic Search field does not support phrase-level matching, the three terms were treated as independent tokens, yielding mostly unrelated studies on generative design or heat-engine optimization. Only 3 results (2.2%) proved relevant. By contrast, Scopus yielded 51 results through exact phrase matching, of which 32 were ultimately included (62.7% post-screening precision), justifying its exclusive use.

The bibliographic search was performed in Scopus, through the TITLE-ABS-KEY field, using the following query:

TITLE-ABS-KEY ("search engine optimization" AND ("generative AI" OR "large language model" OR "ChatGPT" OR "generative engine" OR "AI search" OR "Perplexity" OR "AI Overviews")) OR TITLE-ABS-KEY ("generative engine optimization" OR "answer engine optimization")

The query was constructed to include both papers that address SEO in relation to generative artificial intelligence and those that explicitly discuss concepts such as GEO or AEO. The search was conducted on March 22, 2026. In its initial form, it generated 51 documents. Data collection was performed in a single extraction phase, with no subsequent updates to the corpus. After applying filters for publication type (article or conference paper) and the analysed time interval (2023-2026), 40 documents remained. The PRISMA 2020 protocol (Page *et al*., 2021) was used for the final selection, applied in four stages, as illustrated in Figure 1. The first stage was the identification stage, in which 51 papers were recorded. This was followed by automatic screening, after which 40 documents remained. In the manual screening stage, 8 papers were excluded as irrelevant or only tangentially related to the research topic, leaving a final sample of 32 documents.

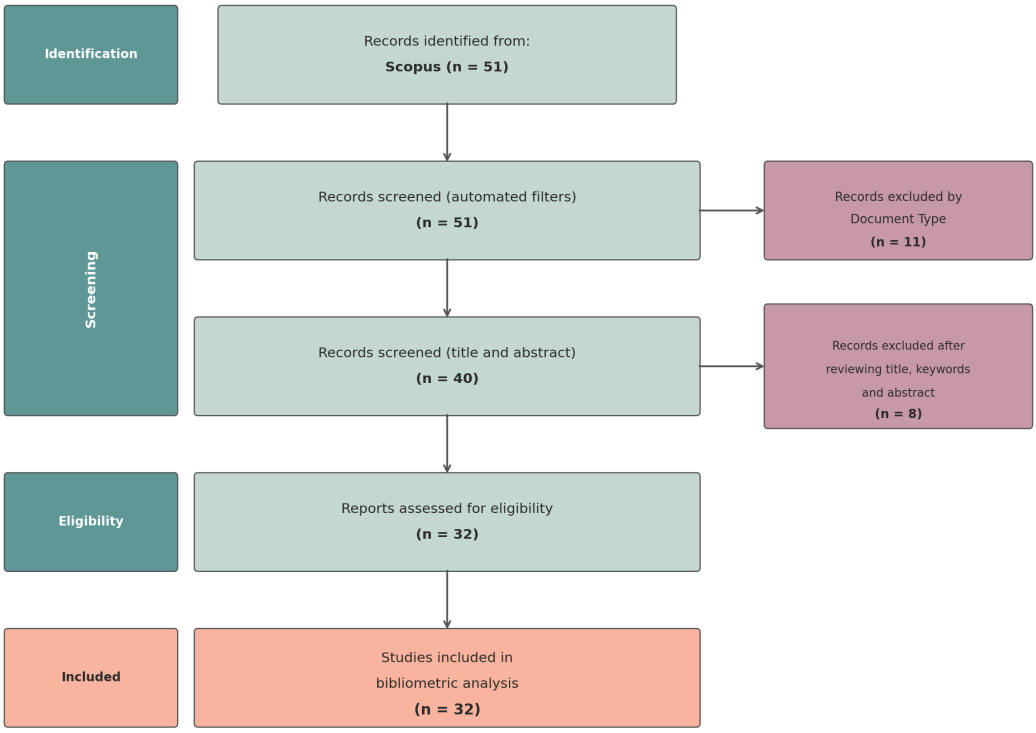


Fig. 1. *Document selection process according to the PRISMA 2020 protocol*

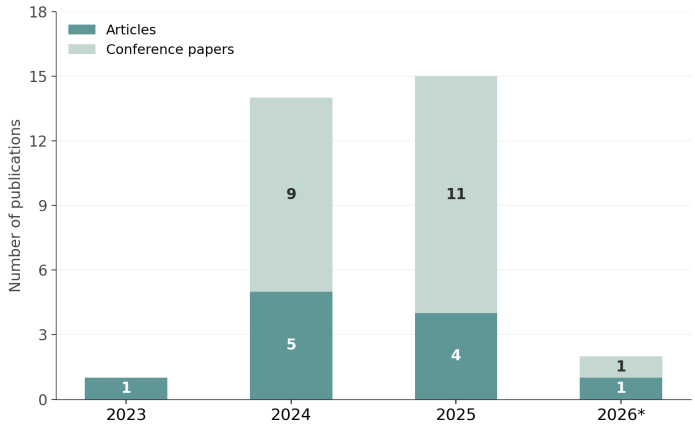
The analysis included articles and conference papers (2023–2026) that explicitly address SEO in the context of generative AI or content visibility in generative engines. Studies on traditional SEO without an AI component, unrelated ML techniques, and

