AI-Powered Academic Search and Research in 2026: What Changed in August, What the New Features Do, and Why the Database Model Is Changing

Executive finding

August 2026 marks a significant transition in academic search. The earlier generation of AI features focused mainly on helping a researcher formulate a query, improve relevance ranking, summarize one document, or ask questions about a retrieved paper. By August, the leading systems increasingly combine several functions: semantic retrieval, query decomposition, reranking, iterative search, citation-based exploration, multi-document synthesis, research-agent workflows, and connections from scholarly databases into general AI environments through the Model Context Protocol, or MCP. Aaron Tay's August 21 analysis captures this transition through Web of Science Smart Search, Google Scholar Labs, and Consensus. Aster Zhao's September 11 comparison shows how the same transition appears inside EBSCO, ProQuest, and JSTOR. [1]

Aaron Tay's LinkedIn post provides an important principle for evaluating these products. He argues that researchers and librarians need to move beyond impressions about interface quality or plausible-looking answers and examine what the system does "under the hood": how retrieval works, what AI changes in the workflow, and where limitations occur. This principle matters more in 2026 because products carrying similar "AI-powered" labels now implement quite different retrieval and synthesis architectures. [2]

The most important distinction is no longer AI versus non-AI search. The useful distinctions now concern where AI enters the information process and how much control it assumes. AI can translate a natural-language question into Boolean syntax, retrieve by semantic similarity, rerank candidate documents, judge relevance, explain individual papers, synthesize evidence from multiple papers, plan successive searches, follow citation networks, or call scholarly databases from another AI application. EBSCO, ProQuest, JSTOR, Web of Science, Scholar Labs, Consensus, Scite, Elicit, and OpenAlex occupy different positions across this continuum. [3]

August also produced a particularly revealing sequence of developments. Elicit announced a substantially expanded Research Agent on August 4. Consensus published a retrieval evaluation on August 4 and continued developing its agentic search infrastructure. Aaron Tay published his comparative analysis on August 21. EBSCO documented AI Insights on August 24. Google Scholar introduced Quick Read on August 25. Scite's August release added one-click MCP connections to ChatGPT and Claude. OpenAlex's August documentation exposed semantic search over scholarly embeddings as an API capability. Taken together, these developments

indicate a transition from "AI inside a database" toward "scholarly retrieval as an AI-accessible research infrastructure." [4]

A second transition is equally important. Search and reading are starting to merge. Google Scholar Quick Read, EBSCO AI Ask, ProQuest Research Assistant, JSTOR AI Research Tool, and Consensus Citation Grounding increasingly make the retrieved document directly interrogable. Retrieval no longer necessarily ends with a list of titles. The result itself becomes an interactive evidence object, with AI explaining relevance, extracting findings, locating supporting passages, and identifying limitations. [5]

A third transition concerns the role of the database. The conventional model assumes a researcher visits Web of Science, EBSCOhost, ProQuest, JSTOR, Scopus, or Google Scholar and searches through its interface. MCP and API integrations challenge this assumption. Consensus, Elicit, Scite, EBSCO, and Clarivate increasingly expose scholarly retrieval to external AI environments. Tay explicitly identifies this possibility in Consensus: the search provider may become a backend service while an LLM environment becomes the user's primary interface. Zhao reaches a similar conclusion when discussing EBSCO Bridge and Clarivate Nexus Connect. [6]

The consequence for research methodology is substantial. Researchers increasingly need to evaluate six layers separately: corpus coverage, query interpretation, retrieval, reranking, generation or synthesis, and provenance. A fluent answer provides little evidence about search completeness. A sophisticated agent cannot retrieve papers absent from its index. A relevant top ten says little about recall. A citation does not guarantee an accurate interpretation of the cited paper. Repeated AI searches may also return different candidates or rankings. Tay's reproducibility tests, Zhao's product testing, and the 2026 OpenScholar research all point toward these methodological distinctions. [7]

What changed during August 2026

The developments around August become clearer when viewed as changes to successive stages of scholarly information work.

