

AI-POWERED RESEARCH QUESTIONS

AI-POWERED ASSISTANCE ACROSS THE RESEARCHER'S WORKFLOW

Ethical Tech CoLab

October 2025

Yorke E. Rhodes III, Nathaniel Fossella, Alexa Shamie, Kirsten Co, Taylor Badt, Amanda Lindsey, Grace Driscoll, Mohagani Townsend, Pegi Bracaj, and Vedant Jain

AI is increasingly used to help academic researchers pose relevant research questions by mining vast literature for insights. This paper is an overview of how AI techniques and tools address key tasks in a researcher's workflow – finding research gaps, generating candidate questions, summarizing the state of the art, and comparing conflicting results.

KEY FIGURES

58M PAPERS

distilled into SCIMUSE's knowledge graph before GPT-4 proposed new research directions

25%

of 4,451 AI-generated ideas were rated highly interesting by expert researchers

800M+

citation statements Scite classifies as supporting, mentioning, or contrasting prior work

5 TASKS

gap-finding, question generation, summarization, contradiction detection, and integration

01

AI FOR IDENTIFYING RESEARCH GAPS

One of the first steps in posing a new research question is recognizing what hasn't been answered yet. AI can analyze large volumes of literature to detect gaps – unexplored or under-investigated topics – much faster than manual reading.

NLP-Based Gap Finders Tools like GapFinder (Pessin et al., 2025) use natural language processing to scan academic texts and map out research gaps. GapFinder mines unstructured paper content (PDFs) to identify areas not well covered by existing studies, acting as an automatic gap analysis that highlights topics or questions which remain unanswered in the current literature.

Keyword Co-occurrence Graphs Another approach leverages network algorithms on publication data. Chatterjee et al. (2024) construct keyword co-occurrence graphs and algorithmically detect structures that indicate a potential gap – strongly connected clusters of concepts that have not been jointly explored, i.e. interconnected yet uncharted research themes. By identifying combinations of topics that appear related but lack corresponding studies, such graph-mining techniques point researchers to plausible gaps that conventional reviews might overlook.

Knowledge Graph Analysis Large-scale knowledge graphs – networks of research concepts and their relationships – can be analyzed to spot missing links. If two areas are connected indirectly through known links but no direct research ties them together, that suggests a gap. Modern AI systems traverse these graphs to suggest which conceptual connections lack supporting studies, aligning with earlier literature-based discovery methods now enhanced by AI for scale and depth.

By automating gap discovery, AI tools give researchers a focused starting point – the not-yet-answered questions or unexplored topic intersections – on which to formulate new research questions.

02

AI FOR FORMULATING NEW RESEARCH QUESTIONS

Once gaps or novel idea spaces are identified, AI can help formulate concrete research questions or hypotheses. Recent advances in large language models and knowledge integration have produced systems that generate research ideas or suggest next steps, effectively acting as brainstorming partners grounded in data.

LLM-Generated Research Ideas A cutting-edge example is SCIMUSE (Gu & Krenn, NeurIPS 2024), which combines a massive literature-derived knowledge graph with GPT-4 to propose new research directions. SCIMUSE distilled knowledge from 58 million papers and prompted GPT-4 to generate novel project ideas tailored to a scientist's interests. Tested with over 100 senior researchers who evaluated 4,451 AI-proposed ideas, roughly 25% were rated highly interesting (a score of 4 or 5 out of 5). Such models can not only suggest topics but formulate them as concrete proposals, often including potential methodologies or collaborations.

Automated Hypothesis Generation Even before the latest LLMs, AI frameworks were generating hypotheses from existing findings. AGATHA (Sybrandt et al., 2020) combines graph mining and deep learning to propose new biomedical hypotheses, analyzing millions of scientific statements – interactions between genes, diseases, and chemicals – and learning to rank which unexplored connections are most plausible. In validation, AGATHA predicted research links later confirmed by papers published after 2015, demonstrating an ability to foresee meaningful research questions early.