editorials without original contributions were excluded. The selected data were exported to CSV format and then imported into VOSviewer 1.6.20, a widely used software tool for constructing and visualising bibliometric networks. VOSviewer was selected for this study due to its established track record in bibliometric research and its capacity to generate co-occurrence, co-authorship, and bibliographic coupling maps from standard bibliographic data formats.

4. Results and Discussions

Regarding the dynamics of publications (RQ1), the distribution of documents clearly shows that the field is still in an early stage of development. Of the 32 documents included in the analysis, only one was published in 2023. The following year, however, their number increased significantly to 14, and in 2025 it remained at a similar level, with 15 papers (Figure 2). Therefore, a rapid increase in interest in this topic is observed, followed by relative stabilisation.

Another relevant aspect is the high share of conference papers. Of the 32 documents, 21 are conference papers, accounting for 65.6% of the corpus. Such a pattern is common in emerging fields, where ideas first circulate through conferences, and later, as the field consolidates, more journal articles begin to appear.



**2026 data are partial (collected March 2026).*

Fig. 2. *Publication distribution by year*

The increase from 1 to 14 publications in a single year coincides with the moment when ChatGPT and Google AI Overviews entered the public consciousness. The most cited article remains Cutler (2023), with 23 citations, followed by Tafesse and Wien (2024) with 20. The founding article of GEO (Aggarwal *et al*., 2024) garnered 8 citations, despite being published at a computer science conference - a sign that its impact transcends disciplinary boundaries.

In terms of author keywords, "search engine optimization" leads by a wide margin (16 occurrences), followed by “artificial intelligence” (9), “ChatGPT” (8), "large language models" and "generative AI" (7 each). The term "generative engine optimization"

appears only twice - most authors still place their contributions under the umbrella of AI-enhanced SEO, the terminology not yet crystallised.

Moving to the co-authorship network (RQ2), the map reveals an extremely dispersed landscape: 82 authors are distributed in 27 clusters, with a cluster/author ratio of approximately 1:3. In bibliometric studies in more established fields, this ratio is usually below 1:10 (Donthu *et al*., 2021). The only connected component of notable size comprises 6 authors - the founding GEO team (Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan, Deshpande, affiliated with Princeton and IIT Delhi). Otherwise, the groups work in isolation, without bridges between them.

This fragmentation deserves careful interpretation. On the one hand, with only 32 documents, the chances of observing dense networks are inherently lower. On the other hand, the dispersion is consistent with a pre-paradigmatic field in which researchers from different disciplines (computer science, marketing, communication sciences) address the same problem without communicating with each other. The two explanations are not mutually exclusive: the small size of the corpus probably amplifies a real and structural fragmentation. A country-level co-authorship analysis was also conducted to explore international collaboration patterns, which reveals a similarly dispersed picture: 18 countries form 13 clusters, with only 6 links (total link strength: 7). The United States leads with 6 documents and 34 citations, followed by Poland (5 documents) and India (4 documents), while a single publication represents 10 countries. This mirrors the fragmentation observed at the individual author level, confirming the absence of established international research networks in this field.

Looking at the thematic clusters (RQ3), the keyword co-occurrence analysis reveals three distinct groups, each with its own logic (Figure 3).

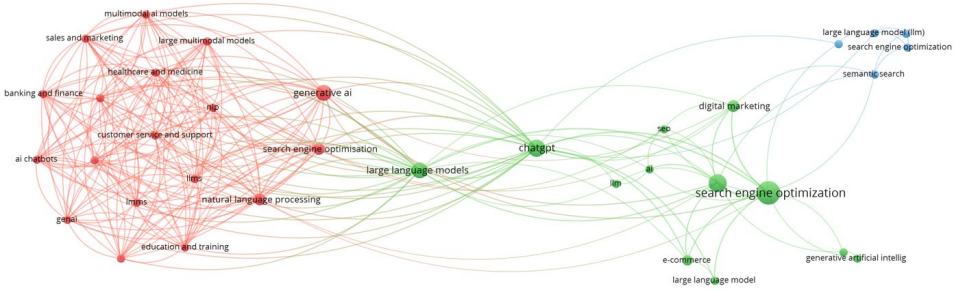


Fig. 3. *Keyword co-occurrence network visualisation*

Cluster 1 (Red) - Applications of AI in vertical domains. It revolves around the terms “generative AI”, "NLP", and "GenAI" and brings together studies on the use of AI chatbots in healthcare, banking, education, journalism, and tourism. It is essentially the exploratory phase - the area where the academic community is trying to understand what generative AI can do.

