https://doi.org/10.59761/RCR5239

Writing review articles with AI: the TRACE digital co-expert approach

Valentine P. Ananikov,[a,b]* Daria M. Arkhipova,[a] Alexey S. Kashin,[a] Artem P. Vorozhtsov[a]

[a] *N.D.Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation*
[b] *RUDN University, 117198 Moscow, Russian Federation*

Review articles remain among the most influential formats for organizing knowledge, evaluating evidence and defining future research directions. However, review preparation has become increasingly difficult as publication volume, interdisciplinarity and methodological heterogeneity continue to grow. Artificial intelligence (AI) is now entering scholarly workflows as a practical instrument for search, screening, classification, summarization, drafting and editorial refinement. However, the usefulness of AI in review preparation depends not on automation alone, but on whether it is embedded into a robust scientific workflow with human oversight. Here, we propose TRACE — Transparent Review Acceleration with Co-Expert AI — a framework for AI-assisted review writing based on the concept of a digital co-expert. We outline a general 12-stage workflow for review preparation, define the functional roles that AI can play at each stage, and analyze the risks introduced by AI at every stage. In the case study, estimated speedups were up to 22-fold for AI-assisted stages and more than threefold for the complete review-preparation workflow. On this basis, we argue that high-quality review preparation should not be framed as human replacement, but as structured collaboration in which human and machine capabilities compensate for each other's weaknesses. TRACE emphasizes transparent division of cognitive labor: AI contributes speed, scale, recall and formal consistency, whereas human researchers provide judgment, conceptual framing, domain understanding, epistemic responsibility and final accountability.



The bibliography includes 66 references.

Keywords: artificial intelligence, review article, literature analysis, scholarly writing, human–AI collaboration, digital co-expert, research workflow.

Contents

1. Introduction 1
2. General strategy for review preparation 3
3. AI support for review preparation 5
4. Risks of AI use at each stage 6
5. Digital co-expert rather than complete human replacement 6
6. Review preparation checklist 8
7. Case study on real review preparation 8
8. TRACE as a complementary framework to existing review-reporting systems 9
9. Conclusion 10
10. List of abbreviations 11
11. References 11

1. Introduction

Review articles play a central role in modern science because they organize fragmented knowledge, identify contradictions, define conceptual boundaries and often shape the next research agenda. Their value has increased as the scientific record has become larger, faster and more interdisciplinary. At the same time, the practical burden of preparing a high-quality review has grown substantially. Researchers must now search across multiple bibliographic databases, screen large and often

V.P.Ananikov. D.Sc., Academician of the RAS, Laboratory Head at ZIOC RAS.
E-mail: val@ioc.ac.ru
Current research interests: AI in science, digital technologies, microscopy, molecular complexity and transformations, metal complexes and nanoparticles, development of new generation of highly active nanosized and molecular catalysts, mechanistic studies of chemical reactions by experimental and theoretical methods.
D.M.Arkhipova. Ph.D., Senior Researcher at the same Institute.
E-mail: arkhipova_daria@ioc.ac.ru
Current research interests: synthetic data, electron microscopy, transition-metal catalysis, ionic liquids.
A.S.Kashin. D.Sc., Leading Researcher at the same Institute.
E-mail: a.kashin@ioc.ac.ru
Current research interests: electron microscopy, microscopy data processing, chemistry of disperse systems, transition-metal catalysis, X-ray spectroscopy.
A.P.Vorozhtsov. Research Engineer at the same Institute.
E-mail: artem2001@ioc.ac.ru
Current research interests: synthetic data generation, multimodal large language models, deep learning, AI in chemistry.

Received 13 May 2026

© 2026 Uspekhi Khimii, ZIOC RAS, Russian Academy of Sciences

redundant collections, retrieve and organize full texts, compare methods across heterogeneous studies, detect gaps and contradictions, and assemble a narrative that is simultaneously comprehensive, selective, critical and readable. Review preparation has therefore become not only an intellectual task, but also a workflow-management problem.

This environment has made AI highly attractive for scholarly work. Recent literature shows rapid growth of AI-assisted practices in academic writing, research synthesis and scientific publishing, with particular emphasis on search support, summarization, language refinement, classification and workflow acceleration.[1–5] AI-assisted literature review methods have also progressed from isolated text-generation utilities toward broader pipeline support, including screening, prioritization, evidence extraction and structured synthesis.[6–8] Several comprehensive surveys have already documented the evolution of AI applications in systematic literature review automation[9] and the broader field of automated machine learning,[10] providing a historical context that situates the emergence of workflow-oriented frameworks. These developments create real opportunities for efficiency, especially in tasks dominated by scale, repetition or formal transformation. As an example, chemistry-focused review of influential digital technologies reaches a similar conclusion at field scale: AI is already reshaping literature mining, analytical data handling, automated experimentation, synthesis planning and integrated data infrastructures, which means that review preparation increasingly occurs inside a broader digital research environment rather than as an isolated writing task.[11]

A substantial methodological ecosystem already exists for improving the transparency, reproducibility and quality of review articles. The most influential example is the PRISMA family, including PRISMA 2020 for systematic reviews and meta-analyses,[12] PRISMA-P for review protocols,[13] PRISMA-S for literature-search reporting,[14] PRISMA-ScR for scoping reviews,[15] and specialized extensions for diagnostic-test-accuracy, network-meta-analysis and individual-participant-data reviews.[16–18] Other important contributions include MOOSE for meta-analyses of observational studies,[19] ENTREQ[20] and eMERGe[21] for qualitative evidence synthesis and meta-ethnography, RAMESES for realist and meta-narrative syntheses,[22,23] SWiM for synthesis without meta-analysis,[24] SANRA for narrative review quality assessment,[25] and AMSTAR 2 (see Ref. 26) and ROBIS[27] for appraisal and risk-of-bias assessment of systematic reviews. More general methodological discussions have also clarified review typologies, the SALSA logic of Search–Appraisal–Synthesis–Analysis, and integrative-review methodology.[28–30] Recently, AI-oriented review-reporting initiatives have also appeared, including PRISMA-AI and PRISMA-trAIce, which emphasize the need to document AI either as the object of review or as a methodological tool in evidence synthesis.[31,32] These preceding systems provide important foundations for transparent and rigorous review reporting. However, the approach proposed in the present article was developed independently. Our aim is different and deliberately practical: to provide a simple, straightforward and human-controlled strategy for integrating AI into review preparation, especially at stages dominated by scale, repetition, classification, synthesis support and drafting, while preserving human responsibility for scope definition, interpretation, verification, originality and final scientific judgment.

Recent systematic evaluations of AI-assisted review preparation, synthesizing empirical studies, provide quantitative benchmarks for both opportunities and risks. It is important to note, however, that these figures are model- and task-dependent and should not be interpreted as exact universal benchmarks. In one large validation of the agentic workflow, large language model (LLM) based literature screening achieved 96.7% sensitivity and 97.9% specificity,[33] matching or exceeding human performance,[33,34] while reducing filtering times from weeks to minutes[35,36] (though broader LLM benchmarks from the same study report lower sensitivity in other settings[33]). Data extraction accuracy reached 93.1% compared to 79.7% for human reviewers.[33] In a separate hybrid two-LLM workflow with human arbitration, automation rates of 91% were achieved with error rates as low as 8%.[37] Yet the scientific value of a review article cannot be reduced to speed.[38] Reviews require interpretive judgment: deciding what matters, how findings should be compared, which contradictions are substantive rather than superficial, and what central message the literature actually supports. These are precisely the points at which AI use becomes risky.

Generative systems can omit relevant studies,[39] fabricate citations,[40] reproduce wording from published articles,[41] confuse specialized terminology,[42] compress nuance, reproduce hidden biases,[43] and produce fluent but epistemically weak text.[6,44–46] In scholarly writing, the danger is not only explicit error, but also misplaced confidence: an output that looks coherent may still be incomplete, distorted or conceptually shallow.[44,45] This caution is consistent with knowledge-centered analyses arguing that present-day AI should be treated primarily as assistive and augmentative: its main value lies in accelerating expert work, whereas unsupervised substitution for domain judgment remains premature.[47,48] Even studies demonstrating technically feasible fully automated review generation identify important limitations and reinforce the need for transparent verification and validation.[40,49] Hybrid approaches combining AI automation with structured human validation have proven most effective,[37,50,51] particularly for reference verification where AI limitations are most severe.

