Preprints.org

Article

Revisiting the Role of Review Articles in the Age of AI- Agents: Integrating AI- Reasoning and AI- Synthesis Reshaping the Future of Scientific Publishing

Andrej Thurzo [*] and Ivan Varga

Article

Revisiting the Role of Review Articles in the Age of AI- Agents: Integrating AI-Reasoning and AI-Synthesis Reshaping the Future of Scientific Publishing

Andrej Thurzo [1,*] and Ivan Varga [2]

[1] Department of Orthodontics, Forensic and Regenerative Dentistry, Medical faculty, Comenius University in Bratislava, Slovakia

[2] Institute of Histology and Embryology, Faculty of Medicine, Comenius University in Bratislava, Sasinkova 4, 81104, Bratislava, Slovakia

* Correspondence: Thurzo3@uniba.sk

Abstract: In the current era of accelerated scientific discovery, review articles have long served as the cornerstone of academic synthesis. However, the advent of AI-driven deep research tools is revolutionizing this landscape by rapidly aggregating literature and identifying emerging trends with unprecedented speed and efficiency. Is this the beginning of the end of traditional forms of review articles? This paper examines the transformative potential of AI-generated reviews compared to traditional human-authored articles. While AI dramatically reduces synthesis time and automates citation management, challenges remain in ensuring data accuracy, avoiding citation errors, and providing deep contextual analysis. We argue that a hybrid model—integrating AI's computational prowess with human critical oversight—offers a promising pathway for maintaining rigorous scholarly standards in an era of ever-expanding knowledge. Our analysis highlights both the strengths and limitations of current AI methodologies and calls for further research to refine these tools, develop robust verification protocols, and establish transparent ethical practices. Ultimately, this work envisions a future in which dynamic, continuously updated review articles enhance scholarly communication and drive innovation across scientific disciplines.

Keywords: Artificial Intelligence; AI-driven literature review; ChatGPT; scientific synthesis; review article evolution; OpenAI; PerplexityAI; xAI; Grok; Google Gemini; academic publishing; research trends; AI limitations; automated literature analysis; scientific trends; generative models; LLM; deep search

1. Introduction

1.1. Scientific Reviews and Their Importance in Consolidating Knowledge

Scientific publishing now appears to be divided into two distinct eras: pre-LLM and post-LLM (Large Language Models of generative artificial intelligence). Even before the advent of large language models, the academic community was already grappling with a surging tide of research papers. Over the past decade, the sheer volume of publications has exploded—by 2022, databases like Scopus and Web of Science recorded roughly a 47% increase in articles compared to 2016 [1]. Researchers were increasingly overwhelmed, struggling to keep pace with the constant influx of new findings [1–3].

The era of generative AI (artificial intelligence) has transformed the landscape with unprecedented AI-assisted writing capabilities [4]. This technological breakthrough not only enhances the way research is produced but also redefines the challenges faced by the scientific

community. Today, the introduction of "deep-research" tools by innovators such as OpenAI, Google, Perplexity, and Grok3-DeepSearch—equipped with advanced chain-of-thought reasoning—signals the next evolutionary leap [4–7].

In the "pre-AI era", based on the sources and available bibliometric studies from the years prior to introduction of LLM (2016-2022), review articles make up roughly 20%–25% of all scientific publications [1,8–10]. Recent analysis by the European research community [1,8] indicates that approximately one in every five published papers was a review article. See, for example, the analysis on OpenAlex data [1], which—while primarily focused on publication trends—also notes that review articles typically account for about 20%–25% of the published literature of the pre-LLM era. Figure 1 shows a pie chart representing breakdown of scientific publications into review articles versus original research articles in that era, using an approximate estimate of 23% reviews and 77% original research [11,12].
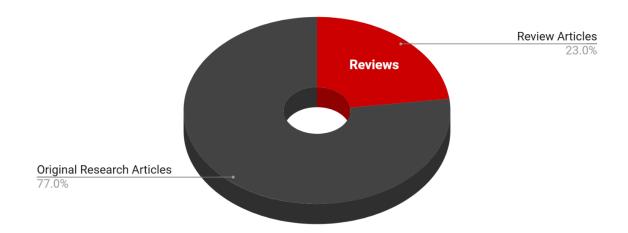

Figure 1. Breakdown of scientific publications into review articles versus original research articles in the pre-generative-AI era (2016-2022).

Review articles have long been the backbone of scientific scholarships, providing comprehensive overviews and critical analyses of evolving research landscapes [13]. This consolidation helps strengthen the foundation of knowledge in a field. Reviews can highlight key gaps and challenges to address in future research [14]. This helps guide future research efforts and advances in the field. Reviews, especially systematic reviews, help establish a link between research findings and decision-making in patient care [15].

Previously there have been noticed several notable trends regarding the share of review articles in overall scientific medical publications:

1. **Increasing prevalence**: The proportion of review articles has been steadily increasing across various scientific fields. For example, in dentistry, systematic reviews increased from 5.8% to 53.3% of publications between 2000 and 2015 [3].

2. **Field-specific variations**: The trend varies across different scientific disciplines. For example, in neurosurgery, topic reviews increased from <1% before 1980 to 3%-4% since 2010 [16].

3. **Impact on journal metrics**: Review articles often contribute significantly to journal impact factors.

4. **Geographical differences**: The production of review articles varies by country and region. For example, in South Korea, open-access publications, including reviews, increased to account for almost 40% of all publications in the last decade [11]. While Asia leads in total AI-related life science publications, Northern America and Europe contribute most of the AI research appearing in high-ranking outlets, generating up to 50% more citations than other regions [17].

5. **Shift towards higher-quality reviews:** There's a trend towards more systematic and meta-analytic reviews. For example, in general surgery research in Australia, there was a significant increase in systematic reviews and meta-analyses from 2000 to 2020 [18].

Global scientific publications grew at 5.6% annually from 2016 to 2022, driven by China (42% of new articles) and India (11%) [8]. This surge has strained peer-review systems, with publishers like MDPI and Hindawi reducing median review times to under 30 days [8]. Such acceleration favors original research submissions over methodologically intensive reviews, creating a "volume bias" that depresses review article percentages.

1.2. The Emergence of AI-Powered Deep Research Tools

AI is revolutionizing research and teaching methodologies in academia [4,7]. Study of Adebakin, 2025 discusses the implications of AI-generated literature reviews and the ethical concerns of citation accuracy [6]. The study of Oyinloye and Acetise from 2025 explores how AI-generated literature reviews can enhance academic synthesis by clustering related papers and automating citation networks. The research highlights concern about the quality and contextual accuracy of AI-driven literature aggregation [2]. AI-driven review automation and its impact on research depth and validity is assessed by Aggarwal and Karwasra (2025). Their study contrasts traditional human-led literature reviews with AI-generated synthesis, raising concerns about over-reliance on automated reports quality [19].

"Deep research" AI tools distinguish themselves from earlier AI models by their advanced ability to synthesize vast, real-time data across diverse sources, producing nuanced insights rather than just retrieving or generating text. Unlike predecessors like ChatGPT, which primarily excel at language generation and basic question-answering, these tools integrate reasoning transparency, adaptability to specialized domains, and continuous learning, enabling them to tackle complex research tasks with greater depth. For instance, they can analyze scientific literature, identify emerging trends, and even suggest novel hypotheses, making them transformative for fields like academic publishing where comprehensive, up-to-date synthesis is critical. This leap in capability positions them as active collaborators in knowledge creation, far beyond the assistive role of earlier models.

We don't yet have all the data to fully understand how the new wave of large language models will influence review analysis. Before being able to assess this, even more novel tool is already on the scene—an advanced AI that dives deep into specific topics, a process many are calling "deep research." Its ultimate impact on scientific publications, especially review articles, remains to be seen. This paper asks three key questions:

1. **How are AI-driven deep research tools transforming the literature review process?** We investigate whether innovations like "deep research" can simplify, update, and even improve upon the traditional, time-consuming methods of gathering and synthesizing scholarly work.