Cluster 2 (Green) - SEO transformation through AI/LLM. Dominated by the largest node (“search engine optimization”), it connects ChatGPT, LLM, digital marketing, e-commerce, and semantic search. This is where the core of the SEO - GEO transition lies - the place where traditional optimization begins to merge with generative technologies.

Cluster 3 (Blue) - Technical SEO and structured data. More narrowly, it is organised around schema.org. Basically, it deals with the “how?” - the technical implementation through which the content becomes readable for linguistic models.

Viewed through the lens of the overlay visualisation, the three clusters fall on a clear time axis: the red cluster emerged earlier (2023–2024), while the green cluster gained consistency in 2024-2025. The field has shifted from “what can generative AI do?” to “how do we adapt our visibility strategies?” - which is precisely the transition from SEO to GEO. The bridging role of the term "ChatGPT" between clusters 1 and 2 confirms that this technology was the catalyst for the transition. What is missing from this thematic structure is equally significant. No cluster has emerged dedicated to ethics or algorithmic transparency, although the manipulation of visibility in generative engines raises questions as pressing as those in adversarial SEO (Gyöngyi and Garcia-Molina, 2005). A cluster focused on the user perspective is also missing. These absences are not surprising in such a young field, but they mark priority research directions.

The bibliographic coupling analysis reinforces the same picture: 19 clusters out of 32 papers, with only 14 links. The strongest coupling was found between two companion papers from the same research group (link strength: 67). Co-citation analysis was also attempted but produced no significant clusters at a threshold of 2 citations, indicating that authors do not yet draw on a shared set of foundational works. Each research group brings its own disciplinary references, and a common bibliographic base has not yet formed.

Taken together, these indicators - conference dominance (65.6%), co-authorship fragmentation (27/82), absence of a dedicated journal, terminological inconsistency, and failure of co-citation - point to a pre-paradigmatic stage. Yet the increase from 1 to 15 publications per year suggests the field is approaching an inflection point, with convergence towards GEO as an anchor term still an ongoing process.

Concerning the research gaps and future agenda (RQ4), Figure 4 summarises the identified gaps and the corresponding directions.

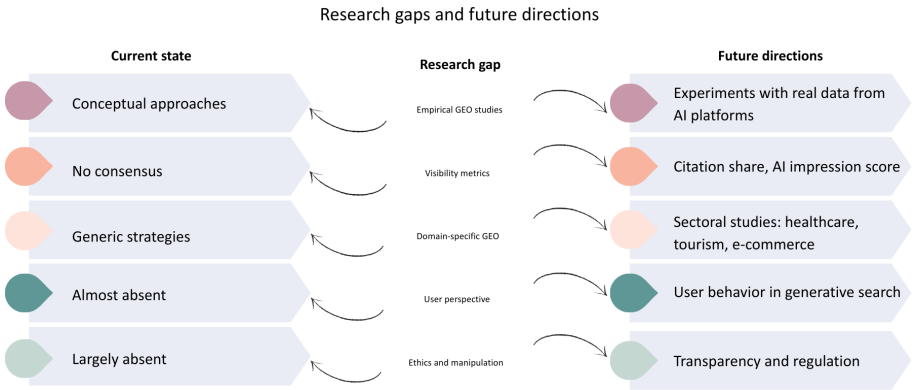


Fig. 4. *Research gap and future directions*

Beyond the gaps identified above, it is worth highlighting a cross-cutting gap: the near absence of empirical data from AI platforms. Most studies use conceptual frameworks

rather than traffic or citation data from real-world uses of generative engines. This gap is partly explained by limited access to proprietary platform data, but it remains the main obstacle to the scientific maturation of GEO.

From a practical standpoint, digital marketing practitioners should prioritise strategies with emerging empirical support, such as schema markup, entity authority, ChatGPT-optimized content, and structured Q&A formats. For researchers, these gaps represent opportunities, with studies based on real platform data being the most pressing need.