Recently published studies comprise empirical evaluations, systematic reviews, scoping reviews, case studies, and qualitative research.[52–54] Across these works, GPT-based models appear in 73.2% of review automation projects,[55] with frontier models achieving 97–98% accuracy in abstract screening,[37] while lightweight open-source models attain around 66% accuracy.[56] A scoping review of 37 studies found that LLMs have been tested across 10 of 13 systematic review steps, most frequently for literature search (41%) and study selection (38%), yet fully validated applications remain scarce, with only 54% of evaluations judged as promising.[50] In contrast to conventional ML tools, LLMs demonstrate superior performance in both sensitivity and the area under the curve (AUC), pointing to a more accurate and time-efficient solution for literature screening.[57] Using the tool for title and abstract screening reduced the reviewer workload to 23% of the original articles, provided that methodological quality was assured through a researcher-in-the-loop algorithm and appropriate stopping criteria.[58] Risk-of-bias assessments have shown variable agreement between AI and human reviewers.[33,59] In a separate publication-screening proof of concept, the AI tool showed higher inter-rater consistency than the human reviewers.[60] Furthermore, AI tools can provide reasoning for their decisions, suggesting that iterative prompting can sometimes induce correction of an initial classification.[61]

Generative AI is most often used to support screening, search, and writing under human oversight, yet concerns about validity,

transparency, and methodological rigor persist, as its application remains rarely embedded in a structured, methodologically guided review process-prompting a phase-specific synthesis that highlights both promise and current limitations to offer practical recommendations for responsible use.[62] Accordingly, the key question is not whether AI can participate in review preparation, but how it should properly participate. The present article addresses this question by proposing a conceptual framework for writing review articles with AI. We do not compare named software products and do not advocate any specific commercial platform. Instead, we focus on functions, workflow roles and responsibilities. Our central proposition is that AI should be used not as an autonomous substitute for scientific authorship, but as a digital co-expert: a structured cognitive partner that can expand human capacity without displacing human responsibility.

To develop this argument, we first define a general workflow for review preparation that can serve as a human baseline. We then analyze which types of AI assistance can support each stage, what risks they introduce, and why their best use lies in complementarity rather than replacement. The result is a unified framework intended for researchers, postdoctoral scholars and academic teams seeking to integrate AI into review writing while preserving rigor, criticality and scientific accountability.

We refer to the proposed approach as TRACE. Unlike existing approaches that either focus narrowly on technical performance benchmarks, offer generic checklists without workflow integration, or remain silent on documentation requirements, TRACE provides a stage-specific, documentation-centered framework that explicitly separates AI actions from human verification decisions while mandating traceability of both. The term reflects the central aim of the framework: to make AI-assisted review preparation faster, but also traceable, documented and scientifically accountable. In current practice, many researchers already use AI tools during review preparation for searching, screening, summarizing, drafting or language improvement, but not uncommonly hesitate to disclose this use because of concerns about negative editorial, reviewer or reader reactions. Such hidden use is problematic because it prevents proper evaluation of how the review was prepared and where human control was applied. We therefore argue not against AI use, but against unreported and unstructured AI use. TRACE provides a simple procedure for transparent integration of AI as a digital co-expert, with explicit documentation of AI-assisted stages, human verification, and final author responsibility for interpretation, originality and scientific judgment.

2. General strategy for review preparation

Review preparation can be represented as a structured multi-stage process that transforms an initial topic idea into a peer-review-ready manuscript. The workflow outlined here is intended as a general guide rather than a rigid prescription. It can be applied in a fully manual human workflow or adapted for AI-assisted preparation. Depending on the field, journal format, team composition and purpose of the review, stages may overlap, repeat, change order or be supplemented by additional operations. Nevertheless, the sequence provides a practical baseline because it captures the core logic of review development: from conceptual framing, through literature collection and analysis, to critical synthesis and final verification.

As illustrated in Fig. 1, review preparation is best understood not as linear writing, but as iterative knowledge engineering. The early stages determine scope and logic, the middle stages govern evidence collection and analytical organization, and the later stages transform that evidence into a coherent critical narrative. Weakness at any early stage propagates onward: an imprecise scope leads to noisy literature retrieval; poor screening affects section logic; weak analytical assignment compromises synthesis; and insufficient verification can undermine the reliability of the final manuscript.

The major stages indicated in Fig. 1 cover the whole review-development process, from topic definition to final revision and update. Although shown as a sequence, the process is iterative and several stages may be revisited during manuscript development. Here, we briefly describe review preparation stages:

Stage I. Define topic, scope and review concept

The process begins with defining what the review is actually about and why it should be prepared. This includes selecting the topic, setting temporal and thematic boundaries, choosing the review type and formulating the central message. At this stage, the authors must determine whether the review is intended to be narrative, critical, conceptual, systematic, tutorial-style, perspective-like or application-oriented. This decision affects every subsequent choice, including search breadth, section logic and standards of completeness. A well-defined scope prevents later drift and helps establish novelty relative to existing reviews.

Result: review title, topic, years covered, literature coverage, review type, central message.

Stage II. Plan review strategy and workflow

Once the concept is defined, authors should establish how the review will be produced. This includes assigning roles within the author team, selecting the target journal, considering formatting constraints, defining data organization practices and setting up manuscript version control. Search keywords and search logic should also be planned at this stage. A review is not only an intellectual product but also a managed project, and weak workflow design often creates avoidable inefficiency later. It is important to make sure that co-authors have the necessary competence and knowledge for the area selected.

Result: search keywords, co-author roles, target journal, data organization and workflow plan.

Stage III. Perform broad literature search

An important evidence-gathering step is a broad literature search across major databases and indexing platforms. The aim is not immediate precision, but comprehensive recall: to build a large bibliographic universe from which the relevant body of literature can later be selected. Search terms should be sufficiently broad to capture disciplinary variation in terminology, and multiple databases are usually required to reduce coverage bias.

Result: an extensive bibliographic database including title, abstract, authors, keywords and other information.

Stage IV. Screen and select relevant studies

The broad literature pool should then be filtered to identify studies that are truly relevant to the review's scope. This stage requires explicit or at least well-defined inclusion and exclusion criteria, such as topic and content, time range, document type, methodological rigor, relevance threshold and handling of duplicates or borderline cases. Screening is one of the most consequential stages because it determines the actual evidentiary foundation of the review.

Result: bibliographic database containing only the articles selected for inclusion in the review.

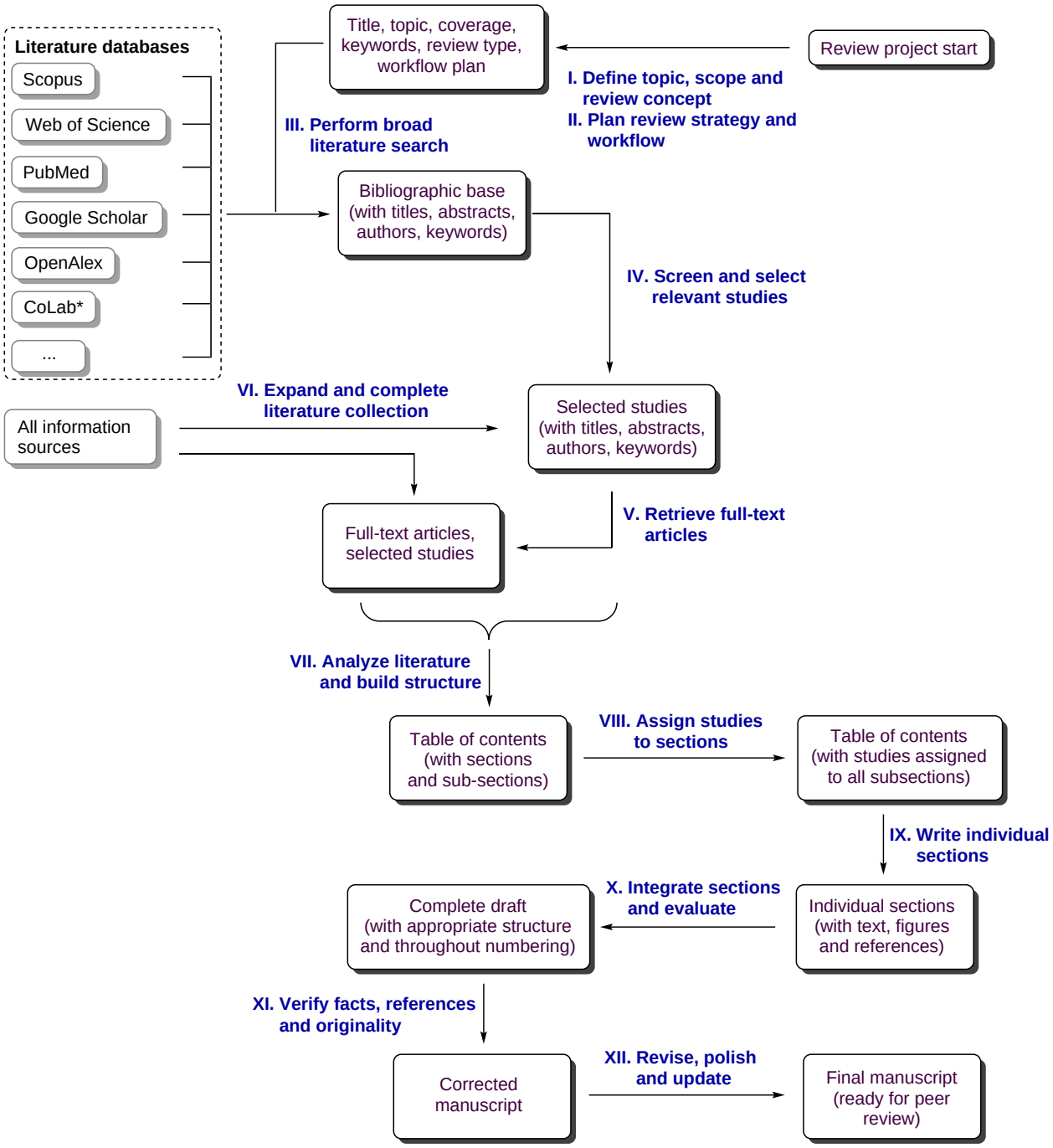

Figure 1. Schematic workflow of review preparation stages (*CoLab refers to https://colab.ws).