Date	Development	What changes in the research workflow
August 4	Elicit Research Agent expansion	Search becomes an agentic research process spanning scholarly papers, structured databases, web sources, uploaded material, analysis, tables, figures, and reports. [8]

Date	Development	What changes in the research workflow
August 4	Consensus publishes scientific-search evaluation	Retrieval relevance becomes an explicit evaluation target separate from final-answer quality. Consensus's own evaluation compares its scholarly retrieval with general-purpose frontier models using 100 anonymized research queries. [9]
August 11	OpenAlex semantic-search documentation updated	Semantic retrieval becomes available as open scholarly infrastructure, including embedding-based matching for paragraph-length inputs, abstracts, and grant descriptions. [10]
August 21	Aaron Tay compares WoS Smart Search, Scholar Labs, Consensus	The comparison makes visible three competing futures: hybrid retrieval, iterative AI retrieval without synthesis, and agentic retrieval plus synthesis. [11]
August 24	EBSCO documents AI Insights	Article-level AI summaries become embedded directly in a large licensed aggregation environment and use RAG over selected full text. [12]
August 25	Google Scholar launches Quick Read	Google Scholar crosses from AI-assisted retrieval

Date	Development	What changes in the research workflow
		toward query-specific interpretation of individual accessible papers. [13]
During August	Scite simplifies MCP integration	A scholarly citation system becomes callable directly from ChatGPT, Claude, and other MCP-compatible environments. [14]

Some EBSCO developments technically began at the end of July but became part of the August product environment Zhao examines. EBSCO announced AI Conversational Search and Ask This Document on July 30. Conversational Search initially launched for Business Source Ultimate through the Business Searching Interface, while Ask This Document grounded answers in an individual full-text item. The important August implication is the integration of retrieval, follow-up questioning, synthesis, and source inspection inside a traditional aggregator. [15]

Google Scholar's August 25 Quick Read release is especially important because it appeared four days after Tay's Katina article. Tay accurately described Scholar Labs at the time of publication as an AI-enhanced deep-search system returning ranked papers rather than synthesized answers. Quick Read changes the boundary. Scholar Labs still does not produce a single literature-wide synthesis, but selected papers can now receive query-specific presentations organized around Answers, Approach, and Considerations. Google limits Quick Read to accessible papers, logged-in users, and, at launch, English-language articles. [16]

This distinction matters. Scholar Labs preserves the ranked-paper model at the collection level while introducing generative interpretation at the document level. Google has therefore separated "Which papers answer my question?" from "What does this paper say about my question?" rather than combining both into one automatically generated literature review. [17]

EBSCO follows another path. AI-assisted Search translates ordinary-language input into a Boolean-style query. AI Insights gives two to five key points from an article. AI Ask supports document-level questions. AI Conversational Search supports question-and-follow-up interaction across relevant EBSCO material. The resulting product architecture spans query formulation, retrieval assistance, relevance assessment, reading support, and conversational evidence exploration. [18]

ProQuest remains more document-centered in Zhao's evaluation. ProQuest Research Assistant can produce takeaways, answer questions, suggest related topics, and create topic mind maps for eligible documents. Ebook Central's assistant works at chapter

level with more constrained prompts. Zhao found no ProQuest equivalent to EBSCO's natural-language-to-query translation at the initial search stage. This means ProQuest's AI contribution sits mainly after retrieval rather than in retrieval itself. [19]

JSTOR occupies a different position. Semantic Results modifies retrieval itself by matching contextual similarity rather than relying only on lexical overlap, while its AI Research Tool supports RAG-based interaction with an open document or the returned semantic result set. Zhao reports a maximum of 25 Semantic Results and found repeated conversational answers often relied on only one or two papers from the retrieved set. This makes JSTOR stronger for exploratory relevance assessment and document interpretation than for comprehensive cross-literature synthesis. [19]

The August Scite development shows why the interface-level comparison captures only part of the 2026 transition. Scite's release notes describe a one-click connection from Scite to ChatGPT and Claude, plus support for other MCP-compatible tools. The same release improved Zotero group and laboratory-library imports. Instead of making a researcher move repeatedly among reference manager, database, citation index, and chatbot, the components increasingly connect into a shared research environment. [14]

OpenAlex provides another part of this infrastructure. Its semantic-search API embeds titles and abstracts and compares their vectors with an embedded user query. It accepts up to 2,000 characters for semantic matching, allowing a researcher or agent to supply a paragraph, grant aim, research description, or abstract rather than a few keywords. OpenAlex documents GTE Large EN embeddings with 1,024 dimensions and cosine similarity for matching. Its semantic results can also combine with most OpenAlex filters. [10]

These developments mean that August 2026 should not be described as a month when databases merely "added chatbots." The deeper change concerns where machine reasoning enters the retrieval pipeline and how scholarly databases expose their content to AI systems. [20]

What the AI-powered features are doing under the hood

Aaron Tay's August article offers one of the clearest frameworks for understanding the new academic-search market. He separates systems along two dimensions: how deeply they search and whether they return ranked results or generated synthesis. This produces four categories: quick search, deep search, quick RAG, and deep research. [21]

Web of Science Smart Search represents quick, hybrid retrieval. According to Tay's analysis, Web of Science processes natural-language input into a structured Web of Science query and also creates a semantic representation. It combines conventional structured retrieval and vector-based semantic retrieval through reranking. The user still receives a ranked result list rather than an AI-written answer. Advanced Search also remains available for researchers needing controlled Boolean syntax. [22]