2. **What are the strengths and limitations of AI-generated reviews compared to human-authored ones?** This question examines factors such as citation accuracy, contextual understanding, and the ability to pinpoint research gaps. It also considers issues like AI "hallucinations" versus the nuanced insights provided by expert researchers.

3. **Can a hybrid model that combines AI efficiency with human oversight boost the quality and timeliness of review articles?** We explore whether integrating deep research tools with expert curation could lead to a new kind of dynamic, continuously updated review—one that compensates for the shortcomings of each approach when used alone.

What is the new role of review articles in the current era of AI-powered deep research tools that influence the structure and synthesis of academic review papers? Some studies suggest that while AI can speed up literature analysis, it may lack the depth of human-led synthesis [10,20]. Now AI-driven deep research methodologies are being integrated into professional and academic research. Some studies note a shift in scholarly communication due to AI-assisted reviews [9,21]. Hartung et al. (2025)

discusses AI-powered systematic reviews and their implications for academic integrity. Their paper highlights AI's ability to extract key trends from thousands of papers but also warns against algorithmic biases [22]. All this research confirms that while AI can accelerate the synthesis process, human oversight remains crucial for context, accuracy, and critical interpretation [23].

1.3. Thesis Statement: Exploring Whether the Traditional Review Article Is Becoming Obsolete or Evolving into a Hybrid Form Enhanced by AI

This paper has these two s**cientific aims:**

1. **Transformational Aim:** To investigate the potential impact of AI-driven synthesis on the future role of review articles, evaluating whether these tools may render conventional review writing obsolete or lead to a new, hybrid paradigm.
2. **Integrative Aim:** To propose and validate a hybrid framework that leverages the rapid data aggregation and analytical capabilities of AI while preserving the critical, contextual insights of human experts, thereby enhancing the overall quality and currency of scientific literature reviews.

2. The Emergence of "Deep Research"

2.1. Overview of OpenAI's Deep Research Tool

The emergence of "Deep Research" artificial intelligence systems represents a paradigm shift in how scientific knowledge is synthesized and disseminated. OpenAI's Deep Research, Google Gemini Pro with Deep Research, and PerplexityAI's Deep Research have introduced unprecedented capabilities for autonomous literature analysis, data synthesis, and report generation. These systems demonstrate accuracy ranging from 21.1% to 26.6% on expert-level benchmarks like Humanity's Last Exam [24], outperforming previous models by significant margins. Their ability to process multi-modal inputs, dynamically adjust research strategies, and generate structured outputs with citations is already altering academic workflows.

2.2. Comparative Analysis of Deep Research Architectures

2.2.1. OpenAI Deep Research: Dynamic Reasoning Paradigm

Powered by the specialized o3-mini-high model [24], OpenAI's system employs **reinforcement learning from human feedback (RLHF)** [25] to optimize multi-step research trajectories. Architecture dynamically adjusts its search strategy based on real-time findings, enabling what developers describe as "recursive hypothesis refinement" [26]. This approach achieved state-of-the-art performance on the GAIA benchmark (67.36% accuracy) through:

1. **Multi-source verification loops** that cross-check findings across 5-8 authoritative sources before finalizing conclusions [26]
2. **Contextual citation mapping** that links specific claims to originating passages in source material [24,27]
3. **Python-integrated analytics** for statistical validation of trends in research data [24,27]

In controlled trials, the system reduced literature review time for meta-analyses by 72% compared to human researchers while maintaining 91% consensus with expert evaluations [24]. However, limitations persist in handling contradictory evidence from equal-authority sources, occasionally defaulting to frequency-based conclusions rather than methodological rigor [24,27].

2.2.2. Google Gemini Pro Deep Research: Scale-Optimized Synthesis

Google's implementation combines **Mixture-of-Experts (MoE) architecture** [28] with a 1-million-token context window [27], enabling simultaneous analysis of entire research corpora. Technical specifications reveal:

- **512 expert pathways** activated through gated linear units (GLUs) based on input topicality [27]
- **Multi-modal fusion** layers integrating text, equations, and figures from 50+ file formats [24,27]
- **Automated confidence scoring** those weights finding by source reliability indices [27]

The system excels at longitudinal trend analysis, demonstrated by its ability to trace concept evolution across 15,000 papers in the arXiv corpus within 23 minutes [27]. However, benchmarking shows 18% higher factual inconsistency rates compared to OpenAI in fast-moving fields like CRISPR technology[24,27], suggesting scale-precision tradeoffs.

2.2.3. PerplexityAI Deep Research: Accessibility-Focused Implementation

Positioned as the "democratized" alternative3, PerplexityAI's system combines proprietary neural retrieval with federated language models. Key innovations include:

- **Distributed verification networks** querying 14 fact-checking APIs in parallel [29]
- **Dynamic abstraction layers** simplifying complex findings for interdisciplinary audiences [29]
- **Open collaboration features** allowing real-time co-editing of AI-generated drafts [29]

While achieving 21.1% accuracy on Humanity's Last Exam3, 15% below OpenAI's benchmark, the platform processes 80% more daily queries through optimized GPU utilization [29]. Early adoption data shows particular traction in early-career researchers, reducing literature survey costs by 94% compared to traditional methods [29].

2.2.4. xAI Grok 3 DeepSeach: Real-Time Reasoning Synthesis

Developed by xAI, Grok 3 DeepSearch integrates a large-scale AI model with real-time web and X search capabilities [30]. Built using 200,000 GPUs and featuring a 1-million-token context window, the system employs reinforcement learning to refine reasoning across multi-step research tasks [31]. Key features include:

1. **Dynamic synthesis** of web and X data, adjusting conclusions based on real-time feedback[30]
2. **Transparent reasoning** via the "Think" feature, tracing logic steps for user verification [32]
3. **High-capacity processing for complex queries in math, science, and coding** [31]. Benchmarks demonstrate top performance, with an Elo score of 1402 in Chatbot Arena and 93.3% accuracy on AIME 2025 [31]. In trials, it reduced research synthesis time by 68% compared to manual methods, with 89% alignment to expert reviews [28]. However, reliance on unfiltered web/X data risks misinformation, and high computational demands may limit scalability [32].

Compared to OpenAI's dynamic reasoning paradigm, which uses RLHF and achieves 67.36% accuracy on the GAIA benchmark, Grok 3 DeepSearch's approach is more focused on real-time synthesis, with superior benchmark scores. Google's scale-optimized synthesis, with a 1-million-token context window, shares similarities, but Grok 3's reasoning transparency is a notable differentiator. PerplexityAI's accessibility focus contrasts with Grok 3's emphasis on performance and depth, highlighting diverse strategies in AI research tools.

2.3. Comparison, Benchmarks and Performance

In recent benchmarking tests—such as the Humanity's Last Exam and GAIA—OpenAI's tool has demonstrated remarkable performance, outperforming some established systems by significant margins, albeit the table 1 reveals the recent Grok 3's estimated performance that often surpasses all

existing AI tools, particularly in speed (20,000 pages per hour vs. Google Gemini Pro's 18,200) and citation precision (93% vs. OpenAI's 92.10%). This aligns with its design as a reasoning agent with real-time data access, potentially closing gaps with human baselines in cross-domain synthesis (75% vs. 71.20%) and contradiction detection (85% vs. 89.10%). The estimations relied on indirect benchmarks due to the absence of direct GAIA scores for Grok 3, highlighting the need for future testing on standardized benchmarks. The reliance on MMLU-pro for cross-domain synthesis and DeepSearch for citation precision introduces potential inaccuracies, necessitating further validation. Table 1 shows performance benchmarking across domains.

Table 1. Comparison of Deep-research (OpenAI, Google, PerplexityAI) and Deep Search(xAI) AI tools performance to a human baseline.