Stage V. Retrieve full-text articles

Selected records should be converted from bibliographic entries into analyzable documents. This requires retrieval of full texts and establishment of a practical system for storing PDFs or XML files, naming files, annotating content and linking notes to manuscript sections. The quality of the later analytical stages depends heavily on how well the full-text set is organized. It is important to note that access to articles should be obtained by the human researcher through open-access sources, institutional or personal licenses, interlibrary services or other authorized routes.

Result: full-text articles for the studies of interest in the planned review.

Stage VI. Expand and complete literature collection

A high-quality review rarely emerges from keyword search alone. After the initial literature set is assembled, authors should actively search for missing studies using reference chasing, citation networks, journal browsing, full-text search and long-horizon manual accumulation. This stage is essential for completeness because important studies are often missed by database indexing, keyword mismatch or disciplinary fragmentation.

Result: completed set of full-text articles for the studies of interest in the planned review.

Stage VII. Analyze literature and build structure

With the literature set in place, the review shifts from collection to interpretation. Authors analyze the literature to identify conceptual clusters, methodological families, unresolved debates and narrative logic. The output of this stage is a structured table of contents that reflects how the literature should be understood rather than merely how it was found.

Result: table of contents with sections and sub-sections.

Stage VIII. Assign studies to sections

Each study should then be assigned to one or more sections according to its intellectual role in the manuscript. This stage helps translate the literature set into a section-specific evidence map and supports the planning of figures, comparative tables, taxonomies and conceptual diagrams. Many studies may belong to multiple sections, especially if they contain both methodological and conceptual contributions.

Result: table of contents with studies assigned to each section and subsection.

Stage IX. Write individual sections

The manuscript is then drafted section by section. At this point, each section should become a self-contained analytical unit with integrated references, figures, tables and author interpretation. Importantly, writing a review is not equivalent to summarizing papers sequentially. The purpose is to synthesize, compare, evaluate and interpret. Critical analysis should therefore be embedded directly into the writing process.

Result: individual sections with text, figures and references.

Stage X. Integrate sections and evaluate

After individual sections are drafted, they should be integrated into a coherent full manuscript. This includes harmonizing style, narrative flow, numbering, terminology and internal cross-references. It is also the stage at which authors should evaluate whether the review actually delivers the intended message and whether earlier decisions about scope, evidence or structure need revision.

Result: complete draft with appropriate structure and consistent numbering throughout.

Stage XI. Verify facts, references and originality

Before finalization, the manuscript should undergo rigorous verification. Facts, references, figure attributions, paraphrasing quality and internal consistency should all be checked carefully. Plagiarism and attribution checks should be performed. Permissions for reproduced figures should be obtained where required, the attribution for openly licensed material should be provided. This stage is especially important for ensuring that claims are traceable to source literature and that the review maintains both accuracy and originality.

Result: corrected manuscript.

Stage XII. Revise, polish and update

The final stage is refinement. Authors should improve readability, sharpen phrasing, remove redundancy and ensure that the review communicates efficiently to its intended audience. A final literature update is also necessary because new papers often appear during the writing process. Only after this last pass does the manuscript become submission-ready.

Result: final manuscript ready for peer review.

3. AI support for review preparation

AI and digital tools can support nearly every stage of review preparation, but their usefulness depends on being assigned to the right task. In this article we intentionally avoid discussing named software products or commercial brands. The purpose is not to compare platforms or advertise particular tools, but to identify functional classes of assistance that are increasingly available across scholarly ecosystems. In practice, these functional classes can be realized using a single class of AI systems — LLM-based agentic systems — whose behavior is steered by prompting and, where needed, augmented with external tools. The same system can perform retrieval, summarization, parsing, verification or language polishing, depending on how the user formulates the instruction. Thus, we do not prescribe specific software; instead, we describe prompt-defined roles that can be instantiated with any accessible LLM. These functional classes range from retrieval-oriented algorithms to generative language models, and from organizational aids to analytical and editorial support. Their value lies in augmenting the workflow, not defining the scholarship.

Across the 12 stages described above, AI contributes most effectively in tasks that involve scale, repetition, prioritization, pattern detection, language transformation or structured reformatting. Examples include expanding search queries, clustering large bibliographic collections, prioritizing likely-relevant records, extracting metadata, grouping studies by theme, generating provisional outlines, summarizing section-specific datasets, improving linguistic clarity and checking for formal inconsistencies. These capabilities can significantly reduce time spent on labor-intensive operations and allow researchers to reallocate effort toward interpretation, critique and conceptual design.[6,7] A caveat is that the initial prompt reflects the researcher's current understanding, which may inadvertently narrow the search space and miss serendipitous associations that manual browsing would reveal. To mitigate this bias, stage VI involves human-driven expansion of the literature set *via* citation tracing and journal browsing, performed iteratively. However, the risk of overlooking unexpected connections persists and should be acknowledged. A useful classification separates AI-Assistant, AI-Analyst and AI-Researcher levels, thereby distinguishing routine language- and workflow-support tasks from custom data-analysis pipelines and, further, from more ambitious systems aimed at hypothesis generation and strategic design.[47] This distinction helps clarify why different review stages require different expectations from AI tools.

Table 1 shows that the most productive role of AI is not isolated text generation, but distributed workflow assistance. Particularly valuable contributions emerge in literature discovery, relevance prioritization, clustering, section-wise organization, draft acceleration and formal consistency control. However, the table also makes clear that every stage retains a human control point. This is not a conservative add-on; it is structurally necessary. Scientific reviews are not only information-processing exercises but acts of judgment, and judgment cannot be outsourced without degrading interpretive quality.[6,8] The broader literature points in the same direction: the current digital transformation is driven not by one universal AI function but by a portfolio of capabilities, from natural language processing (NLP) and chemoinformatics to high-throughput experimentation and digital twins, and their value depends on disciplined integration into human research workflows.[11]

Table 1. Functional roles of AI at each stage of review preparation. The table is organized by workflow stage and focuses on functions.

Stage	AI function	Potential support	Human control point
I. Define topic, scope and review concept	Topic-scoping, similarity search, trend mapping	Survey adjacent topics, detect overlapping reviews, map emerging themes	Approve scope, define novelty, decide review type and message
II. Plan review strategy and workflow	Workflow planning, keyword expansion, taxonomy suggestion	Propose search terms, workflow structures and role distributions	Validate terminology and approve operational plan
III. Perform broad literature search	Semantic search, query expansion, deduplication	Broaden retrieval and merge outputs from multiple sources	Audit coverage and ensure recall
IV. Screen and select relevant studies	Screening prioritization, relevance ranking, abstract triage	Accelerate first-pass selection and identify likely-inclusion records	Make final inclusion or exclusion decisions
V. Retrieve full-text articles	Metadata extraction, file organization, document parsing	Rename, sort, annotate and structure full-text collections	Check source integrity and text corpus completeness
VI. Expand and complete literature collection	Citation-network analysis, gap detection	Identify likely missing studies and undercovered subfields	Judge relevance and completeness threshold
VII. Analyze literature and build structure	Topic modeling, thematic clustering, outline drafting	Transform large datasets into candidate structures and subsection candidates	Construct the intellectual logic
VIII. Assign studies to sections	Section-wise tagging, evidence mapping	Assign studies to likely sections and detect cross-cutting papers	Decide where each study matters
IX. Write individual sections	Summarization, comparison support, draft assistance	Generate section drafts and comparative matrices	Interpret evidence and preserve nuance
X. Integrate sections and evaluate	Consistency checking, transition drafting	Improve coherence and align terminology	Decide narrative order and emphasis
XI. Verify facts, references and originality	Reference checking, claim-to-source matching	Detect citation mismatches and unsupported claims	Perform factual verification and attribution control
XII. Revise, polish and update	Language polishing, update-search support	Refine wording and update recent literature	Approve final text and take submission responsibility