This design keeps AI primarily on the retrieval side of the boundary. For research methodology, such a model has an attractive property: the researcher remains responsible for interpretation of the documents. The system helps decide which papers rise in the result list but does not replace reading with a synthesized narrative. Tay interprets this as Clarivate preserving Web of Science's role as a trusted retrieval layer. [23]

Google Scholar Labs represents deep search rather than quick search. Google describes Labs as identifying key topics, aspects, and relationships in a research question, searching Scholar for these components, evaluating retrieved papers against the overall question, and explaining how each selected paper may help. Google's June 2026 update reported roughly tenfold faster operation, up to three times more papers scanned, and ten times more searches allowed per day. Tay observed Scholar Labs running more than ten searches for a query and progressively evaluating candidate records for relevance. [24]

The central innovation here is AI relevance judgment. Traditional academic search mainly computes relevance from query-document relationships using lexical, fielded, citation, or increasingly vector signals. Scholar Labs adds another layer in which AI evaluates candidate papers against a more complex interpretation of the research question. Researchers therefore give up some transparency in exchange for potentially improved relevance. Google does not publicly expose every relevance decision rule or every underlying subquery. [25]

Consensus moves further toward full research automation. Tay reports a multi-stage retrieval architecture combining lexical and semantic retrieval, reranking and rescoring, followed by synthesis. Deep Search runs multiple sub-searches and can combine conventional paper retrieval with citation-graph exploration before producing a longer report. Its interface exposes subqueries and search history more explicitly than many competing systems. [25]

Tay's distinction between "deep search" and "agentic search" is critical. A deep-search tool can execute a predetermined sequence of multiple searches. An agentic system lets an LLM decide which search or analysis tool to call next based on intermediate results. Consensus increasingly follows the latter approach. Its current Research Agent documentation describes planning sub-searches, applying filters, following citations and authors, reading and comparing papers, retaining conversational context, and identifying research gaps. [26]

This distinction changes the unit of search. In a conventional database the unit is a query. In a deep-search system the unit becomes a set of machine-generated queries. In an agentic system the unit becomes a research objective from which the agent derives a sequence of retrieval and analysis actions. [26]

Zhao's EBSCO comparison reveals a simpler but pedagogically important AI function: natural language to Boolean translation. EBSCO's AI-assisted Search can translate a question such as a causal or topical research question into expanded concepts and

Boolean combinations. The interface can show the refined query, which makes part of the transformation visible, although Zhao notes users cannot directly edit the generated form from the AI workflow and must change their question or return to Advanced Search. [19]

This function matters because it changes how novice users enter academic databases. Historically, database literacy required learning keyword decomposition, synonyms, Boolean operators, subject headings, field codes, phrase searching, truncation, proximity operators, and database-specific syntax before achieving a good search. AI-assisted search can move part of query formulation from prerequisite skill to inspectable output. The educational challenge then shifts toward teaching students to evaluate the generated concepts, omissions, and Boolean relationships. Zhao explicitly recommends EBSCO's visible translation as a search-skills teaching opportunity rather than a substitute for expert search design. [19]

Semantic search solves another problem. Lexical search asks whether documents contain the terms or variants specified in a query. Embedding-based search represents texts as numerical vectors and retrieves texts with similar meanings even when vocabulary differs. OpenAlex, for example, documents semantic retrieval based on title and abstract embeddings. JSTOR Semantic Results adds contextual matching to its conventional search. Web of Science combines semantic and structured retrieval. [27]

Reranking adds another hidden layer. A retrieval system may collect hundreds or thousands of candidate papers through keyword and semantic methods, then apply a more computationally expensive model to reorder a smaller candidate set. Tay's description of Consensus illustrates this architecture: broad retrieval first, richer relevance and quality judgment later. This pattern helps explain why "semantic search" alone tells us little about final results. The candidate-generation process, ranking model, quality signals, filters, and final reranker can all change what the researcher sees. [22]

RAG changes the workflow again. Instead of asking an LLM to answer from pretraining alone, a RAG system first retrieves relevant text and supplies retrieved material to the generation model. Zhao reports that EBSCO, ProQuest, and JSTOR ground their AI assistance in licensed content and provide pathways from generated outputs back to sources. The 2026 OpenScholar study offers independent research evidence for the importance of this architecture. OpenScholar retrieves from a large scholarly corpus, reranks evidence, generates citation-backed responses, and iteratively checks its own output. [28]