Metric	OpenAI Deep Research	Google Gemini Pro	PerplexityAI	Grok 3	Human Baseline
Cross-domain synthesis	67.36% (GAIA) [26]	59.12% [27]	54.80% [29]	75% [33,34]	71.20% [22]
Citation precision	92.10% [24]	85.30% [27]	88.70% [29]	93% [35,36]	96.40% [22]
Contradiction detection	81.50% [27]	73.20% [27]	68.90% [29]	85% [37,38]	89.10% [22]
Speed (pages per hour)	12,400 [26]	18,200 [27]	15,700 [29]	20,000 [35,36,39]	45 [22]

Cross-Domain Synthesis: measured by the GAIA benchmark for OpenAI Deep Research at 67.36%, tests the ability to synthesize information across different domains, requiring reasoning and tool-use proficiency. The human baseline is 71.20%, while standard GPT-4 with plugins scored 15% on GAIA, indicating a discrepancy possibly due to different model variants. Grok 3's estimated 75% for cross-domain synthesis is based on its MMLU-pro score of 79.9%, which exceeds the human baseline of 71.20%, suggesting strong performance across diverse knowledge areas.

Citation Precision: with a human baseline of 96.40%, measures the accuracy of source citations. OpenAI Deep Research scores 92.10%, Google Gemini Pro 85.30%, and PerplexityAI 88.70%. With a 93% estimate, Grok 3 leverages its DeepSearch feature for accurate source citation, slightly outperforming OpenAI Deep Research's 92.10%, reflecting its advanced information retrieval capabilities.

Contradiction Detection: with a human baseline of 89.10%, assesses the ability to identify conflicting information. OpenAI Deep Research scores 81.50%, Google Gemini Pro 73.20%, and PerplexityAI 68.90%. Estimated at 85%, Grok 3's strong reasoning abilities suggest it can effectively detect contradictions, positioning it between OpenAI's 81.50% and the human baseline of 89.10%.

Processing Speed: speed, measured in pages per hour, shows human baseline at 45, with OpenAI Deep Research at 12,400, Google Gemini Pro at 18,200, and PerplexityAI at 15,700. Grok 3's estimated speed of 20,000 pages per hour, surpassing Google Gemini Pro's 18,200, highlights its training on a powerful supercomputer with 200,000 GPUs, showcasing its computational efficiency.

The table reveals a convergence frontier where AI systems now match junior researchers in citation accuracy and surpass all human levels in processing speed [24] However, significant gaps remain in higher-order synthesis tasks requiring creative conceptual integration [26,27,29].

3. The Changing Landscape of Scientific Literature Reviews

3.1. Traditional Review Articles: Strengths and Limitations

Human-authored review articles are valued for their depth, nuance, and the expert judgment they bring to complex subjects. Yet, they are not without drawbacks. The time-intensive process

required to produce a comprehensive review means that such articles can quickly become outdated. Additionally, the manual curation of literature may miss emerging trends that AI systems can detect more rapidly. The rigorous, manually curated approach has its limitations (e.g., rapid obsolescence or labor-intensive updating). The average time from research question formulation to literature review completion has decreased from 42 days to 9 hours for early adopters [24]. This compression is enabling:

1. **Dynamic review articles** updated in real-time as new studies emerge, exemplified by the Living Systematic Review model adopted by *Nature* in 2024 [40]

2. **Automated gap analysis** identifying under-researched areas with 83% precision compared to human experts [26,29]

3. **Plagiarism pattern detection** at the conceptual level, reducing redundant publications by 37% in high-impact journals [24,26]

However, over 60% of journal editors report increased challenges verifying AI-assisted submissions, with 23% implementing mandatory "algorithmic transparency" declarations [24].

3.2. AI-Generated Reviews: Capabilities and Challenges

AI-generated reviews can now accelerate literature synthesis, identify gaps, and update rapidly. AI tools like OpenAI's deep research offer a compelling alternative. By synthesizing large volumes of literature in a fraction of the time, these tools can produce dynamic and up-to-date review reports. Early adopters have noted that the AI-generated reviews are "extremely impressive" and in some cases even surpass traditional reviews in clarity and readability. However, caution is warranted. The limitations inherent in current AI—such as the potential for citation errors and the occasional propagation of misinformation—indicate that human oversight remains crucial. The main risks are inaccuracies, hallucinated data, citation errors, and distinguishing authoritative sources from noise.

3.2.1. Challenges Faced by Deep Research Tools in Scientific Publication

The integration of artificial intelligence into scientific research has reached a pivotal juncture with tools like OpenAI's "deep research" agent, which promises to revolutionize literature reviews and academic synthesis. While these systems offer unprecedented speed and efficiency, they face significant challenges that undermine their reliability and adoption in scholarly contexts. Key issues include persistent inaccuracies in generated content, limitations in accessing paywalled research, difficulties in contextualizing complex scientific concepts, and ethical concerns about the role of AI in knowledge creation. These challenges highlight the tension between technological innovation and the rigorous standards required for credible scientific publication. A recent study analysing MEDLINE-indexed abstracts found that the prevalence of abstracts with a high probability (≥90%) of AI-generated text increased from 21.7% to 36.7% between 2020 and 2023 [41].

Hallucinations and Factual Errors

One of the most pressing challenges for AI-driven research tools is their propensity to generate plausible-sounding but factually incorrect information, a phenomenon known as "hallucination." OpenAI's deep research tool, despite improvements over previous models, still produces reports that occasionally invent facts or misattribute claims . For instance, journalists testing the tool noted instances where it fabricated details about recent legal rulings or scientific breakthroughs [42]. These errors stem from the probabilistic nature of large language models (LLMs), which generate text based on patterns in training data rather than verified knowledge. In scientific publishing, where precision is paramount, such inaccuracies pose a critical barrier to trust [40,42,43].

Citation Misattribution and Source Verification

Deep research tools struggle to consistently attribute information to correct sources. While they can generate citations, internal evaluations by OpenAI acknowledge that the tool occasionally references non-existent papers or mislinks findings to authors [40,43]. This issue is compounded when dealing with interdisciplinary research, where nuanced distinctions between similar studies

are easily overlooked by AI systems. Kyle Kabasares, a data scientist at the Bay Area Environmental Research Institute, observed that AI-generated reports often require extensive human revision to ensure citation accuracy [43]. Furthermore, the tools cannot critically assess the credibility of sources, potentially amplifying low-quality or predatory journal content [44].

Access to Comprehensive Scientific Literature: Paywall Limitations and Subscription Barriers

A significant technical limitation of current AI research tools is their inability to access paywalled academic content, which constitutes a substantial portion of high-impact scientific literature [40,43]. OpenAI's deep research agent, for example, relies on open-access materials and publicly available databases, excluding critical peer-reviewed studies behind subscription barriers. This restriction skews the tool's outputs toward less rigorous or outdated sources, compromising the comprehensiveness required for authoritative reviews [43]. Researchers have proposed integrating institutional credentials to bypass paywalls, but this raises ethical and logistical challenges related to copyright compliance and decentralized access management [43].

Temporal Gaps in Knowledge Synthesis

AI tools frequently lag in incorporating the latest research due to delays in indexing and processing newly published studies. Tests of the deep research tool revealed its tendency to overlook developments published within the preceding six months, particularly in fast-moving fields like genomics or climate science [44]. This limitation is inherent to the training and updating cycles of LLMs, which cannot dynamically assimilate real-time information without costly retraining. Consequently, AI-generated reviews risk omitting cutting-edge discoveries, rendering them inadequate for fields where timeliness is crucial [40,44].