4. Risks of AI use at each stage

The same properties that make AI useful also could make it risky, especially when researchers are unaware of these limitations. Systems that accelerate retrieval can amplify blind spots; systems that summarize can erase nuance; systems that generate fluent text can conceal unsupported claims. Therefore, responsible AI use in review writing requires stage-specific risk awareness in addition to generic caution. The key question is not whether a given tool is 'good' or 'bad', but what kind of error it tends to introduce at a specific point in the workflow.[8]

Among the risks described in Table 2, citation hallucination and unsupported synthesis are particularly consequential because they directly undermine scientific trust. The table emphasizes that AI-related failure modes differ across workflow stages and therefore require targeted control. Recent discussions in scholarly publishing show that fabricated or corrupted references are not rare edge cases but a growing integrity problem in AI-assisted writing.[63] More subtle, but equally important, is the risk of interpretive flattening: AI often produces consensus-shaped summaries that suppress disagreement, ambiguity or contradictory opinions. This is especially problematic in review articles, where the purpose is often not to erase disagreement but to explain it. This problem also has a practical dimension: discussions from chemical research emphasize that the competitive advantage belongs to researchers who learn to use AI critically and productively, not to those who delegate scientific responsibility to it.[48] False trust in machine-generated authority therefore becomes both an epistemic and a professional risk.

A practical implication is as follows: AI should be used under a control-loop logic rather than a one-pass generation logic (Fig. 2). Outputs should be treated as provisional analytical material that must be checked against source literature, reinterpreted by the authors and either accepted, revised or discarded. This turns AI from an authority into an instrument. For reviews that are predominantly conceptual — where the core value lies in comparing theoretical frameworks, methodological paradigms, or competing interpretations — the risk of flattening is especially acute. Generative models tend to harmonize terminology, smooth contradictions, and reduce competing approaches to an averaged schema. To counter this, we recommend a stronger human presence at stages involving critical interpretation and synthesis. These measures help preserve intellectual diversity and depth, which are the hallmarks of a high-quality conceptual review.

5. Digital co-expert rather than complete human replacement

The analysis above leads to a clear conclusion: AI can be highly useful in review preparation, but only if its role is defined correctly. A fully automated 'press-button review' is not an appropriate model for scholarly analysis. The problem is not simply that AI sometimes makes mistakes. The deeper issue is that review writing is not a single task but a chain of heterogeneous cognitive operations, some of which are well-suited to automation and some of which are fundamentally dependent on human interpretation, contextual judgment and responsibility.

We therefore propose the digital co-expert approach as a conceptual framework for AI-assisted review writing. In this framework, AI is not treated as an autonomous author, nor merely as a passive editing utility, but as a structured analytical partner embedded into a human-led workflow. The term 'digital

Table 2. Stage-specific risks introduced by AI use in review preparation.

Stage	Typical AI-related risk[a]	How the risk appears	Minimum mitigation
I. Define topic, scope and review concept	Topic misframing	Overly generic, fashionable or derivative topic selection	Manual novelty check against existing reviews (AI-retrieved, human-verified)
II. Plan review strategy and workflow	Workflow overconfidence	Premature trust in automated planning or keyword sets	Human review of strategy and search logic
III. Perform broad literature search	Retrieval bias	Missing relevant studies due to keyword or ranking bias	Use multiple databases and manual search expansion
IV. Screen and select relevant studies	False screening confidence	Incorrect exclusion of relevant papers or inclusion of irrelevant ones	Human audit of borderline and excluded records
V. Retrieve full-text articles	Literature set corruption	Broken metadata, wrong file linking, missing or duplicate full texts	Spot-check file integrity and citation matching
VI. Expand and complete literature collection	Completeness illusion	Belief that AI-found literature is complete when important work is absent	Backward or forward citation chasing and manual gap analysis
VII. Analyze literature and build structure	Superficial clustering	Topically neat but conceptually weak structure generation	Rebuild structure around scientific questions, not lexical similarity
VIII. Assign studies to sections	Misassignment of evidence	Papers assigned to sections based on keywords rather than meaning	Human section-level evidence review
IX. Write individual sections	Fluent but shallow drafting	Text that sounds polished but lacks critique, balance or source fidelity	Line-by-line verification against source literature
X. Integrate sections and evaluate	Narrative homogenization	Loss of author voice, over-standardized transitions, artificial smoothness	Human editorial pass for argument strength and style
XI. Verify facts, references and originality	Citation hallucination and attribution error	Invented references, incorrect metadata, unsupported claims	Direct reference verification and claim tracing
XII. Revise, polish and update	Polished unreliability	Final text appears publication-ready despite hidden factual weaknesses	Final human read-through and independent verification

[a] Retrieval bias is a systematic distortion in search results due to query or tool limitations; false screening confidence is unjustified trust in AI-selected records; superficial clustering is shallow thematic grouping without conceptual depth; misassignment of evidence is incorrect categorization of papers by topic or relevance; fluent but shallow drafting is grammatically correct but analytically weak text; citation hallucination is non-existent or fabricated references; unsupported claims are statements not grounded in cited sources; polished unreliability is convincing but inaccurate output.

co-expert' is used here to denote a system that can contribute meaningfully to expert work without itself constituting the expert authority. Its value lies in complementarity: AI expands human processing capacity, while the human author preserves epistemic direction, conceptual judgment and final accountability. A closely related chemistry-oriented conceptualization describes the evolution from AI as a routine assistant to AI as a deeper analytical participant in scientific work, and even introduces the idea of 'technology-as-participant' to capture this shift.[47] Our digital co-expert model may be

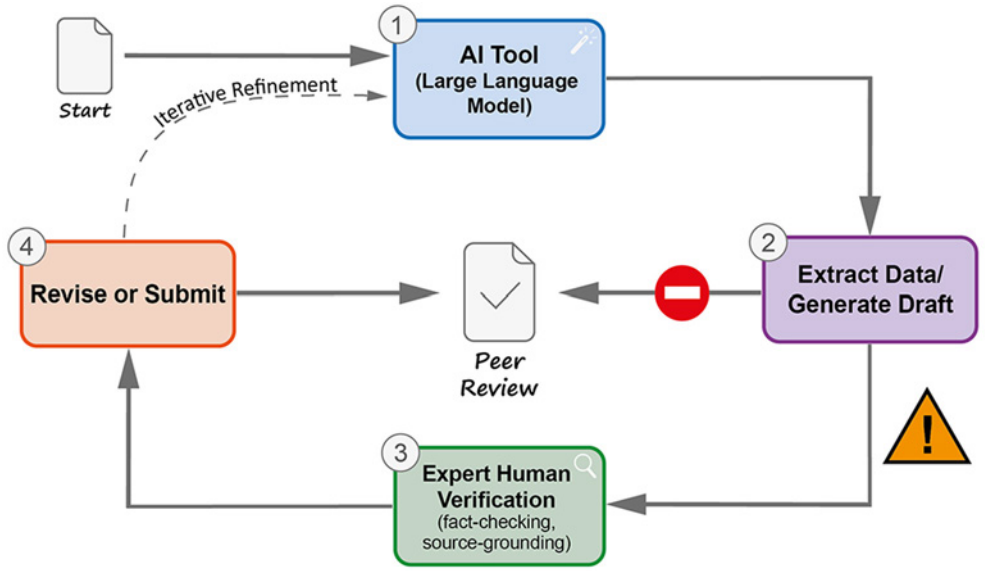

Figure 2. Control loop for responsible AI-assisted review preparation: AI outputs should pass through iterative human verification, source-grounding and revision before being integrated into the manuscript (the figure emphasizes process control rather than one-step automation).

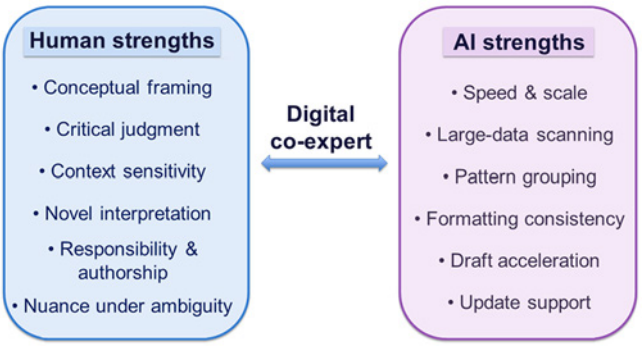

Figure 3. Complementarity of human and AI capabilities in review writing: the digital co-expert framework emerges from complementarity — AI contributes speed, scale and structural assistance, whereas humans provide interpretation, critical judgment and accountability.

viewed as a review-writing-specific formulation of the same broader movement toward structured human-machine collaboration.