OpenScholar's Nature paper illustrates why retrieval quality deserves as much attention as generation quality. Its authors found high citation fabrication rates when an ungrounded GPT-4o setup was asked to cite recent literature in their experiments, while the specialized retrieval, reranking, and citation-validation pipeline substantially improved citation accuracy. The authors also emphasize that scientific literature synthesis still cannot be fully automated and depends on coverage, retrieval quality, attribution, and verification. [29]

Document-level Q&A represents a narrower and often safer use of the same principle. AI Ask, ProQuest Research Assistant, JSTOR AI Research Tool, and Google Scholar Quick Read focus the model on one accessible document or a restricted retrieved set. For many research tasks, this may deliver more value than broad autonomous synthesis because the researcher can compare the generated statement with the underlying paper. [30]

Google's Quick Read is an instructive design. Instead of offering a generic summary of a paper, it creates a query-specific reading aid. Its structure separates what the paper says about the user's question, the relevant methodological approach, and considerations or caveats. This moves beyond "summarize this PDF" toward "explain this paper in relation to my research question." [13]

The final technical shift is MCP and API access. Here AI stops being a feature embedded in the search interface and becomes an orchestrator. Scite's MCP lets external AI applications call Scite. Elicit provides Research Agent capabilities through its own application and API. Consensus has also invested in MCP and agentic search. EBSCO is developing external AI connections through Bridge, while Zhao describes Clarivate's Nexus Connect strategy for bringing licensed scholarly resources into external AI environments. [31]

The technical progression in 2026 therefore looks approximately like this:

natural-language question → query interpretation → Boolean and/or semantic retrieval → candidate pooling → reranking → source selection → document interrogation → multi-document synthesis → iterative agent actions → externally callable scholarly services.

The platforms differ mainly in where they enter this chain, where they stop, and how much of the intermediate process they reveal. [32]

From academic databases to research agents

The strongest signal across the additional August sources is the transition from search applications toward research agents.

Elicit's August 4 Research Agent announcement illustrates the endpoint more clearly than a database-specific chatbot. Elicit says its agent can combine scientific literature, clinical trials, patents, structured scientific databases, regulatory information, conference material, public web sources, and uploaded internal data. It can then create cited reports, tables, visualizations, calculations, slides, and documents while preserving project context. The company also exposes the system through an API. These are vendor descriptions rather than independent validation, but they show where the product category is moving. [8]

The change is conceptual. A literature-search system answers, "Which scholarly documents match this question?" A research agent receives something closer to, "Investigate this question, decide which evidence sources are required, search them,

compare the findings, analyze the data, and produce an evidence-linked output." Elicit's product architecture explicitly spans these later stages. [8]

Consensus has moved in a similar direction. Its summer 2026 update describes multi-step tool use combining semantic search, paper search, DOI lookup, citation-graph traversal, larger corpus coverage, citation grounding, and a library integrated with its Research Agent. Its stated corpus expansion and performance claims come from Consensus itself and should therefore be treated as vendor-reported information, but the architectural direction is clear. Search, reading, evidence organization, synthesis, and reference-library work are converging. [33]

Scite shows another version of the same change. Its August release concentrated on allowing ChatGPT, Claude, and other AI assistants to call Scite through MCP, while also strengthening Zotero integration. In this architecture the researcher may spend less time inside Scite itself. Scite supplies citation and evidence services to an external conversational environment. [14]

OpenAlex takes the infrastructure argument further. Its February 2026 release introduced semantic search alongside traditional advanced-search capabilities, long queries, and bulk access to open full text. Its August semantic-search documentation describes an API designed for meaning-based retrieval from descriptions substantially longer than conventional database keyword queries. OpenAlex therefore serves both researchers and machine agents. [34]

This supports Tay's prediction of a market split. One route preserves specialized research interfaces with explicit filters, retrieval controls, methods, and evidence-management features. The other route makes scholarly search a backend service available from the AI environment where the user already works. Consensus is pursuing both at once. Scite, Elicit, EBSCO, and Clarivate show elements of the same strategy. [6]

This development poses a strategic issue for libraries. The historical library technology stack assumed a discovery layer or database interface as the access point to licensed scholarship. MCP-style connections invert the relationship. The user's AI environment can become the front end; the library's licensed database becomes an authenticated evidence provider behind it. Zhao identifies EBSCO Bridge and Clarivate Nexus Connect as early examples of this outward movement of licensed content. [19]

The competitive advantage of a scholarly database may consequently shift away from interface design toward five assets: corpus coverage, licensing and full-text rights, metadata quality, retrieval quality, and machine-readable access. Consensus's own August retrieval evaluation makes a related argument: specialized corpus coverage, metadata, and ranking remain major determinants of scientific search quality even when general-purpose models have substantial reasoning capacity and web access. Since Consensus conducted the evaluation itself, its reported performance advantage should be read as vendor evidence rather than neutral benchmarking. [9]