5. Conclusions

Overall, the study shows that the transition from SEO to GEO is not a simple terminological extension, but a sign of a broader shift in how digital visibility is understood. Where the dominant logic was positioning in search results and attracting organic traffic, generative engines shift attention to citation, source selection, and the ability of content to be integrated into synthesised answers. The bibliometric analysis suggests that the field is still in an incipient phase, marked by rapid growth, conceptual fragmentation, and a poorly defined collaboration structure. It is precisely this early stage that makes it relevant for research, as it allows us to capture how new themes, actors, and vocabularies associated with optimization for generative search are forming.

At the same time, the available literature is not yet sufficiently consolidated to provide stable analysis models or unanimously accepted metrics. The development of the GEO field will depend not only on the proliferation of publications but also on the maturation of the theoretical and methodological tools for understanding this phenomenon, with the emphasis likely shifting from defining concepts to empirically validating strategies and building clearer analytical benchmarks. GEO appears today as a field with high potential for both academic research and digital marketing practice. As generative engines continue to reshape the relationship between users, content, and information sources, understanding these processes will become increasingly important.

References

Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., Deshpande, A. (2024). GEO: Generative Engine Optimization. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*, pp. 5–16. https://doi.org/10.1145/3637528.3671900

Aria, M., Cuccurullo, C. (2017). "Bibliometrix: An R-tool for comprehensive science mapping analysis." *Journal of Informetrics*, 11(4), pp. 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Barenholtz, L. (2025). "Zero-Click Searches and How They Impact Traffic." *Similarweb Blog*. Available at: https://www.similarweb.com/blog/marketing/seo/zero-click-searches/, [Accessed on 6 March 2026].

BrightEdge. (2025). "AI search visits surging in 2025." *BrightEdge Research Reports*. Available at: https://www.brightedge.com/resources/research-reports/ai-search-visits-in-surging-2025 , [Accessed on 20 March 2026].

BrightEdge. (2026). "AI Overviews at the One-Year Mark: Presence, Size, and What They're Citing." *BrightEdge Weekly AI and Search Insights*. Available at: https://www.brightedge.com/resources/weekly-ai-search-insights/ai-overviews-one-year-presence-size-citing, [Accessed on 20 March 2026].

Brin, S., Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1–7), pp. 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Cutler, K., 2023. ChatGPT and search engine optimisation: The future is here. *Applied Marketing Analytics*, 9(1), pp. 8–22. https://doi.org/10.69554/EQBM2625

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*, pp. 4171–4186. https://doi.org/10.18653/v1/N19-1423

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M. (2021). "How to conduct a bibliometric analysis: An overview and guidelines." *Journal of Business Research*, 133, pp. 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Gartner. (2024). Gartner predicts search engine volume will drop 25% by 2026, due to AI chatbots and other virtual agents [Press release]. Available at: https://www.gartner.com/en/newsroom/press-releases/2024-02-19-gartner-predicts-search-engine-volume-will-drop-25-percent-by-2026-due-to-ai-chatbots-and-other-virtual-agents, [Accessed on 10 March 2026].

Granka, L.A., 2010. The politics of search: A decade retrospective. *The Information Society*, 26(5), pp. 364–374.

Gyöngyi, Z., Garcia-Molina, H. (2005). "Web spam taxonomy." In *Proceedings of the First International Workshop on Adversarial Information Retrieval on the Web (AIRWeb 2005)*, Chiba, Japan, May 2005, (pp. 39–47).

Killoran, J.B. (2013). "How to use search engine optimization techniques to increase website visibility." *IEEE Transactions on Professional Communication*, 56(1), pp. 50–66.

Mongeon, P., Paul-Hus, A. (2016). "The journal coverage of Web of Science and Scopus: a comparative analysis." *Scientometrics*, 106, pp. 213–228. https://doi.org/10.1007/s11192-015-1765-5

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., *et al*. (2021). "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews." *BMJ*, 372(71). https://doi.org/10.1136/bmj.n71

SparkToro and Datos. (2024)." 2024 Zero-Click Search Study: For every 1,000 US Google Searches, only 374 clicks go to the Open Web. In the EU, it's 360." *SparkToro Blog.* Available at: https://sparktoro.com/blog/2024-zero-click-search-study-for-every-1000-us-google-searches-only-374-clicks-go-to-the-open-web-in-the-eu-its-360/, [Accessed on 5 March 2026].

Tafesse, W., Wien, A. (2024). "ChatGPT's applications in marketing: a topic modeling approach." *Marketing Intelligence & Planning*, 42(4), pp. 666–683. https://doi.org/10.1108/MIP-10-2023-0526

Zupic, I., Čater, T. (2015). "Bibliometric methods in management and organization." *Organizational Research Methods*, 18(3), pp. 429–472. https://doi.org/10.1177/1094428114562629