This framing is important because the strongest and weakest points of human and AI performance are not the same (Fig. 3). Human researchers are comparatively slow in repetitive large-volume operations such as bulk screening, metadata organization, section-wise tagging, language compression and repeated formatting. Performance can also degrade when attention should be sustained across large amounts of monotonous material. AI, by contrast, is efficient in high-volume scanning, pattern grouping, consistency support and provisional drafting. However, AI is weak in exactly the areas where scholarly reviews become scientifically meaningful: deciding what is conceptually important, distinguishing superficial similarity from substantive relevance, identifying when disagreement matters, recognizing methodological asymmetry, and generating genuinely original framing rather than statistically probable text.[6–8]

The digital co-expert framework therefore rests on division of cognitive labor. AI should carry the burden of scale-sensitive, repetitive and formally transformable operations. Human researchers should retain ownership of scope, inclusion logic, interpretation, critique, synthesis, narrative argument and all final decisions affecting scientific meaning. This is not simply a compromise position between enthusiasm and skepticism. It is a structurally stronger model because it assigns each actor to the class of work it performs best. This division is also consistent with broader analyses of digital chemistry, where AI contributes through scalable handling of data-rich and formally structured tasks, while scientific direction remains anchored in human expertise, interpretation and responsibility.[11,47]

An important consequence is that effective AI use does not reduce the need for expertise; in many cases, it increases it. Poorly supervised AI can generate outputs that are polished enough to evade casual detection but weak enough to distort the review. Only a sufficiently knowledgeable researcher can recognize when an apparently plausible summary has omitted a crucial debate, merged incompatible concepts or smoothed over an unresolved contradiction. In that sense, AI assistance is most powerful in the hands of experts rather than in place of them.

At the same time, the digital co-expert model is not anti-automation. On the contrary, it explicitly embraces digital augmentation where it improves scholarly efficiency and breadth. What it rejects is category error: the mistaken assumption that because AI can generate review-like text, it can therefore replace the scientific act of reviewing. A review article is not merely a descriptive document; it is a dedicated research project. That argument requires authorship in the full intellectual sense.

6. Review preparation checklist

To make the proposed workflow practically usable and transparent, we suggest accompanying the review preparation process with a structured review preparation checklist.[64] We have set up a dedicated webpage on our website for the TRACE framework,[64] where we provide the latest version of the checklist and supplementary materials. This page will be updated as the framework evolves. The purpose of this checklist is explicit documentation of how the review was designed, executed, controlled and finalized. In many review articles, readers can evaluate the final narrative but cannot easily understand how the literature collection was assembled, how scope boundaries were set, how screening decisions were made, or how completeness and reliability were ensured. A structured checklist helps expose this hidden methodological layer of review preparation and thereby strengthens confidence in the resulting article.

The checklist is organized according to the 12-stage workflow described above and is intended to be filled in by the authors during manuscript development. It captures practical information such as review scope, years covered, search keywords, databases used, number of records retrieved, deduplication and screening outcomes, full-text coverage, section assignment strategy, verification steps, and final update procedures. In this way, the checklist functions as a compact operational record of the review-building process. It may also help co-authors coordinate their work more efficiently by making explicit which decisions were taken, which stages were completed, and where additional effort or verification may still be needed.

Such planning and self-control are important for any review article because review quality depends not only on writing skill, but also on the rigor of the upstream evidence workflow. A review may appear polished while still being incomplete, weakly structured or biased by untracked selection decisions. The checklist therefore acts as a form of methodological self-audit. It encourages the authors to think in advance about completeness, traceability, inclusion logic and reproducibility of the review process. Even in narrative or conceptual reviews, where full formal systematic-review protocols may not be appropriate, explicit internal control remains highly valuable for maintaining scientific reliability. This becomes even more important in the context of the digital co-expert approach.

For these reasons, we recommend that the completed review preparation checklist should increasingly become a normal companion document for review articles, especially when digital and AI-assisted methods are used. Its inclusion as Supplementary Information can improve transparency for editors, reviewers and readers, while also helping authors themselves produce stronger, more reliable and better-documented reviews.

7. Case study on real review preparation

The workflow proposed in this article was evaluated in practice during preparation of the review article entitled ‘Synthetic data in electron microscopy. AI co-expert-assisted review’.[65] This review contained 198 references, 30 figures, 6 tables and 56 pages (without references), providing a realistic test case for estimating the effect of the digital co-expert approach on review

preparation time. The time intervals estimated for each stage are summarized in Table 3.

The results show that AI acceleration was not uniform across the workflow. The digital co-expert was applied selectively, mainly to stages where large-volume processing, classification, structuring or drafting created the greatest time burden. These stages included screening and selection of relevant studies, analysis of literature and structure building, assignment of studies to sections, drafting of individual sections, and final revision and polishing. In these AI-assisted stages, the acceleration was substantial: screening was accelerated by 10 times, literature analysis and structure building by 10 times, study assignment by 10 times, section writing by 22 times, and revision and polishing by 3 times. The time-weighted average acceleration across these five AI-assisted stages was approximately 8.8 times.

Table 3. Comparison of time intervals that would take the review written by a human (without AI use) and with AI-co-expert approach (human+AI) for the preparation of 'Synthetic data in electron microscopy. AI co-expert-assisted review' article.[65]

Stage	Time needed to complete		Estimated acceleration (times)
	Human only[a]	With digital co-expert[b]	
I. Define topic, scope and review concept	8 h (*1 w.day*[c])	8 h (*1 w.day*) [AI not used]	1
II. Plan review strategy and workflow	8 h (*1 w.day*)	8 h (*1 w.day*) [AI not used]	1
III. Perform broad literature search	40 h (*1 week*[d])	40 h (*1 week*) [AI not used]	1
IV. Screen and select relevant studies	80 h (*2 weeks*)	8 h (*1 w.day*) [AI used]	10
V. Retrieve full-text articles	8 h (*1 w.day*)	8 h (*1 w.day*) [AI not used]	1
VI. Expand and complete literature collection	24 h (*3 w.days*)	24 h (*3 w.days*) [AI not used]	1
VII. Analyze literature and build structure	80 h (*2 weeks*)	8 h (*1 w.day*) [AI used]	10
VIII. Assign studies to sections	40 h (*1 week*)	4 h [AI used]	10
IX. Write individual sections	352 h (*2 months*[e])	16 h (*2 w.days*) [AI used]	22
X. Integrate sections and evaluate	40 h (*1 week*)	40 h (*1 week*) [AI not used]	1
XI. Verify facts, references and originality	40 h (*1 week*)	40 h (*1 week*) [AI not used]	1
XII. Revise, polish and update	120 h (*3 weeks*)	40 h (*1 week*) [AI used]	3
Total	**840 h (*21 weeks*)**	**244 h (*~6 weeks*)**	**3.4**

[a] Estimated based on experience of preparing similar review articles. [b] Real time taken for the preparation of the review noted. Includes AI usage time and human co-work or verification (*i.e.*, overall time AI+human). [c] w.day is a working day. [d] 1 week equals 5 w.days. [e] 1 month equals 22 w.days.

At the level of the complete review-preparation workflow, the acceleration was lower but still considerable. The estimated human-only preparation time was 840 h, whereas the AI co-expert-assisted workflow required 244 h. Thus, the overall estimated acceleration was approximately 3.44 times. This difference between stage-specific and overall acceleration is important. It shows that AI produced very strong acceleration where it was used, but the full workflow still included several human-only stages that were not shortened. These stages included topic definition, workflow planning, broad literature search, full-text retrieval, expansion of the literature set, manuscript integration, and verification of facts, references and originality.

The human-only stages should not be interpreted as inefficient or unnecessary. On the contrary, they remained essential for scientific quality, responsibility and reliability. Conceptual framing, search strategy, completeness control, final integration and verification are critical stages where human expertise, judgment and accountability are especially important. The present case study therefore supports the central idea of the digital co-expert approach: AI should not replace the researcher, but should accelerate the most time-consuming and scale-sensitive parts of the workflow while leaving key scientific decisions under human control.

Further acceleration is likely possible if AI is integrated more deeply into additional stages, for example through more automated literature retrieval, stronger reference-verification systems, improved evidence mapping and better manuscript-level consistency checking. However, such integration should be accompanied by stronger validation procedures. The time intervals in Table 3 should be regarded as case-specific estimates rather than universal constants because the human-only baseline is derived from expert judgment and the analysis is based on a single case. Actual time will vary with the specific LLM, review topic, team size, author experience, and literature volume. For example, larger corpora or less capable models may increase manual verification overhead, whereas well-crafted prompting can shorten routine tasks. At the same time, deep conceptual work (framing, synthesis, critical appraisal) remains inherently time-consuming and less compressible, so the relative gain will differ across projects.