The Nature OpenScholar study supports the broader principle from an independent research setting. Its results show that domain-specific retrieval infrastructure, evidence selection, reranking, citation attribution, and iterative verification can materially affect literature-synthesis performance. Larger generative models alone do not eliminate the information-retrieval problem. [29]

This helps explain a central paradox of academic AI in 2026. As the visible interface becomes simpler, the hidden search pipeline becomes more complex. A researcher may type one sentence, while the system generates numerous subqueries, performs semantic and lexical searches, follows citations, filters candidates, reranks papers, reads selected passages, compares claims, and generates a report. The user's query becomes easier to write, while the method becomes harder to audit. [35]

How AI changes search and research practice in 2026

The first major change concerns query formulation.

Researchers no longer always need to begin with a fully specified keyword strategy. EBSCO can transform plain language into Boolean concepts. Scholar Labs can decompose a complex research question into multiple searches. Semantic systems such as OpenAlex can retrieve by meaning from a sentence, paragraph, abstract, or grant description. Consensus Research Agent can interpret a request, generate sub-searches, and apply filters. [36]

For exploratory research, this reduces the cost of entering an unfamiliar literature. A student can begin with a conceptual question and use AI to identify terminology, synonyms, research designs, authors, and related questions. Zhao identifies vocabulary development, relevance assessment, and passage location as especially strong use cases for database-embedded AI. [19]

For expert research, the benefit is different. AI can serve as an additional recall mechanism. Semantic retrieval can retrieve work using unexpected terminology; iterative search can investigate several aspects of a question; citation graphs can surface earlier or later connected studies. These mechanisms complement rather than eliminate controlled vocabulary, field searching, citation searching, and database-specific strategies. [37]

The second change concerns relevance assessment.

Traditional search places much screening work on the researcher. AI increasingly pre-screens candidate papers and explains why each may be relevant. Scholar Labs gives paper-specific relevance descriptions. EBSCO AI Insights summarizes key points. Google Scholar Quick Read answers a query in relation to a particular paper. JSTOR explains relationships between a document and the current search. [38]

This creates a new research behavior: AI-assisted triage. Instead of opening every promising title and reading its abstract or full text, researchers can use machine-generated relevance explanations to decide which documents deserve

attention. The efficiency gain may be large when dealing with dozens or hundreds of candidates. The risk is selection bias. Zhao observes that EBSCO visually identifies AI-enabled content, which may encourage users to favor documents with AI assistance over equally relevant documents lacking compatible hosted full text. [19]

The third change concerns reading.

"Search result" and "document reader" increasingly form one workflow. EBSCO AI Ask, ProQuest Research Assistant, JSTOR AI Research Tool, Scholar Quick Read, Consensus Citation Grounding, and Elicit's agent all connect generated explanations to underlying evidence to varying degrees. Researchers can ask for findings, methods, relevance, limitations, comparisons, or supporting passages without leaving the research environment. [39]

This does not make close reading obsolete. It changes where close reading starts. AI can direct attention toward passages likely to matter; the researcher still needs to inspect methods, operational definitions, sample characteristics, statistical evidence, limitations, and context before using a claim in scholarly work. Zhao's testing found uneven document-assistant performance and source-navigation quality across vendors, reinforcing the need for verification. [19]

The fourth change concerns literature synthesis.

Consensus Deep Search, Elicit Research Agent, OpenScholar, Scopus AI Deep Research, and other systems move beyond source discovery toward multi-document reports. Tay's taxonomy treats this as a separate category from retrieval because the system now determines both which literature to use and how to describe the literature. [40]

This concentration of functions creates methodological dependence. A generated literature review inherits errors from every upstream layer:

question interpretation → corpus coverage → retrieval → deduplication → ranking → paper selection → passage extraction → evidence interpretation → synthesis.

A well-written final report can therefore conceal poor recall or biased candidate selection. Tay's focus on retrieval quality and Zhao's insistence on understanding what each feature really does are methodologically important precisely because polished generation does not validate preceding search stages. [41]

The fifth change concerns reproducibility.

Tay repeated identical searches and found substantial differences among systems. In his limited test, Web of Science Smart Search showed low top-ten stability, Consensus Deep Search produced a much more stable larger candidate pool than its precise ranking positions, and Scholar Labs returned identical top-ten results across the same-day repetitions. Tay explicitly cautions that his experiment used only one query

per system and same-day runs, so caching or query-specific behavior may explain part of the result. [23]

The methodological lesson is stronger than any ranking of the products. In a conventional systematic-search protocol, a researcher can report the database, date, fields, query string, limits, and resulting record count. An AI search may contain additional hidden or dynamic variables: model version, embedding model, generated subqueries, reranker, candidate cutoff, agent decisions, caching state, corpus snapshot, full-text availability, and random generation. Reproducibility therefore requires logging more than the visible user prompt. Tay explicitly argues that unstable search steps require acknowledgement when used in systematic or scoping review workflows. [23]

The sixth change concerns the difference between discovery and systematic evidence retrieval.