Frameworks and Methodologies The general methodology is to leverage vast corpora – papers, databases, knowledge graphs – and use algorithms to recombine existing knowledge into new statements. Some rely on co-occurrence patterns that reveal an intermediate concept linking two previously unrelated ideas; others use language models to creatively assemble known facts into an innovative question. The common thread is that AI can connect dots across the literature in ways a human might overlook.

Used this way, AI gives researchers a systematic means of ideation. Instead of starting from scratch, they can explore a menu of AI-suggested questions that align with identified gaps, then apply their expertise to refine or pursue the most promising ones.

03

AI FOR SUMMARIZING THE STATE OF THE ART

To pose a relevant question, a researcher must understand current knowledge – what is already known and what remains uncertain. AI assists through advanced summarization, condensing an ever-growing literature into digestible syntheses of the state of the art.

Automated Literature Summaries AI tools can ingest multiple papers on a topic and produce concise summaries of key findings, methods, and consensus. LLM-powered search platforms such as Elicit, Scispace, and Consensus let users ask a research question in natural language and return relevant papers, each with a brief summary of its contribution – giving a quick overview of current knowledge without reading each paper in full.

State-of-the-Art Overviews Some systems perform multi-document summarization, synthesizing information across many papers into one coherent overview. Work here includes using models like BERT and SciBERT for scientific text summarization and efforts to generate literature reviews automatically. Altmami & Menai (2022) survey machine-learning techniques for summarizing scientific articles, and shared tasks such as Allen AI's Multi-Document Summarization for Literature Review aim to train AI to write an aggregated summary from dozens of papers – effectively a draft related-work section.

Knowledge Graphs and Semantic Summaries Structured representations such as AIDA or the Open Research Knowledge Graph (ORKG) encode scholarly knowledge as interconnected data – papers, methods, and results as linked entries. AI

can traverse these graphs to produce a comparative view of the state of the art (for example, listing all known approaches to a problem and their outcomes) and, importantly, highlight which nodes are missing, tying back into gap detection.

AI-driven summarization dramatically speeds up literature review. A researcher can quickly learn what is known about a topic, or which methods have been tried, rather than spending weeks reading. This focused understanding helps ground new questions in current knowledge – building on what is known and specifically targeting what is not yet well understood.

04

AI FOR COMPARING AND CONTRASTING RESEARCH FINDINGS

Researchers often need to reconcile conflicting results. AI can compare publications and flag where their findings diverge, which is crucial for refining research questions and identifying controversies to investigate.

Citation Analysis for Disagreements Scite, an AI-powered citation index, uses deep learning to classify the context of citations as supporting, mentioning, or contrasting the cited work. Drawing on over 800 million citation statements, its algorithms determine when a study's results have been challenged by subsequent work. Instead of manually collating hints of disagreement, a researcher can see directly that one paper contradicts another's conclusion – making it far easier to pinpoint debates or inconsistencies in a field.

Natural Language Contradiction Detection Beyond citation context, AI can compare the content of findings. Tawfik & Spruit (2018) developed a model for automated contradiction detection in biomedical literature that first extracts key outcome statements (e.g. 'Drug A improves condition B' versus 'Drug A has no effect on condition B'), then uses semantic features – negation, antonyms, and numeric comparisons – to judge whether two statements conflict. Applied to clinical studies, it identified when different papers answered the same question with opposing conclusions, surfacing controversies that often lead directly to follow-up questions.

Cross-Paper Evidence Synthesis AI is also used in systematic reviews to ensure all angles are considered. Machine-learning classifiers can cluster papers by outcome or stance, semi-automatically separating studies that

support a hypothesis from those that do not, and alerting the researcher to clusters of contradictions worth investigating.

By exposing inconsistencies and debates, AI tools show researchers where the science is not settled. These contested areas often make for important research questions – for instance, 'What explains the divergent results regarding X?' AI does the heavy lifting of identifying knowledge conflicts so the researcher can focus on resolving them.