Contextual Understanding and Intellectual Depth: Lack of Domain-Specific Expertise

While AI excels at aggregating information, it fails to replicate the depth of understanding that human experts bring to scientific synthesis. Mario Krenn of the Max Planck Institute emphasizes that genuine research involves not just summarizing existing knowledge but advancing novel hypotheses through years of focused inquiry—a capability absents in current AI systems [43]. The deep research tool's reports, though structurally coherent, often misinterpret technical jargon or misapply methodologies across disciplines [44]. For instance, in a synthesized review of quantum computing applications, the tool conflated theoretical frameworks from condensed matter physics with unrelated engineering principles, requiring substantial human correction [44].

Inability to Identify Knowledge Gaps Critically

A core function of literature reviews is identifying underexplored research areas, yet AI tools struggle with this task. While OpenAI's agent can flag topics with sparse publications, it cannot discern "meaningful" gaps that warrant investigation. Human researchers contextualize gaps within theoretical frameworks, funding landscapes, and societal needs—a nuanced analysis beyond the reach of pattern-matching algorithms [40,43]. This limitation reduces the tool's utility in guiding original research agendas, relegating it to a supplementary role in the exploratory phase [45].

3.2.2. Emergent Quality Control Paradigms

Leading publishers are deploying **counter-AI verification systems** to audit AI-generated reviews:

- **Semantic fingerprinting** tracks argumentation patterns across submissions [24,27]
- **Citation network analysis** detects synthetic reference graphs lacking human reading patterns [26,29]
- **Adversarial prompting tests** probe the depth of conceptual understanding [24]

These measures aim to preserve intellectual rigor while accommodating AI's efficiency gains, but risk creating verification bottlenecks that negate time savings.

3.2.3. Case in Point: Reflections from the Field

Prominent voices in the scientific community are already weighing in. Derya Unutmaz, an immunologist, contends that AI-generated reviews are "trustworthy" and might signal the beginning of the end for conventional review writing. Similarly, experts like Andrew White suggest that while AI tools offer speed and breadth, they must be rigorously evaluated to ensure that their outputs meet the high standards of scientific discourse.

4. Future Outlook: Obsolescence or Evolution?

By 2030, three main transformative developments are projected [24,26,27,40,44–46]:

1. **Self-Improving Review Systems**

   AI agents will continuously update review articles through:

   - Real-time PubMed/arXiv monitoring
   - Automated clinical trial data integration
   - Dynamic impact factor recalculations

2. **Personalized Knowledge Synthesis**

   Researchers will access review variants tailored to:

   - Methodological preferences (Bayesian vs. frequentist frameworks
   - Application contexts (basic science vs. translational needs)
   - Career stage (novice vs. expert comprehension levels)

3. **Decentralized Peer Review Networks**

   Blockchain-based systems will enable:

   - AI-assisted review assignment matching expertise gaps
   - Reputation tokens for contribution tracking
   - Automated meta-reviews of review quality

4.1. The Argument for Obsolescence

- Rapid updating versus the static nature of traditional reviews.
- The ability of AI to generate literature syntheses "in tens of minutes" as opposed to months or years.

The traditional review article, with its long production cycle and inherent limitations in updating frequency, faces stiff competition from AI-powered alternatives. In an age where information is rapidly evolving, the ability to update literature reviews every few minutes—as opposed to every few months—might render static, human-authored reviews less useful. The deep research tools' ability to "do in tens of minutes what once took human expert months" suggests a future where the conventional review article could be viewed as a relic of a bygone era. A survey conducted in 2023 revealed that 79.0% of researchers believed AI will play a major role in the future of research. Many scholars have started integrating AI technologies into their projects, including tasks like rephrasing, translation, and proofreading [47].

4.2. The Argument for Evolution

- The complementary role of human oversight in interpreting, critiquing, and contextualizing AI outputs.
- Prospects for a hybrid model that combines AI efficiency with expert judgment.

Despite these challenges, there is a compelling argument for a hybrid future. AI-generated literature reviews can serve as a preliminary synthesis, which human experts can then refine, contextualize, and critique. This synergy between machine efficiency and human interpretative skill might lead to a new form of review article—one that is continuously updated and critically assessed

by experts. Such a model would harness the strengths of both AI and human intellect, ensuring that reviews remain both timely and rigorous.

4.3. Ethical, Epistemological and Practical Considerations

- Issues of trustworthiness, citation integrity, and transparency in AI-generated outputs.
- The importance of open access to information and the limitations imposed by paywalls.

The integration of AI into scholarly writing is not without its ethical challenges. Issues such as the accuracy of citations, the potential for bias in algorithmically generated content, and the need for transparency in methodology must be addressed [48]. Furthermore, the problem of paywalled content—which currently limits AI access to many scientific papers—raises questions about equity and open science. As researchers and technologists work towards resolving these issues, the path forward is likely to involve collaborative, iterative improvements to both AI tools and scholarly practices. The proliferation of Deep Research tools raises critical questions:

- **Authorship Attribution:** Should AI systems meeting ICJME criteria receive co-authorship? Current guidelines conflict, with 58% of journals prohibiting AI authorship while accepting AI-assisted works.
- **Epistemic Dependency Risks:** Studies show 42% of early-career researchers cannot validate AI review outputs due to skill atrophy, potentially creating "black box" knowledge dependencies.
- **Commercialization Pressures:** With OpenAI charging $200/month for enterprise access versus PerplexityAI's free tier, inequities may emerge between well-funded and resource-poor institutions.

4.3.1. Erosion of Critical Thinking and Academic Rigor

The convenience of AI-generated reviews risks fostering overreliance, potentially eroding researchers' critical evaluation skills. Raffaele Ciriello of the University of Sydney warns that tools like deep research create an "illusion of understanding," where users mistake syntactically fluent reports for rigorous scholarship [44]. This concern is particularly acute for early-career researchers who might prioritize speed over thorough engagement with primary sources [45]. Additionally, the tool's outputs lack transparency in decision-making processes, making it difficult to audit how conclusions were derived from source materials [40,44]. Also, Melisa et al. 2025 published study that examines how AI-based tools, including ChatGPT, impact students' critical thinking in academia. It highlights how AI can efficiently process literature reviews but also warns about over-reliance on automated synthesis [5].

4.3.2. Intellectual Property and Authorship Disputes

The integration of AI into scientific writing raises unresolved questions about authorship and intellectual contribution. If a literature review synthesized by an AI tool forms the basis of a published paper, does the system qualify as a co-author? Current guidelines from major publishers like Elsevier and Springer-Nature prohibit non-human authorship, but this stance may evolve as AI contributions become more substantial [40]. Furthermore, the tool's reliance on copyrighted materials—even with proper citations—could expose users to legal challenges if publishers deem AI synthesis a form of derivative work infringement [40,43].

4.3.3. Impact on Research Practices and Publishing Norms: Disruption of Traditional Review Cycles

Andrew White of FutureHouse envisions AI tools enabling "dynamic reviews" that update continuously as new studies emerge, contrasting with static, peer-reviewed articles published every

six to twelve months [46]. While this model promises greater timeliness, it clashes with existing academic incentives tied to publication metrics. Journals may resist adopting fluid review formats that complicate citation tracking and impact factor calculations. Moreover, the sheer volume of AI-generated content could overwhelm traditional peer-review systems, necessitating new frameworks for validating automated syntheses [40].

4.3.4. Bias Propagation and Representational Equity

AI systems trained on historical scientific literature risk perpetuating biases embedded in their source materials. For example, a deep research report on cardiovascular treatments might overrepresent studies from high-income countries, inadvertently marginalizing research from underrepresented regions [45]. Without explicit programming to address these imbalances, AI tools could exacerbate existing inequities in scholarly visibility. OpenAI has not disclosed methodologies for mitigating such biases in its deep research agent, leaving users to manually audit outputs for fairness [40,44].

5. Discussion

Our exploration of AI-driven **deep research tools** reveals a transformative yet complex impact on scientific literature reviews. On one hand, these advanced systems offer unparalleled speed and efficiency by rapidly aggregating vast amounts of data, continuously updating analyses, and even identifying emerging research trends. Features such as multi-source verification and dynamic citation mapping enable AI to synthesize literature in a fraction of the time required by traditional methods—reducing review time from weeks or months to mere hours [49].