Zhao concludes that EBSCO, ProQuest, and JSTOR's current AI features work best as guided starting points. She recommends them for vocabulary development, relevance assessment, document understanding, and exploratory work, while retaining advanced database searching, citation searching, structured review methods, and close reading for advanced research. [19]

This distinction should govern academic use in 2026. For exploratory research, the optimization target is often rapid learning: find useful concepts and representative papers quickly. For a systematic review, the target is defensible coverage under an explicit protocol. The best tool for the first problem is not necessarily the best tool for the second. Semantic ranking and agentic search may improve discovery while making the exact search pathway less reproducible. [42]

The seventh change concerns the researcher's core skill set.

Search literacy in 2026 increasingly includes AI retrieval literacy. Researchers still need vocabulary, source evaluation, Boolean logic, controlled vocabularies, and citation searching. They also need to understand lexical versus semantic retrieval, generated queries, reranking, RAG grounding, document versus corpus-level Q&A, agentic search, corpus coverage, source attribution, and run-to-run stability. [43]

A useful practical rule follows from these developments: use AI to expand your search space and reduce mechanical work, but preserve human control over evidence inclusion, interpretation, and methodological reporting when research conclusions depend on comprehensiveness or reproducibility. Zhao's recommendations and Tay's reproducibility findings strongly support this division of labor. [42]

Reading the attached figure

Figure 1, overview of AI functions across EBSCOhost, ProQuest, and JSTOR

Figure 1 maps the AI tools/features in EBSCO, ProQuest, and JSTOR to the stages of the search process.

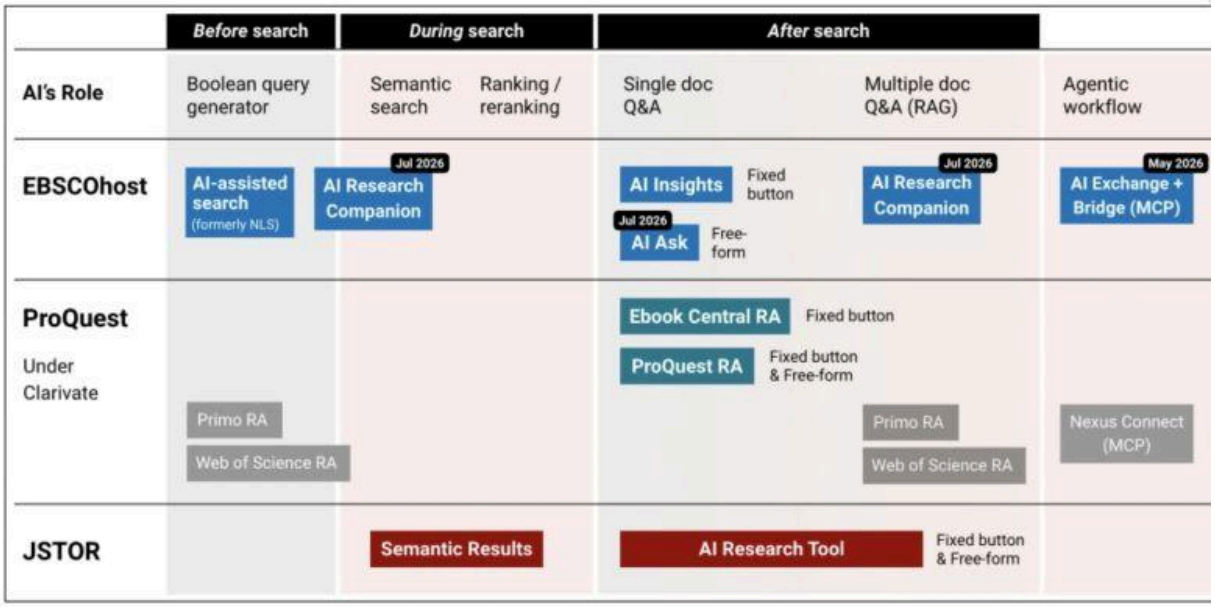

FIGURE 1 Overview of AI features in EBSCO, ProQuest, JSTOR

The attached figure comes from Zhao's analysis and offers a useful three-stage model of AI in database research: before search, during search, and after search. Its greatest strength is conceptual clarity. Instead of labeling an entire platform "AI-powered," it asks exactly where AI intervenes. Zhao uses the same structure in her text, identifying query translation before search, semantic retrieval during search, and document assistance or RAG after retrieval. [19]