A natural question is whether the estimated acceleration comes at the cost of reduced scientific quality. In this work, we did not conduct a controlled experiment to measure the quality of the AI-assisted review against a conventional counterpart. However, we performed an informal internal evaluation. The experts judged the AI-assisted version to be at least comparable in coverage and clarity, while the conceptual framing remained author-driven.

8. TRACE as a complementary framework to existing review-reporting systems

As TRACE provides a stage-based logic for organizing review preparation, it could be viewed as complementary to existing review-preparation, review-reporting and review-appraisal systems. The PRISMA family and its extensions have made a major contribution to transparency and reproducibility in systematic reviews, protocols, search reporting, scoping reviews and specialized evidence-synthesis formats.[12–18] Other established systems, including MOOSE, ENTREQ, eMERGe,

RAMESES, SWiM, SANRA, AMSTAR 2 and ROBIS have further strengthened reporting, synthesis quality and critical appraisal across observational meta-analyses, qualitative syntheses, realist and meta-narrative reviews, narrative reviews and systematic-review evaluation.[19–27] More recently, AI-oriented initiatives such as PRISMA-AI and PRISMA-trAIce have begun to address the need for transparent reporting of AI as a methodological tool, rather than solely as a subject of research — an area not fully covered by previous guidance.[31,32] In parallel, the RAISE (Responsible use of AI in evidence SynthEsis) recommendations — a three-part framework developed by Cochrane, the Campbell Collaboration, JBI and the Collaboration for Environmental Evidence — provide complementary guidance on the responsible, transparent and ethically sound use of AI across the evidence-synthesis ecosystem, covering practice recommendations (RAISE 1), tool development and evaluation (RAISE 2), and tool selection and use (RAISE 3).[66] These systems provide important foundations and should continue to be used whenever they are appropriate for the review type.

TRACE has a different organizing principle. Most existing systems are organized primarily around what should be reported in the final article, *viz*., title, abstract, eligibility criteria, information sources, search strategy, selection process, data collection, risk of bias, synthesis, results, discussion and transparency items. TRACE is organized around the actual work of preparing a review manuscript: from concept definition and workflow planning through literature collection, structure building, assignment of studies to sections, drafting, integration, verification, revision and final update. Thus, TRACE addresses a different object. It is not only a reporting checklist for a completed review, but a workflow concept for how the review is built, controlled and documented during preparation.

An important issue is the stage-wise AI architecture. TRACE does not simply recommend that AI use should be disclosed. Instead, it maps AI functions, possible contributions, human control points, stage-specific risks and minimum mitigation measures across all 12 stages of review preparation. This makes the framework broader than search-reporting tools and more operational for manuscript preparation than AI-disclosure frameworks alone. Its purpose is to help authors decide where AI is useful, where it is risky, and where human verification must remain decisive.

A distinguishing feature is the digital co-expert concept. In TRACE, AI is framed neither as an autonomous author nor as a passive language-editing utility. It is treated as a structured analytical partner embedded in a human-led workflow. This division of cognitive labor is central: AI contributes speed, scale, recall, clustering, drafting support and formal consistency, whereas human researchers retain responsibility for scope, relevance, interpretation, critique, synthesis, originality, verification and final scientific judgment. This human-AI co-expert logic is not the organizing principle of other reporting approaches.

TRACE includes practical evaluation of acceleration. In the case study presented above, estimated speedups were up to 22-fold for AI-assisted stages and approximately 3.44-fold for the complete review-preparation workflow. Such timing analysis is usually outside the scope of conventional reporting guidelines.

It is important to note that TRACE is intended as a preparation-oriented workflow and transparency aid for AI-assisted review preparation; it is not a substitute for PRISMA family or other reporting standards when the review type requires them.

9. Conclusion

Review preparation has become not only an intellectual task but also an increasingly complex workflow-management problem — one that is more demanding than ever before. In this article, we proposed TRACE — Transparent Review Acceleration with Co-Expert AI — as an approach for integrating AI into review writing without replacing human scientific responsibility. TRACE is based on the idea that AI should function as a digital co-expert: a structured partner that accelerates selected operations, while human authors retain control over scope, interpretation, verification, originality and final scientific judgment.

The article first formulated a general 12-stage workflow for review preparation, covering the complete path from topic definition, strategy planning and literature search to screening, full-text collection, structure building, drafting, integration, verification, revision and final updating. This workflow provides a common baseline for both manual and AI-assisted review preparation. We then described the types of support that AI can provide at each stage, including semantic search, query expansion, deduplication, screening prioritization, thematic clustering, evidence mapping, section drafting, consistency checking and language polishing. This analysis shows that the greatest value of AI lies not in isolated text generation, but in distributed workflow acceleration.

At the same time, the TRACE approach recognizes that AI-assisted review preparation introduces specific risks. These include retrieval bias, false screening confidence, superficial clustering, misassignment of studies, overcompression of nuance, fluent but shallow drafting, citation hallucination, unsupported claims and polished unreliability. Therefore, the proposed framework emphasizes a control-loop logic: AI outputs should be treated as a provisional material that must be checked, revised, rejected or integrated by human experts.

A central practical element of TRACE is transparency. Many researchers already use AI tools during review preparation, but may hesitate to disclose this use because of possible negative reactions from editors, reviewers or readers. Hidden AI use is problematic because it prevents proper evaluation of how literature was selected, summarized, transformed and verified. TRACE therefore argues for structured and documented AI use rather than informal, invisible assistance. We provide checklist[64] that supports this aim by recording the review scope, search strategy, screening decisions, literature coverage, AI-assisted stages, verification procedures and final update steps. In this way, the checklist converts AI-assisted review writing from an undocumented practice into a traceable and auditable workflow.

The practical evaluation of the approach was performed during preparation of the exemplary review. The case study showed that AI was most useful when applied selectively to the most time-consuming stages: screening and selecting relevant studies, analyzing literature and building structure, assigning studies to sections, writing individual sections, and revising/polishing the manuscript. For these AI-assisted stages, the time-weighted average acceleration was 8.8 times. At the level of the complete review-preparation workflow, the estimated human-only duration was 840 hours, whereas the TRACE-assisted workflow required 244 hours, corresponding to an overall acceleration of approximately 3.44 times.

This difference between stage-specific and overall acceleration is important. It shows that TRACE does not simply maximize speed at every point. Several stages remained human-only, including topic definition, workflow planning, broad

literature search, full-text retrieval, expansion of the literature collection, manuscript integration and final verification. These stages are essential safeguards, not inefficiencies. They preserve conceptual framing, literature completeness, interpretive coherence, factual reliability and author accountability. Thus, the practical case study supports the main principle of TRACE: AI should accelerate the stages dominated by scale, repetition and formal transformation, while human experts retain responsibility for the stages where scientific meaning is defined and validated.

Further acceleration is certainly possible with deeper AI integration into literature retrieval, evidence mapping, full-text comparison, claim-to-source verification, figure and table support, and manuscript-level consistency control. However, such development should proceed together with stronger validation, better source-grounding, transparent documentation and clear disclosure of AI use. The time intervals reported in the case study should also be interpreted as estimates rather than universal constants, because acceleration will depend on topic, literature volume, review type, author experience, tool availability and workflow strategy. Nevertheless, the results indicate that TRACE can provide substantial estimated efficiency gains while maintaining the essential qualities of scholarly review writing: rigor, critical judgment, transparency and responsibility.

Methodology

AI tools were used for the preparation of the present article. AI tools were used for preliminary analysis and processing, to polish the language and to improve the writing; all content, analysis, and conclusions remain the responsibility of the authors.

Funding

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (Agreement No. 075-15-2024-531).

10. List of abbreviations

AI — artificial intelligence;
GPT — Generative Pre-trained Transformer;
LLM — large language model;
NLP — natural language processing;
RAISE — Responsible use of AI in evidence SynthEsis;
TRACE — Transparent Review Acceleration with Co-Expert AI.