05

INTEGRATED AI SYSTEMS FOR RESEARCH-QUESTION GENERATION

Bringing these capabilities together, modern AI research assistants provide end-to-end support – from mapping existing knowledge to highlighting gaps and inconsistencies, often in an interactive workflow.

Consolidated Platforms New research-assistant tools – commercial and experimental platforms such as Semantic Scholar's AI, IBM Watson Discovery for science, and academic prototypes – integrate search, summarization, and analysis. Some AI writing assistants (e.g. Jenni.ai, Silatus) let a user enter a tentative topic, then automatically fetch relevant literature, summarize it, and suggest next questions. As Bolaños et al. (2024) describe, such systems can iteratively build a literature review, surfacing open questions during the writing process – noting, for instance, that 'few studies address Z,' cueing a potential new question.

Hypothesis and Idea Suggestion Integrated agents can propose questions in context. A researcher might upload a set of papers and ask 'what is a good next research question here?' The AI combines the findings, points out a missing piece, and phrases a candidate question. Some systems use retrieval-augmented generation (RAG) to pull specifics from documents and have an LLM draft open-problem paragraphs, explicitly listing unsolved problems gleaned from the literature.

Workflow Efficiency The overall impact is on efficiency and comprehensiveness. By the time a researcher formulates a question, they have already seen the summarized state of the art, the known gaps, and the controversial findings – reducing the chance of proposing something trivial or already answered. As surveys note, current tools are not flawless (hallucinations and gaps in coverage exist), but the trend is for AI to handle more of the heavy lifting, leaving researchers to concentrate on

creative and critical thinking. Human insight remains crucial: the AI provides options and evidence, but the researcher chooses and refines the question.

AI contributes at every stage of question formulation – scouting the landscape of knowledge, spotlighting what is unknown or disputed, and even drafting candidate questions. For a researcher, these tools act like a smart assistant, ensuring the questions they pose are informed, significant, and address genuine gaps: a faster, more focused path to high-impact research.

06

THE EMOTIONAL DIMENSION OF RESEARCH IN THE AGE OF AI

While AI-powered tools can streamline the formulation of research questions and accelerate literature analysis, their growing presence does not diminish the emotional gratification traditionally associated with academic work. Rather, it redefines how that sense of relief, pride, and satisfaction is experienced. The fulfillment that arises from researching and writing – through effort, reflection, and perseverance – remains intact when scholars engage intentionally and thoughtfully with AI. The time invested in interpreting, refining, and shaping ideas continues to enhance both intellectual rigor and personal reward.

Instead of reducing the depth or authenticity of scholarly effort, AI research assistants can preserve and even amplify the emotional dimensions of academic inquiry. By relieving researchers of repetitive or mechanical tasks, AI allows greater focus on the creative, analytical, and interpretive aspects that form the heart of scholarship. The gradual movement from uncertainty to clarity – the hallmark of intellectual discovery – still occurs; it simply unfolds through collaboration between human judgment and machine assistance. The researcher's sense of accomplishment now stems not only from producing results but from orchestrating and curating the knowledge process itself.

Protecting and cultivating emotional satisfaction in this environment requires recognizing that human agency remains central. AI can efficiently identify relevant literature or summarize findings, but the researcher continues to decide what matters, why it matters, and how it contributes to the broader conversation. The source of pride thus shifts from labor-intensive information gathering to higher-order critical thinking – evaluating, connecting, and creating meaning.

Many researchers also choose to maintain elements of manual engagement, such as drafting outlines or annotating papers independently before consulting AI tools. These deliberate choices reinforce the personal connection to one's work and sustain the emotional gratification that comes from intellectual perseverance. The presence of AI does not erase these experiences; it reframes them within a more collaborative and dynamic process of knowledge creation.