The first column, before search, identifies the Boolean query generator. EBSCO's AI-assisted Search occupies this space because the system converts a natural-language question into a Boolean-style expression and passes the result through EBSCO's normal search infrastructure. AI therefore changes query construction more than the underlying retrieval logic. Zhao's technical comparison explicitly notes that EBSCO's AI-assisted search relies on Boolean and keyword retrieval rather than full semantic retrieval. [19]

This is a valuable distinction for teaching AI literacy. A natural-language interface can look like semantic or generative search even when the system behind the interface ultimately executes a conventional query. The visible interface does not reveal the

retrieval architecture. Tay's LinkedIn observation about investigating what happens under the hood applies directly here. [44]

The second column, during search, includes semantic search and ranking or reranking. JSTOR Semantic Results belongs clearly here. The system changes which documents enter and rise within the result set. The same conceptual category extends beyond Zhao's three databases to Web of Science Smart Search, OpenAlex semantic search, Consensus hybrid retrieval and reranking, and Scholar Labs' iterative relevance assessment. [45]

This stage deserves more attention than summary features because it determines the evidence available downstream. A perfect summarizer cannot synthesize a relevant paper the retrieval system never found. The Consensus August benchmark, despite being vendor-run, explicitly frames specialized retrieval and ranking as core academic-search problems. The OpenScholar study independently demonstrates the importance of retrieval and reranking within a literature-synthesis architecture. [46]

The third column, after search, shows three increasingly ambitious functions: single-document Q&A, multiple-document Q&A through RAG, and agentic workflows.

Single-document Q&A represents the most mature pattern among the legacy database products. EBSCO AI Ask and AI Insights, ProQuest Research Assistant, Ebook Central Research Assistant, and JSTOR's document-level Research Tool work largely in this zone. They assist reading after a source has already been identified. [19]

Google Scholar Quick Read, introduced after Tay's August 21 article and after the July product snapshot underlying Zhao's figure, would now fit naturally into this part of the model. It operates on an individual accessible paper and tailors its explanation to the current research query. Adding Quick Read to an expanded version of the figure would show how rapidly traditional search products are moving into AI-assisted reading. [13]

Multiple-document RAG is more demanding. A system must retrieve an appropriate evidence set, select passages, represent disagreements, and synthesize findings while preserving provenance. Zhao's figure positions EBSCO's Research Companion around this area, while her text emphasizes its initial limitation to Business Source Ultimate during the evaluated period. JSTOR can answer over up to 25 Semantic Results, but Zhao found its responses often concentrated on only one or two documents. [19]

Consensus Deep Search, Elicit Research Agent, and OpenScholar represent more advanced versions outside the three core platforms in Zhao's figure. They place multi-source evidence synthesis closer to the center of the product rather than treating it mainly as an add-on to a database interface. [47]

The agentic-workflow box is the figure's most forward-looking element, but it also exposes a limitation of the linear "before, during, after" framework.

In an agentic system, "after search" stops being a distinct phase. The agent may search, inspect findings, formulate another query, follow a citation, retrieve another

paper, compare findings, apply a filter, return to search, and then revise its synthesis. Search and analysis form a loop rather than a sequence. Tay defines agentic search in terms of an LLM deciding which tools to call and when, based on intermediate results. Consensus Research Agent and Elicit Research Agent increasingly follow this model. [48]

I would therefore extend Zhao's figure from a linear three-stage model to a two-layer model.

The first layer would preserve her stages:

question formulation → retrieval → evidence interaction → synthesis.

The second layer would sit across all stages:

agentic orchestration, provenance, authentication, licensed-content access, citation tracing, and external AI integration.

This extension follows directly from the August 2026 developments in Consensus, Elicit, Scite, EBSCO, and Clarivate. MCP allows an external agent to invoke a scholarly service at different points in a research process rather than only after a conventional database search has finished. [49]

A second extension should add "verification" after generation. Current diagrams often end with Q&A or synthesis, but 2026 systems increasingly expose exact supporting text, document locations, search histories, citations, or agent traces. Consensus's citation-grounding work, Elicit's sentence-level source links, EBSCO's cited conversational answers, and Scholar Quick Read's structured treatment of methods and considerations all move in this direction. [50]

A third extension should add "corpus boundary" as a dimension outside the workflow. EBSCO, ProQuest, JSTOR, Web of Science, Google Scholar, Consensus, OpenAlex, and Elicit do not search the same scholarly universe. Full-text availability also differs from metadata availability. Zhao emphasizes licensed-platform boundaries; Tay emphasizes differences among indexes and publisher partnerships. Corpus composition therefore belongs in any evaluation model alongside the visible AI features. [51]

The attached figure is therefore an excellent map of first-generation database AI integration. The August 2026 developments suggest a second-generation figure needs loops, external connectors, and an explicit evidence-verification layer.