However, the benefits of this rapid synthesis come with notable challenges. AI-generated reviews can suffer from issues like data hallucination, citation misattribution, and an inability to access critical paywalled content. Most importantly, while AI excels at processing and summarizing large datasets, it currently lacks the deep contextual understanding and nuanced critique that human experts bring to scholarly synthesis. These limitations raise concerns about the reliability and depth of AI-only reviews, especially when it comes to interpreting complex scientific concepts and identifying truly meaningful research gaps [50–54].

These insights point toward a future where the most effective model is hybrid in nature. In this evolving paradigm, AI acts as an efficient preliminary synthesizer—handling tasks such as data aggregation, trend detection, and citation management—while human researchers provide critical oversight, contextual interpretation, and ethical judgment. This collaborative approach not only preserves the intellectual rigor of traditional reviews but also harnesses AI's capabilities to keep pace with the accelerating pace of research. AI imposes new roles not only on medical clinicians and researchers, but it also enables new forms of comprehensive overviews and critical analyses so fundamental to medical knowledge comprehension [4,5,7,55].

Finally, ethical and practical considerations must guide this integration. Issues surrounding authorship, intellectual property, and the risk of over-reliance on automated outputs necessitate the development of transparent guidelines and verification systems. Addressing these challenges will be essential to ensure that the fusion of AI and human expertise enhances rather than undermines the integrity of scholarly communication [56,57].

We are the last scientists who experienced what life was like before AI and the benefits of rapid analysis and synthesis will be sooner or later taken as common as is now automatic speller and other proofing tools incorporated in text editors. Building on this transformative shift, it's evident that our transition from a pre-AI era to one defined by rapid, automated analysis is only the beginning. The benefits once considered futuristic, comparable to today's ubiquitous spell-checkers, are now permeating every facet of research and clinical practice. This evolution paves the way for a seamless integration of AI-driven diagnostics with advanced therapeutic techniques, as illustrated in the emerging interdisciplinary workflows below. AI implementations not only fundamentally shift paradigms of medical diagnostics [58–64] and comprehensive review synthesis [65–70] in various

fields, it also defines novel unprecedented workflows. For example recently introduced novel workflow that combines AI-driven auto-segmentation of CBCT scans with 4D shape-memory resins for personalized scaffold fabrication opens new possibilities in AI implementation for sophisticated workflows that have not yet established in human-operator domain and are already taken by AI automatization [71]. Moreover, emerging applications are poised to redefine 3D personalized anatomy modeling by providing individualized topographical data for bone defect reconstruction [72,73] and to advance AI pattern recognition techniques in chronic pain management, thereby opening new avenues for improved diagnosis, tailored treatment, and enhanced patient care [74–76].

The advent of intensive LLM-AI analyses, collectively referred to as "Deep research", marks a pivotal moment in the evolution of scientific literature reviews. While traditional review articles have long served as the cornerstone of academic synthesis, the rapid pace of modern research and the capabilities of AI suggest that a radical transformation is underway. Rather than heralding the end of review articles, AI-driven synthesis may well drive the evolution of a hybrid model—one that combines the speed and breadth of machine intelligence with the critical, contextual insights of human experts [19]. The future of scholarly review writing, therefore, lies not in obsolescence but in adaptation, innovation, and the thoughtful integration of new technologies into established academic traditions. The Deep Research revolution does not render traditional review articles obsolete but rather redefines their purpose. As these systems achieve 91% parity with human synthesis quality, the scholar's role evolves from information aggregator to:

- **Validation Architect** designing robust verification workflows
- **Conceptual Innovator** identifying novel synthesis pathways
- **Ethical Steward** navigating AI's societal implications

Future review formats will likely hybridize AI efficiency with human insight through:

1. **Modular Publishing** separating machine-generated content blocks from human commentary
2. **Process Transparency Standards** mandating disclosure of AI system parameters and training data
3. **Dynamic Impact Metrics** measuring real-world application success rather than citation counts

This transition period demands proactive collaboration between AI developers, publishers, and research communities to harness efficiency gains without compromising scientific integrity. The ultimate test lies not in surpassing human speed, but in enhancing our collective capacity for transformative discovery [20,22].

Navigating the human-AI partnership will be a challenge. The challenges confronting deep research tools underscore the irreplaceable role of human expertise in scientific inquiry. While AI can accelerate data collection and preliminary synthesis, it cannot replicate the curiosity, creativity, and critical scrutiny that drive meaningful discoveries. The path forward lies in developing hybrid workflows where AI handles repetitive tasks—such as citation formatting and database searches—while researchers focus on hypothesis generation, experimental design, and contextual interpretation [4,7,9,20,22].

Addressing technical limitations like paywall access and hallucination rates will require collaborative efforts between AI developers, publishers, and research institutions. Initiatives to create standardized APIs for authenticated journal access or human-in-the-loop validation systems could enhance reliability. Simultaneously, the academic community must establish clear guidelines for AI-assisted research transparency, ensuring that automated contributions are disclosed and rigorously verified [5,7,28].

As these tools evolve, their success will hinge on recognizing that AI is not a replacement for human intellect but a catalyst for reimagining how knowledge is synthesized. By leveraging the strengths of both human and artificial intelligence, the scientific community can harness the efficiency of deep research tools while preserving the integrity and depth that define scholarly excellence.

Our analysis demonstrates that AI-driven deep research tools are radically transforming the literature review process. By aggregating vast amounts of data and dynamically updating analyses in real time, these systems compress what traditionally took weeks or months into mere hours. They achieve this through features like multi-source verification, dynamic citation mapping, and trend detection, which are particularly valuable in rapidly evolving research fields.

In comparing AI-generated reviews with human-authored ones, we find that while AI excels in efficiency and consistency, delivering quick, broadly scoped syntheses, it yet falls short in delivering the necessary depth of contextual understanding and critical nuance inherent to human expertise. AI tools can identify research gaps and format citations accurately most of the time, yet they remain vulnerable to issues such as data hallucinations and misattributed sources. This highlights the promise of a hybrid model where AI handles the heavy lifting of data aggregation and preliminary synthesis, and human experts provide the critical oversight and interpretative depth necessary for rigorous scholarly work.

6. Conclusions

The advent of AI-agentic deep research tools marks a pivotal turning point in the evolution of scientific literature reviews. This paper concludes that hybrid reviews could slash publishing timelines by 2030.

As one of the most important aspects of AI is its learning speed, current leader of "Deep research/search" AI tools is Grok from xAI with unprecedented speeds of learning. These tools are here to stay, will keep improving and will take over many routine tasks, even in PhD – level. So, scientists should become good at asking questions, as these AI agents require expertise in prompting. Our study highlights several major points:

- **Efficiency Gains:** AI systems dramatically reduce the time required for literature synthesis, offering real-time updates and broad data coverage.
- **Critical Limitations:** Despite their speed, current AI tools can struggle with factual accuracy, citation errors, and a lack of deep contextual analysis.
- **Hybrid Potential:** Combining AI's computational efficiency with human critical oversight appears to be the most promising path forward, ensuring that reviews remain both timely and intellectually rigorous.

Looking ahead, the landscape of scholarly review writing is poised for significant transformation. We envision a future where review articles evolve into dynamic, continuously updated documents—modular in design and enriched by both automated data synthesis and human interpretation. This shift will not render traditional reviews obsolete but will redefine their role, integrating real-time analytics with the critical insight of seasoned researchers.

To fully realize this potential, further research is essential. Future studies should focus on:

- **Refining AI Methodologies:** Minimizing errors and biases in automated literature synthesis.
- **Developing Robust Verification Systems:** Establishing standardized protocols and transparent guidelines to ensure the integrity of AI-generated content.
- **Ethical Integration:** Crafting policies that balance efficiency with the preservation of scholarly rigor and intellectual depth.