11. References

1. M.Khalifa, M.Albadawy. *Comput. Methods Programs Biomed. Update*, **5**, 100145 (2024); https://doi.org/10.1016/j.cmpbup.2024.100145
2. J.Clark, B.Barton, L.Albarqouni, O.Byambasuren, T.Jowsey, J.Keogh, T.Liang, C.Moro, H.M.O'Neill, M.Jones. *Res. Synth. Methods*, **16**, 601 (2025); https://doi.org/10.1017/rsm.2025.16
3. S.Harasgama, H.Pearce, C.Appel, L.Loftus, H.Painter, I.Kuhn, J.Karpusheff, A.Ceesay, J.Ford. *J. Med. Internet Res.*, **28**, e81597 (2026); https://doi.org/10.2196/81597
4. A.Cheng, A.W.Calhoun, G.Reedy. *Adv. Simul.*, **10**, 22 (2025); https://doi.org/10.1186/s41077-025-00350-6
5. C.Ganjavi, M.B.Eppler, A.Pekcan, B.Biedermann, A.L.Abreu, G.S.Collins, I.S.Gill, G.E.Cacciamani. *BMJ*, **384**, e077192 (2024); https://doi.org/10.1136/bmj-2023-077192
6. A.Adel, N.Alani. *AI & Soc.*, **40**, 6799 (2025); https://doi.org/10.1007/s00146-025-02406-7
7. A.Asai, J.He, R.Shao, W.Shi, A.Singh, J.C.Chang, K.Lo, L.Soldaini, S.Feldman, M.D'Arcy, D.Wadden, M.Latzke, J.Sparks, J.D.Hwang, V.Kishore, M.Tian, P.Ji, S.Liu, H.Tong, B.Wu, Y.Xiong, L.Zettlemoyer, G.Neubig, D.S.Weld, D.Downey, W.-t.Yih, P.W.Koh, H.Hajishirzi. *Nature*, **650**, 857 (2026); https://doi.org/10.1038/s41586-025-10072-4
8. L.Huang, W.Yu, W.Ma, W.Zhong, Z.Feng, H.Wang, Q.Chen, W.Peng, X.Feng, B.Qin, T.Liu. *ACM Trans. Inf. Syst.*, **43**, 42 (2025); https://doi.org/10.1145/3703155
9. J. de la Torre-López, A.Ramírez, J.R.Romero. *Computing*, **105**, 2171 (2023); https://doi.org/10.1007/s00607-023-01181-x
10. R.Barbudo, S.Ventura, J.R.Romero. *Knowl. Inf. Syst.*, **65**, 5097 (2023); https://doi.org/10.1007/s10115-023-01935-1
11. V.P.Ananikov. *Artif. Intell. Chem.*, **2**, 100075 (2024); https://doi.org/10.1016/j.aichem.2024.100075
12. M.J.Page, J.E.McKenzie, P.M.Bossuyt, I.Boutron, T.C.Hoffmann, C.D.Mulrow, L.Shamseer, J.M.Tetzlaff, E.A.Akl, S.E.Brennan, R.Chou, J.Glanville, J.M.Grimshaw, A.Hróbjartsson, M.M.Lalu, T.Li, E.W.Loder, E.Mayo-Wilson, S.McDonald, L.A.McGuinness, L.A.Stewart, J.Thomas, A.C.Tricco, V.A.Welch, P.Whiting, D.Moher. *BMJ*, **372**, n71 (2021); https://doi.org/10.1136/bmj.n71
13. D.Moher, L.Shamseer, M.Clarke, D.Ghersi, A.Liberati, M.Petticrew, P.Shekelle, L.A.Stewart. *Syst. Rev.*, **4**, 1 (2015); https://doi.org/10.1186/2046-4053-4-1
14. M.L.Rethlefsen, S.Kirtley, S.Waffenschmidt, A.P.Ayala, D.Moher, M.J.Page, J.B.Koffel. *Syst. Rev.*, **10**, 39 (2021); https://doi.org/10.1186/s13643-020-01542-z
15. A.C.Tricco, E Lillie, W.Zarin, K.K.O'Brien, H.Colquhoun, D.Levac, D.Moher, M.D.J.Peters, T.Horsley, L.Weeks, S.Hempel, E.A.Akl, C.Chang, J.McGowan, L.Stewart, L.Hartling, A.Aldcroft, M.G.Wilson, C.Garritty, S.Lewin, C.M.Godfrey, M.T.Macdonald, É.V.Langlois, K.Soares-Weiser, J.Moriarty, T.Clifford, Ö.Tunçalp, S.E.Straus. *Ann. Intern. Med.*, **169**, 467 (2018); https://doi.org/10.7326/m18-0850
16. M.D.F.McInnes, D.Moher, B.D.Thombs, T.A.McGrath, P.M.Bossuyt, T.Clifford, J.F.Cohen, J.J.Deeks, C.Gatsonis, L.Hooft, H.A.Hunt, C.J.Hyde, D.A.Korevaar, M.M.G.Leeflang, P.Macaskill, J.B.Reitsma, R.Rodin, A.W.S.Rutjes, J.P.Salameh, A.Stevens, Y.Takwoingi, M.Tonelli, L.Weeks, P.Whiting, B.H.Willis. *JAMA*, **319**, 388 (2018); https://doi.org/10.1001/jama.2017.19163
17. B.Hutton, G.Salanti, D.M.Caldwell, A.Chaimani, C.H.Schmid, C.Cameron, J.P.A.Ioannidis, S.E.Straus, K.Thorlund, J.P.Jansen, C.D.Mulrow, F.Catalá-López, P.C.Gøtzsche, K.Dickersin, I.Boutron, D.G.Altman, D.Moher. *Ann. Intern. Med.*, **162**, 777 (2015); https://doi.org/10.7326/m14-2385
18. L.A.Stewart, M.Clarke, M.M.Rovers, R.D.Riley, M.Simmonds, G.Stewart, J.F.Tierney. *JAMA*, **313**, 1657 (2015); https://doi.org/10.1001/jama.2015.3656
19. D.F.Stroup, J.A.Berlin, S.C.Morton, I.Olkin, G.D.Williamson, D.Rennie, D.Moher, B.J.Becker, T.A.Sipe, S.B.Thacker. *JAMA*, **283**, 2008 (2000); https://doi.org/10.1001/jama.283.15.2008
20. A.Tong, K.Flemming, E.McInnes, S.Oliver, J.C.Craig. *BMC Med. Res. Methodol.*, **12**, 181 (2012); https://doi.org/10.1186/1471-2288-12-181
21. E.F.France, M.Cunningham, N.Ring, I.Uny, E.A.S.Duncan, R.G.Jepson, M.Maxwell, R.J.Roberts, R.L.Turley, A.Booth, N.Britten, K.Flemming, I.Gallagher, R.Garside, K.Hannes, S.Lewin, G.W.Noblit, C.Pope, J.Thomas, M.Vanstone, G.M.A.Higginbottom, J.Noyes. *J. Adv. Nurs.*, **75**, 1126 (2019); https://doi.org/10.1111/jan.13809
22. G.Wong, T.Greenhalgh, G.Westhorp, J.Buckingham, R.Pawson. *BMC Med.*, **11**, 21 (2013); https://doi.org/10.1186/1741-7015-11-21
23. G.Wong, T.Greenhalgh, G.Westhorp, J.Buckingham, R.Pawson. *J. Adv. Nurs.*, **69**, 987 (2013); https://doi.org/10.1111/jan.12092
24. M.Campbell, J.E.McKenzie, A.Sowden, S.V.Katikireddi, S.E.Brennan, S.Ellis, J.Hartmann-Boyce, R.Ryan, S.Shepperd,