Ultimately, AI research assistants do not weaken the emotional core of academic work – they evolve it. Relief, pride, and satisfaction remain grounded in the same human capacities that have always driven scholarship: curiosity, discernment, and creativity. As AI becomes more deeply integrated into research workflows, the challenge is not to preserve emotion against technology but to recognize how these tools can extend the very sense of fulfillment that comes from thinking deeply, questioning critically, and contributing meaningfully to human understanding.

07

GUARDRAILS FOR AI-ASSISTED RESEARCH QUESTIONING

These advances demonstrate the growing role of AI in shaping scholarly inquiry and research design. However, these capabilities introduce parallel risks that necessitate structured guardrails. While AI can help align research questions with documented gaps in the literature, it is equally capable of hallucinating citations, fabricating conceptual linkages, or overstating consensus where none exists. The same generative power that lets AI suggest high-value inquiries can also produce plausible-sounding but incorrect premises, or subtly shift a researcher's framing toward what the model predicts rather than what the literature supports. When AI tools summarize or frame a body of work, researchers may be tempted to rely on the output without independently engaging the foundational texts – risking a second-order understanding and a loss of methodological rigor.

Given these dynamics, responsible adoption requires AI systems that do more than generate ideas: they must support verification, transparency, and careful acknowledgement of uncertainty. Guardrails are not barriers to innovation; they are mechanisms to ensure that efficiency gains do not erode scholarly integrity. These controls include expectations that researchers:

- Independently review core sources rather than relying solely on AI summaries.

- Confirm that AI-generated citations and claims are accurate and grounded in the original texts.
- Treat AI-suggested research questions as exploratory starting points, not final conclusions.
- Reflect on and document how AI shaped their thinking, assumptions, or direction.

Establishing these practices is essential for maintaining academic rigor, particularly as graduate-level research increasingly intersects with intelligent systems capable of shaping epistemic direction.

08

AI-ENABLED DETECTION OF CONFLICTS OF INTEREST

Conflicts of interest undermine the credibility of academic research and weaken trust in scientific institutions. Financial ties, organizational affiliations, or personal relationships can influence how results are framed or interpreted, and manual review alone is not equipped to keep pace with the volume and complexity of today's scholarly output. AI systems offer a scalable, systematic way to identify signals of undisclosed or inaccurately reported conflicts, strengthening transparency across the publication process.

AI can scan manuscripts for references to companies, products, funding sources, and affiliations, and compare those details with the information listed in author disclosures. When it notices a discrepancy – such as prominent discussion of a commercial product without a corresponding disclosure – it can flag the passage for further review.

Beyond individual papers, AI can detect broader patterns across publications, including recurring collaborations or concentrated citation behavior that may indicate undisclosed relationships. These signals are not conclusions, but they prompt a closer look on which a researcher can decide whether follow-up is needed – expanding analytical capacity and strengthening transparency in the academic publishing process.

References

SOURCES

- 01 Pessin et al. (2025) – GapFinder: NLP mining of unstructured papers to surface unanswered questions.
- 02 Chatterjee et al. (2024) – Keyword co-occurrence graph mining to detect uncharted research themes.
- 03 Gu & Krenn (2024) – SCIMUSE: a 58M-paper knowledge graph paired with GPT-4 to propose research ideas (NeurIPS 2024, ML4Physical Sciences).
- 04 Sybrandt et al. (2020) – AGATHA: graph mining and deep learning for biomedical hypothesis generation.
- 05 Altmami & Menai (2022) – A survey of machine-learning techniques for summarizing scientific articles (Springer).
- 06 Scite – Smart Citations classify citation context as supporting, mentioning, or contrasting (Journal of the Medical Library Association).
- 07 Tawfik & Spruit (2018) – Automated contradiction detection in biomedical literature (Springer).
- 08 Bolaños et al. (2024) – AI systems that iteratively build cited literature reviews (Springer).

This paper is supported by AI-enabled research assistance: the Microsoft 365 Copilot Researcher agent, running OpenAI's GPT-5 model, was used to partially generate and check content.