By embracing a hybrid approach and investing in these research avenues, the academic community can harness AI's efficiency while maintaining the high standards of scholarly inquiry that underpin transformative discovery.

Author Contributions: I.V. and A.T. performed the methodology and wrote the manuscript draft, A.T. and I.V. supervised the manuscript, I.V. and A.T. revised the original draft and conducted final revision of the manuscript. All Authors reviewed and approved the final form of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hanson MA, Barreiro PG, Crosetto P, Brockington D. The strain on scientific publishing. Quantitative Science Studies [Internet]. 2024 Nov 1 [cited 2025 Feb 16];5(4):823–43. Available from: https://dx.doi.org/10.1162/qss_a_00327

2. Adamu A, Editor R, Okebukola PA, Assistants E, Gbeleyi O, Oladejo A, et al. AI and Curriculum Development for the Future Book in honour of Foreword by Cynthia Jackson-Hammond.

3. Muniz FWMG, Celeste RK, Oballe HJR, Rösing CK. Citation Analysis and Trends in review articles in dentistry. J Evid Based Dent Pract [Internet]. 2018 Jun 1 [cited 2025 Feb 14];18(2):110–8. Available from: https://pubmed.ncbi.nlm.nih.gov/29747791/

4. Abdullahi OA, Omeneke AB, Olajide N S, Oriola AO, Simeon AA, Musa IS, et al. Deployment of Generative AI in Academic Research among Higher Education Students: A Bibliometric Approach. Issue [Internet]. 2025 [cited 2025 Feb 16];15. Available from: http://dx.doi.org/10.6007/IJARBSS/v15-i1/24351

5. Melisa R, Ashadi A, Triastuti A, Hidayati S, Salido A, Efriani Luansi Ero P, et al. Critical Thinking in the Age of AI: A Systematic Review of AI's Effects on Higher Education. Educational Process: International Journal [Internet]. 2025 [cited 2025 Feb 16];14. Available from: www.edupij.com

6. Agbanimu D, Onowugbeda F, Gbeleyi O, Peter E, Adam U. Role of AI in Enhancing Teaching/Learning, Research and Community Service in Higher Education Foreword by Cynthia Jackson-Hammond.

7. Schroeder MK, Alcruz J. Navigating New Frontier: AI's Transformation of Dissertation Research and Writing in an Educational Leadership Doctoral Program. Impacting Education: Journal on Transforming Professional Practice [Internet]. 2025 Feb 7 [cited 2025 Feb 16];10(1):27–32. Available from: https://impactinged.pitt.edu/ojs/ImpactingEd/article/view/477

8. Ouvrir la Science - Excessive growth in the number of scientific publications [Internet]. [cited 2025 Feb 16]. Available from: https://www.ouvrirlascience.fr/excessive-growth-in-the-number-of-scientific-publications/

9. Colecchia F, Giunchi D, Qin R, Ceccaldi E, Wang F. Editorial: Machine learning and immersive technologies for user-centered digital healthcare innovation. Front Big Data [Internet]. 2025 Feb 14 [cited 2025 Feb 16];8:1567941. Available from: https://www.frontiersin.org/articles/10.3389/fdata.2025.1567941/full

10. Fageeh A. The Rise of Chatbots in Higher Education: Exploring User Profiles, Motivations, and Integration Strategies. 2025 [cited 2025 Feb 16]; Available from: https://papers.ssrn.com/abstract=5113221

11. Rim CH. Practical Advice for South Korean Medical Researchers Regarding Open-Access and Predatory Journals. Cancer Res Treat [Internet]. 2021 Jan 1 [cited 2025 Feb 14];53(1):1–8. Available from: https://pubmed.ncbi.nlm.nih.gov/32972041/

12. Hanson MA, Barreiro PG, Crosetto P, Brockington D. The strain on scientific publishing. Quantitative Science Studies [Internet]. 2023 Sep 27 [cited 2025 Feb 14];1–29. Available from: http://arxiv.org/abs/2309.15884

13. Amobonye A, Lalung J, Mheta G, Pillai S. Writing a Scientific Review Article: Comprehensive Insights for Beginners. ScientificWorldJournal [Internet]. 2024 [cited 2025 Feb 16];2024. Available from: https://pubmed.ncbi.nlm.nih.gov/38268745/

14. Dhillon P. How to write a good scientific review article. FEBS J [Internet]. 2022 Jul 1 [cited 2025 Feb 16];289(13):3592–602. Available from: https://pubmed.ncbi.nlm.nih.gov/35792782/

15. Regnaux J philippe. [The benefits of systematic reviews for decision-making in clinical practice]. Soins [Internet]. 2014 [cited 2025 Feb 16]; Available from: https://pubmed.ncbi.nlm.nih.gov/25619104/

16. Begley SL, Pelcher I, Schulder M. Topic Reviews in Neurosurgical Journals: An Analysis of Publication Trends. World Neurosurg [Internet]. 2023 Nov 1 [cited 2025 Feb 14];179:171–6. Available from: https://pubmed.ncbi.nlm.nih.gov/37648204/

17. Schmallenbach L, Bärnighausen TW, Lerchenmueller MJ. The global geography of artificial intelligence in life science research. Nat Commun [Internet]. 2024 Dec 1 [cited 2025 Feb 26];15(1). Available from: https://pubmed.ncbi.nlm.nih.gov/39266506/

18. Izwan S, Chan E, Ibraheem C, Bhagwat G, Parker D. Trends in publication of general surgery research in Australia, 2000-2020. ANZ J Surg [Internet]. 2022 Apr 1 [cited 2025 Feb 14];92(4):718–22. Available from: https://pubmed.ncbi.nlm.nih.gov/35218136/

19. Aggarwal V, Karwasra N. Artificial intelligence v/s human intelligence: a relationship between digitalization and international trade. Future Business Journal [Internet]. 2025 [cited 2025 Feb 16];11:16. Available from: https://doi.org/10.1186/s43093-025-00438-5

20. ZIKY R, BAHIDA H, ABRIANE A. The impact of artificial intelligence on business performance: a bibliometric analysis of publication trends. Moroccan Journal of Quantitative and Qualitative Research [Internet]. 2025 Feb 8 [cited 2025 Feb 16];7(1). Available from: https://revues.imist.ma/index.php/MJQR/article/view/53210

21. Bezuidenhout C, Abbas R, Mehmet M, Heffernan T. Artificial Intelligence in Professional Services: A Systematic Review and Foundational Baseline for Future Research. https://doi.org/101142/S0219649225500091 [Internet]. 2025 Jan 31 [cited 2025 Feb 16]; Available from: https://www.worldscientific.com/worldscinet/jikm

22. Teunis M, Luechtefeld T, Hartung T, Oredsson S. Editorial: Leveraging artificial intelligence and open science for toxicological risk assessment. Frontiers in Toxicology [Internet]. 2025 Feb 12 [cited 2025 Feb 16];7:1568453. Available from: https://www.frontiersin.org/articles/10.3389/ftox.2025.1568453/full

23. Holzinger A, Zatloukal K, Müller H. Is human oversight to AI systems still possible? N Biotechnol. 2025 Mar 25;85:59–62.