J.Thomas, V.Welch, H.Thomson. *BMJ*, **368**, l6890 (2020); https://doi.org/10.1136/bmj.l6890
25. C.Baethge, S.Goldbeck-Wood, S.Mertens. *Res. Integr. Peer Rev.*, **4**, 5 (2019); https://doi.org/10.1186/s41073-019-0064-8
26. B.J.Shea, B.C.Reeves, G.A.Wells, M.Thuku, C.Hamel, J.Moran, D.Moher, P.Tugwell, V.Welch, E.Kristjansson, D.A.Henry. *BMJ*, **358**, j4008 (2017); https://doi.org/10.1136/bmj.j4008
27. P.Whiting, J.Savović, J.P.T.Higgins, D.M.Caldwell, B.C.Reeves, B.Shea, P.Davies, J.Kleijnen, R.Churchill. *J. Clin. Epidemiol.*, **69**, 225 (2016); https://doi.org/10.1016/j.jclinepi.2015.06.005
28. M.J.Grant, A.Booth. *Health Inf. Libr. J.*, **26**, 91 (2009); https://doi.org/10.1111/j.1471-1842.2009.00848.x
29. R.Whittemore, K.A.Knafl. *J. Adv. Nurs.*, **52**, 546 (2005); https://doi.org/10.1111/j.1365-2648.2005.03621.x
30. R.J.Torraco. *Hum. Resour. Dev. Rev.*, **15**, 404 (2016); https://doi.org/10.1177/1534484316671606
31. G.E.Cacciamani, T.N.Chu, D.I.Sanford, A.Abreu, V.Duddalwar, A.A.Oberai, C.C.J.Kuo, X.Liu, A.K.Denniston, B.Vasey, P.McCulloch, R.F.Wolff, S.Mallett, J.Mongan, C.E.Kahn, V.Sounderajah, A.Darzi, P.Dahm, K.G.M.Moons, E.J.Topol, G.S.Collins, D.Moher, I.S.Gill, A.J.Hung. *Nat. Med.*, **29**, 14 (2023); https://doi.org/10.1038/s41591-022-02139-w
32. D.Holst, K.Moenck, J.Koch, O.Schmedemann, T.Schüppstuhl. *JMIR AI*, **4**, e80247 (2025); https://doi.org/10.2196/80247
33. C.Cao, R.Arora, P.Cento, A.Budak, K.Manta, E.Farahani, M.Cecere, A.Selemon, J.Sang, L.X.Gong, R.Kloosterman, S.Jiang, R.Saleh, D.Margalik, J.Lin, J.Jomy, J.Xie, D.Chen, J.Gorla, S.Lee, K.Zhang, J.Kuang, H.Ware, M.Whelan, B.Teja, A.A.C.Leung, R.K.Arora, J.Pillay, L.Hartling, M.Noetel, D.B.Emerson, A.S.Detsky, A.C.Tricco, G.M.Church, D.Moher, N.Bobrovitz. *medRxiv* (2025); https://doi.org/10.1101/2025.06.13.25329541
34. R.Sanghera, A.J.Thirunavukarasu, M.E.Khoury, J.O'Logbon, Y.Chen, A.Watt, M.Mahmood, H.Butt, G.Nishimura, A.Soltan. *arXiv preprint arXiv:2411.02451* (2024); https://doi.org/10.48550/arxiv.2411.02451
35. L.Joos, D.A.Keim, M.T.Fischer. Presented in part at the EuroVis Workshop on Visual Analytics, Luxembourg, 2025; https://doi.org/10.2312/eurova.20251105
36. R.Luo, Z.Sastimoglu, A.I.Faisal, M.J.Deen. *medRxiv* (2024); https://doi.org/10.1101/2024.06.03.24308405
37. C.H.Krag, T.Balschmidt, F.Bruun, M.Brejnebøl, J.J.Xu, M.Boesen, M.B.Andersen, F.C.Müller. *medRxiv* (2024); https://doi.org/10.1101/2024.10.01.24314702
38. T.Oami, Y.Okada, T.Nakada. *JAMA Netw. Open*, **7**, e2420496 (2024); https://doi.org/10.1001/jamanetworkopen.2024.20496
39. N.Bernard, Y.Sagawa Jr, N.Bier, T.Lihoreau, L.Pazart, T.Tannou. *BMC Med. Res. Methodol.*, **25**, 75 (2025); https://doi.org/10.1186/s12874-025-02528-y
40. M.A.Kacena, L.I.Plotkin, J.C.Fehrenbacher. *Curr. Osteoporos. Rep.*, **22**, 115 (2024); https://doi.org/10.1007/s11914-023-00852-0
41. T.J.Margetts, S.J.Karnik, H.S.Wang, L.I.Plotkin, A.L.Oblak, J.C.Fehrenbacher, M.A.Kacena, A.Movilă. *Curr. Osteoporos. Rep.*, **22**, 177 (2024); https://doi.org/10.1007/s11914-023-00853-z
42. R.Ramchandani, E.Guo, E.Rakab, J.Rathod, J.Strain, W.Klement, R.Shorr, E.Williams, D.Jones, S.Gilbert. *PeerJ Comput. Sci.*, **11**, e2822 (2025); https://doi.org/10.7717/peerj-cs.2822
43. S.Dogra, S.Arabshahi, J.Wei, E.Hu, L.Saidenberg, S.Sharma, Z.Gu, M.Siriruchatanon, S.K.Kang. *Acad. Radiol.*, **32**, 7023 (2025); https://doi.org/10.1016/j.acra.2025.08.014
44. C.Muddana, B.Wang, P.-T.Sun, Y.J.Tang. *Biotechnol. Adv.*, **88**, 108814 (2026); https://doi.org/10.1016/j.biotechadv.2026.108814
45. J.Zybaczyńska, M.Norris, S.Modi, J.Brennan, P.Jhaveri, T.J.Craig, T.Al-Shaikhly. *J. Allergy Clin. Immunol. Pract.*, **12**, 106 (2024); https://doi.org/10.1016/j.jaip.2023.10.010
46. O.D.Awosanya, A.Harris, A.Creecy, X.Qiao, A.J.Toepp, T.R.McCune, M.A.Kacena, M.V.Ozanne. *Curr. Osteoporos. Rep.*, **22**, 146 (2024); https://doi.org/10.1007/s11914-023-00855-x
47. V.P.Ananikov. *ChemRxiv* (2025); https://doi.org/10.26434/chemrxiv-2025-fbk8w
48. V.P.Ananikov. *Chim. Oggi – Chem. Today*, **42** (4), 14 (2024); https://www.teknoscienze.com/tks_article/will-artificial-intelligence-ai-replace-chemists/
49. L.McLaughlin, M.S.Walz, C.Arries. *medRxiv* (2026); https://doi.org/10.64898/2026.02.18.26346559
50. J.-L.Lieberum, M.Toews, M.-I.Metzendorf, F.Heilmeyer, W.Siemens, C.Haverkamp, D.Böhringer, J.J.Meerpohl, A.Eisele-Metzger. *J. Clin. Epidemiol.*, **181**, 111746 (2025); https://doi.org/10.1016/j.jclinepi.2025.111746
51. L.Ge, R.Agrawal, M.Singer, P.Kannapiran, J.A.De Castro Molina, K.L.Teow, C.W.Yap, J.A.Abisheganaden. *Syst. Rev.*, **13**, 269 (2024); https://doi.org/10.1186/s13643-024-02682-2
52. F.Bolaños, A.Salatino, F.Osborne, E.Motta. *Artif. Intell. Rev.*, **57**, 259 (2024); https://doi.org/10.1007/s10462-024-10902-3
53. H.S.Yun, I.J.Marshall, T.A.Trikalinos, B.C.Wallace. *arXiv preprint arXiv:2305.11828* (2023); https://doi.org/10.48550/arxiv.2305.11828
54. M.S.A.Sousa, S.Peiris, M.F.Figueiró, M.M.Haby, A.C.Baraldi, L.Reveiz, J.P.Souza. *Syst. Rev.*, **15**, 82 (2026); https://doi.org/10.1186/s13643-025-02842-y
55. D.Scherbakov, N.Hubig, V.Jansari, A.Bakumenko, L.A.Lenert. *J. Am. Med. Inform. Assoc.*, **32**, 1071 (2025); https://doi.org/10.1093/jamia/ocaf063
56. A.Bayani, L.P.E.Ewane, D.S.O.dos Anjos, M.Mac-Seing, J.N.Nikiema. *Int. J. Med. Inform.*, **204**, 106048 (2025); https://doi.org/10.1016/j.ijmedinf.2025.106048
57. Z.-Y.Dai, F.-Q.Wang, C.Shen, Y.-L.Ji, Z.-Y.Li, Y.Wang, Q.Pu. *J. Med. Internet Res.*, **27**, e67488 (2025); https://doi.org/10.2196/67488
58. S.H.B. van Dijk, M.G.J.Brusse-Keizer, C.C.Bucsán, J. van der Palen, C.J.M.Doggen, A.Lenferink. *BMJ Open*, **13**, e072254 (2023); https://doi.org/10.1136/bmjopen-2023-072254
59. B.Hasan, S.Saadi, N.S.Rajjoub, M.Hegazi, M.Al-Kordi, F.Fleti, M.Farah, I.B.Riaz, I.Banerjee, Z.Wang, M.H.Murad. *BMJ Evid. Based Med.*, **29**, 394 (2024); https://doi.org/10.1136/bmjebm-2023-112597
60. Y.Ye, M.Colombo, J.Meessen, N.Hajizadeh, L.D.Vore, S.Krovvidi, S.Meresh, H.G.Sigarchian, S.Jakka, B.Tyl. *Eur. Heart J. Digit. Health*, **7**, ztaf143.008 (2026); https://doi.org/10.1093/ehjdh/ztaf143.008
61. E.Guo, M.Gupta, J.Deng, Y.J.Park, M.Paget, C.Naugler. *J. Med. Internet Res.*, **26**, e48996 (2024); https://doi.org/10.2196/48996
62. J.Köhler, M.V.Harl, M.Westner, S.Strahringer. Presented in part at the 27th International Conference on Business Informatics (CBI), Lisbon, Portugal, 2025; https://doi.org/10.1109/cbi68102.2025.00012
63. S.P.Thomas. *Issues Ment. Health Nurs.*, **46**, 1175 (2025); https://doi.org/10.1080/01612840.2025.2587574
64. TRACE: Transparent Review Acceleration with Co-Expert AI, https://ananikovlab.ru/TRACE (Last accessed November 2026)
65. D.Arkhipova, A.Kashin, A.Vorozhtsov, V.Ananikov. *Russ. Chem. Rev.*, **95** (11), RCR5240 (2026); https://doi.org/10.59761/RCR5240
66. J.Thomas, E.Flemyng, A.Noel-Storr. Responsible use of AI in evidence SynthEsis (RAISE): recommendations and guidance, https://osf.io/fwaud/overview. (Last accessed September 2026); https://doi.org/10.17605/OSF.IO/FWAUD