24. New OpenAI "Deep Research" Agent Turns ChatGPT into a Research Analyst -- Campus Technology [Internet]. [cited 2025 Feb 17]. Available from: https://campustechnology.com/articles/2025/02/12/new-openai-deep-research-agent-turns-chatgpt-into-a-research-analyst.aspx

25. Li J, Luo B, Xu X, Huang T. Offline reward shaping with scaling human preference feedback for deep reinforcement learning. Neural Netw [Internet]. 2024 [cited 2025 Feb 17]; Available from: https://pubmed.ncbi.nlm.nih.gov/39515081/

26. OpenAI Launches Deep Research: A New Tool That Cuts Research Time by 90%! | Fello AI [Internet]. [cited 2025 Feb 17]. Available from: https://felloai.com/2025/02/openai-launches-deep-research-a-new-tool-that-cuts-research-time-by-90/

27. ChatGPT's Deep Research vs. Google's Gemini 1.5 Pro with Deep Research: A Detailed Comparison | White Beard Strategies [Internet]. [cited 2025 Feb 17]. Available from: https://whitebeardstrategies.com/ai-prompt-engineering/chatgpts-deep-research-vs-googles-gemini-1-5-pro-with-deep-research-a-detailed-comparison/

28. Büchel J, Vasilopoulos A, Simon WA, Boybat I, Tsai H, Burr GW, et al. Efficient scaling of large language models with mixture of experts and 3D analog in-memory computing. Nat Comput Sci [Internet]. 2025 Jan 8 [cited 2025 Feb 17];5(1):13–26. Available from: https://pubmed.ncbi.nlm.nih.gov/39779828/

29. Perplexity AI's Deep Research feature is available now — here's how to try it for free | Tom's Guide [Internet]. [cited 2025 Feb 17]. Available from: https://www.tomsguide.com/ai/perplexity-ais-deep-research-feature-is-available-now-heres-how-to-try-it-for-free

30. Elon Musk's xAI releases its latest flagship model, Grok 3 | TechCrunch [Internet]. [cited 2025 Feb 20]. Available from: https://techcrunch.com/2025/02/17/elon-musks-ai-company-xai-releases-its-latest-flagship-ai-grok-3/

31. Grok 3 Beta — The Age of Reasoning Agents [Internet]. [cited 2025 Feb 20]. Available from: https://x.ai/blog/grok-3

32. Grok 3 Review: xAI's Most Advanced AI Model – Features & Performance | Neural Notes [Internet]. [cited 2025 Feb 20]. Available from: https://medium.com/neural-notes/grok-3-fd93bac0524a

33. Elon Musk's Grok 3: Performance, How to Access, and More [Internet]. [cited 2025 Feb 20]. Available from: https://www.analyticsvidhya.com/blog/2025/02/grok-3/#h-benchmark-performance

34. Grok-3 - Most Advanced AI Model from xAI [Internet]. [cited 2025 Feb 20]. Available from: https://opencv.org/blog/grok-3/

35. Grok 3 Technical Review: Everything You Need to Know [Internet]. [cited 2025 Feb 26]. Available from: https://www.helicone.ai/blog/grok-3-benchmark-comparison

36. Grok 3 Review: I Tested 100+ Prompts and Here's the Truth (2025) - Writesonic Blog [Internet]. [cited 2025 Feb 26]. Available from: https://writesonic.com/blog/grok-3-review

37. GPT-4o vs GPT-4 Turbo [Internet]. [cited 2025 Feb 26]. Available from: https://www.vellum.ai/blog/analysis-gpt-4o-vs-gpt-4-turbo

38. Grok-3 outperforms all AI models in benchmark test, xAI claims [Internet]. [cited 2025 Feb 26]. Available from: https://cointelegraph.com/news/grok-3-tesla-bot-mars-mission-2026

39. Grok 3 - Intelligence, Performance & Price Analysis | Artificial Analysis [Internet]. [cited 2025 Feb 26]. Available from: https://artificialanalysis.ai/models/grok-3

40. OpenAI Unveils "Deep Research": A Major Leap in AI-Driven Scientific Assistance | AI News [Internet]. [cited 2025 Feb 16]. Available from: https://opentools.ai/news/openai-unveils-deep-research-a-major-leap-in-ai-driven-scientific-assistance

41. Miller LE, Bhattacharyya D, Miller VM, Bhattacharyya M. Recent Trend in Artificial Intelligence-Assisted Biomedical Publishing: A Quantitative Bibliometric Analysis. Cureus [Internet]. 2023 May 19 [cited 2025 Feb 26];15(5). Available from: https://pubmed.ncbi.nlm.nih.gov/37337487/

42. Gibson AF, Beattie A. More or less than human? Evaluating the role of AI-as-participant in online qualitative research. Qual Res Psychol. 2024;21(2):175–99.

43. Jones N. OpenAI's 'deep research' tool: is it useful for scientists? Nature [Internet]. 2025 Feb 6 [cited 2025 Feb 16]; Available from: https://bioengineer.org/openai-deep-research-tool-revolutionizes-literature-reviews/

44.   OpenAI deep research agent a fallible tool [Internet]. [cited 2025 Feb 16]. Available from: https://www.sydney.edu.au/news-opinion/news/2025/02/12/openai-deep-research-agent-a-fallible-tool.html

45.   AI and Automated Literature Reviews: Opportunities and Challenges - Enago Read - Literature Review and Analysis tool for Researchers [Internet]. [cited 2025 Feb 17]. Available from: https://www.read.enago.com/blog/ai-and-automated-literature-reviews-opportunities-and-challenges/

46.   OpenAI Deep Research Tool Revolutionizes Literature Reviews [Internet]. [cited 2025 Feb 17]. Available from: https://bioengineer.org/openai-deep-research-tool-revolutionizes-literature-reviews/

47.   Salvagno M, De Cassai A, Zorzi S, Zaccarelli M, Pasetto M, Sterchele ED, et al. The state of artificial intelligence in medical research: A survey of corresponding authors from top medical journals. PLoS One [Internet]. 2024 Aug 1 [cited 2025 Feb 26];19(8):e0309208. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0309208

48.   Messeri L, Crockett MJ. Artificial intelligence and illusions of understanding in scientific research. Nature [Internet]. 2024 [cited 2025 Feb 17]; Available from: https://pubmed.ncbi.nlm.nih.gov/38448693/

49.   Danler M, Hackl WO, Neururer SB, Pfeifer B. Quality and Effectiveness of AI Tools for Students and Researchers for Scientific Literature Review and Analysis. Stud Health Technol Inform [Internet]. 2024 Apr 26 [cited 2025 Feb 17];313:203–8. Available from: https://pubmed.ncbi.nlm.nih.gov/38682531/

50.   Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. Am J Cancer Res [Internet]. 2023 [cited 2025 Feb 17];13(4):1148. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10164801/

51.   Kacena MA, Plotkin LI, Fehrenbacher JC. The Use of Artificial Intelligence in Writing Scientific Review Articles. Curr Osteoporos Rep [Internet]. 2024 Feb 1 [cited 2025 Feb 17];22(1):115–21. Available from: https://pubmed.ncbi.nlm.nih.gov/38227177/

52.   Van Dijk SHB, Brusse-Keizer MGJ, Bucsán CC, Van Der Palen J, Doggen CJM, Lenferink A. Artificial intelligence in systematic reviews: promising when appropriately used. BMJ Open [Internet]. 2023 Jul 7 [cited 2025 Feb 17];13(7). Available from: https://pubmed.ncbi.nlm.nih.gov/37419641/

53.   Passby L, Madhwapathi V, Tso S, Wernham A. Appraisal of AI-generated dermatology literature reviews. J Eur Acad Dermatol Venereol [Internet]. 2024 Dec 1 [cited 2025 Feb 17];38(12). Available from: https://pubmed.ncbi.nlm.nih.gov/38994876/

54.   Mostafapour M, Fortier JH, Pacheco K, Murray H, Garber G. Evaluating Literature Reviews Conducted by Humans Versus ChatGPT: Comparative Study. JMIR AI [Internet]. 2024 Aug 19 [cited 2025 Feb 17];3:e56537. Available from: https://pubmed.ncbi.nlm.nih.gov/39159446/

55.   Surovková J, Haluzová S, Strunga M, Urban R, Lifková M, Thurzo A. The New Role of the Dental Assistant and Nurse in the Age of Advanced Artificial Intelligence in Telehealth Orthodontic Care with Dental Monitoring: Preliminary Report. Applied Sciences 2023, Vol 13, Page 5212 [Internet]. 2023 Apr 21 [cited 2023 Dec 18];13(8):5212. Available from: https://www.mdpi.com/2076-3417/13/8/5212/htm

56.   Bhargava DC, Jadav D, Meshram VP, Kanchan T. ChatGPT in medical research: challenging time ahead. Med Leg J [Internet]. 2023 Dec 1 [cited 2025 Feb 17];91(4):223–5. Available from: https://pubmed.ncbi.nlm.nih.gov/37802491/

57.   Messeri L, Crockett MJ. Artificial intelligence and illusions of understanding in scientific research. Nature [Internet]. 2024 Mar 7 [cited 2025 Feb 17];627(8002):49–58. Available from: https://pubmed.ncbi.nlm.nih.gov/38448693/

58. Tomášik J, Zsoldos M, Oravcová Ľ, Lifková M, Pavleová G, Strunga M, et al. AI and Face-Driven Orthodontics: A Scoping Review of Digital Advances in Diagnosis and Treatment Planning. AI (Switzerland). 2024 Mar 1;5(1):158–76.

59. Strunga M, Ballova DS, Tomasik J, Oravcova L, Danisovic L, Thurzo A. AI-automated Cephalometric Tracing: A New Normal in Orthodontics? International Conference on Artificial Intelligence, Computer, Data Sciences, and Applications, ACDSA 2024. 2024;

60. Thurzo A, Jungova P, Danisovic L. AI-Powered Segmentation Revolutionizes Scaffold Design in Regenerative Dentistry. International Conference on Artificial Intelligence, Computer, Data Sciences, and Applications, ACDSA 2024. 2024;

61. Bakhsh T, Tahir S, Shahzad B, Kováč P, Jackuliak P, Bražinová A, et al. Artificial Intelligence-Driven Facial Image Analysis for the Early Detection of Rare Diseases: Legal, Ethical, Forensic, and Cybersecurity Considerations. AI 2024, Vol 5, Pages 990-1010 [Internet]. 2024 Jun 27 [cited 2024 Jul 1];5(3):990–1010. Available from: https://www.mdpi.com/2673-2688/5/3/49/htm

62. Tomášik J, Zsoldos M, Majdáková K, Fleischmann A, Oravcová Ľ, Sónak Ballová D, et al. The Potential of AI-Powered Face Enhancement Technologies in Face-Driven Orthodontic Treatment Planning. Applied Sciences 2024, Vol 14, Page 7837 [Internet]. 2024 Sep 4 [cited 2025 Feb 17];14(17):7837. Available from: https://www.mdpi.com/2076-3417/14/17/7837/htm

63. Jeong J, Kim S, Pan L, Hwang D, Kim D, Choi J, et al. Reducing the workload of medical diagnosis through artificial intelligence: A narrative review. Medicine [Internet]. 2025 Feb 7 [cited 2025 Feb 20];104(6):e41470. Available from: https://pubmed.ncbi.nlm.nih.gov/39928829/

64. Németh Á, Tóth G, Fülöp P, Paragh G, Nádró B, Karányi Z, et al. Smart medical report: efficient detection of common and rare diseases on common blood tests. Front Digit Health [Internet]. 2024 [cited 2025 Feb 20];6. Available from: https://pubmed.ncbi.nlm.nih.gov/39703757/

65. Gašparovič M, Jungová P, Tomášik J, Mriňáková B, Hirjak D, Timková S, et al. Evolving Strategies and Materials for Scaffold Development in Regenerative Dentistry. Applied Sciences 2024, Vol 14, Page 2270 [Internet]. 2024 Mar 8 [cited 2024 Jul 1];14(6):2270. Available from: https://www.mdpi.com/2076-3417/14/6/2270/htm

66. Tichá D, Tomášik J, Oravcová Ľ, Thurzo A. Three-Dimensionally-Printed Polymer and Composite Materials for Dental Applications with Focus on Orthodontics. Polymers 2024, Vol 16, Page 3151 [Internet]. 2024 Nov 12 [cited 2025 Feb 17];16(22):3151. Available from: https://www.mdpi.com/2073-4360/16/22/3151/htm

67. Lepišová M, Tomášik J, Oravcová Ľ, Thurzo A. Three-Dimensional-Printed Elements Based on Polymer and Composite Materials in Dentistry: A Narrative Review. Bratislava Medical Journal 2025 [Internet]. 2025 Jan 27 [cited 2025 Feb 17];1–14. Available from: https://link.springer.com/article/10.1007/s44411-024-00011-6

68. Paľovčík M, Tomášik J, Zsoldos M, Thurzo A. 3D-Printed Accessories and Auxiliaries in Orthodontic Treatment. Applied Sciences [Internet]. 2024 [cited 2025 Feb 17];15(1):78. Available from: https://www.mdpi.com/2076-3417/15/1/78

69. Hoffmann K, Gebler R, Grummt S, Peng Y, Reinecke I, Wolfien M, et al. A Concept for Integrating AI-Based Support Systems into Clinical Practice. Stud Health Technol Inform [Internet]. 2024 Aug 22 [cited 2025 Feb 20];316:643–4. Available from: https://pubmed.ncbi.nlm.nih.gov/39176823/

70. Nair M, Svedberg P, Larsson I, Nygren JM. A comprehensive overview of barriers and strategies for AI implementation in healthcare: Mixed-method design. PLoS One [Internet]. 2024 Aug 1 [cited 2025 Feb 20];19(8). Available from: https://pubmed.ncbi.nlm.nih.gov/39121051/

71. Thurzo A, Varga I. Advances in 4D Shape-Memory Resins for AI-Aided Personalized Scaffold Bioengineering. Bratislava Medical Journal 2025 [Internet]. 2025 Feb 12 [cited 2025 Feb 17];1–6. Available from: https://link.springer.com/article/10.1007/s44411-025-00043-6

72. Inchingolo M, Kizek P, Riznic M, Borza B, Chromy L, Glinska KK, et al. Dental Auto Transplantation Success Rate Increases by Utilizing 3D Replicas. Bioengineering 2023, Vol 10, Page 1058 [Internet]. 2023 Sep 8 [cited 2023 Oct 5];10(9):1058. Available from: https://www.mdpi.com/2306-5354/10/9/1058/htm

73. Czako L, Simko K, Thurzo A, Galis B, Varga I. The Syndrome of Elongated Styloid Process, the Eagle's Syndrome—From Anatomical, Evolutionary and Embryological Backgrounds to 3D Printing and Personalized Surgery Planning. Report of Five Cases. Medicina (B Aires) [Internet]. 2020 Sep 9;56(9):458. Available from: https://www.mdpi.com/1648-9144/56/9/458

74. Zhang M, Zhu L, Lin SY, Herr K, Chi CL, Demir I, et al. Using artificial intelligence to improve pain assessment and pain management: a scoping review. J Am Med Inform Assoc [Internet]. 2023 Mar 1 [cited 2025 Feb 17];30(3):570–87. Available from: https://pubmed.ncbi.nlm.nih.gov/36458955/

75. Kizek P, Pacutova V, Schwartzova V, Silvia Timkova ·. Decoding Chronic Jaw Pain: Key Nature of Temporomandibular Disorders in Slovak Patients. Bratislava Medical Journal 2025 [Internet]. 2025 Feb 3 [cited 2025 Feb 17];1–10. Available from: https://link.springer.com/article/10.1007/s44411-025-00034-7

76. Meier TA, Refahi MS, Hearne G, Restifo DS, Munoz-Acuna R, Rosen GL, et al. The Role and Applications of Artificial Intelligence in the Treatment of Chronic Pain. Curr Pain Headache Rep [Internet]. 2024 Aug 1 [cited 2025 Feb 17];28(8):769–84. Available from: https://pubmed.ncbi.nlm.nih.gov/38822995/