What researchers and librarians should evaluate now

The combined evidence from Tay, Zhao, Google Scholar, EBSCO, Elicit, Scite, OpenAlex, Consensus, and the OpenScholar paper suggests a change in evaluation practice. Asking "Does this database have AI?" has almost no analytical value in 2026.

The more useful questions concern the exact transformations between a research question and the evidence eventually presented to the researcher. [44]

Corpus coverage should come first. Which journals, books, conference proceedings, dissertations, preprints, repositories, guidelines, patents, or other sources enter the index? Does the system have metadata only, abstracts, open full text, licensed full text, or publisher-provided content? Can the AI use material available through the researcher's institution? A synthesis system's ceiling is determined partly by what it can retrieve. [52]

Retrieval architecture comes next. Researchers should determine whether a system performs lexical search, Boolean translation, vector semantic search, citation traversal, hybrid retrieval, multiple generated subqueries, reranking, or some combination. Terms such as "AI Search" or "Research Assistant" conceal these distinctions. Zhao and Tay both demonstrate why feature-level descriptions are more informative than product labels. [53]

Search transparency should receive separate evaluation. Does the system expose the generated Boolean query? Does it show AI-generated subqueries? Can the researcher inspect retrieved candidates before synthesis? Can each retrieval step be rerun? Consensus exposes more of its Deep Search history than many competitors, while Scholar Labs exposes considerably less of its query-generation process. EBSCO exposes its generated Boolean expression, although users do not directly edit it within the AI workflow. [54]

Generation transparency is another dimension. A citation next to an AI sentence is helpful, but the researcher also needs to know whether the source supports the claim. Passage highlighting, exact excerpts, page references, and claim-level grounding offer stronger verification than a citation to an entire document. Zhao found differences among EBSCO, ProQuest, and JSTOR on this point, while newer systems such as Consensus and Elicit place increasing emphasis on evidence-linked output. [55]

Reproducibility deserves formal testing. A practical evaluation can run the same query repeatedly, record titles and rankings, compare overlap at several cutoffs, and repeat on different days. Tay's experiment shows why this matters and also demonstrates the importance of reporting test limitations. Researchers conducting systematic work should preserve prompts, dates, filters, generated subqueries where available, included-record sets, and the tool and mode used. [23]

Search quality and answer quality should remain separate metrics. A tool may retrieve strong evidence and produce a weak summary. Another may write an excellent answer from an incomplete evidence set. Another may retrieve a broad candidate pool but rank the best papers poorly. Evaluating only the final prose collapses distinct errors into one subjective judgment. The OpenScholar research explicitly evaluates retrieval, answer correctness, and citation quality as related but separable components. [29]

Libraries should also distinguish learning support from research automation. EBSCO's query translation can help students see how a natural-language concept maps into Boolean synonyms. Google Scholar Quick Read can help a student examine what a paper contributes and what methodological considerations matter. Document Q&A can help locate passages. These features can strengthen information literacy when instructors make the transformations visible. They can weaken learning when generated output replaces inspection of the underlying source. Zhao and Tay both raise versions of this concern. [42]

For advanced research, a sensible 2026 workflow is increasingly hybrid. Begin with a transparent structured search when the project requires reproducibility. Add semantic or AI deep search to find terminology and papers the structured strategy may miss. Use citation searching to expand from strong seed papers. Use document-level AI for relevance assessment and navigation. Use multi-document synthesis for hypothesis generation, comparison, or orientation. Then verify substantive claims in primary sources and document the AI-assisted stages separately from reproducible database searches. This workflow follows the complementary strengths and limitations identified by Zhao and Tay rather than treating one AI system as a replacement for the full search process. [56]

For systematic or scoping reviews, researchers should resist a common category error: equating a high-quality ranked list with comprehensive retrieval. Semantic ranking optimizes relevance ordering, while systematic evidence retrieval often prioritizes recall and explicit documentation. Scholar Labs, Consensus, Web of Science Smart Search, and similar systems can contribute useful supplementary searches, but their role needs to be reported in a way reflecting dynamic ranking and hidden AI operations. [57]

For exploratory reviews, teaching, research ideation, and unfamiliar fields, the benefits are larger. AI can reduce the time required to learn terminology, locate representative studies, understand how a paper addresses a question, and move iteratively among related concepts. These tasks match the strengths Zhao observes in database-embedded AI and the design objectives Google describes for Scholar Labs and Quick Read. [58]

The emerging model of academic search

The three sources supplied in the prompt, supplemented by the August product releases and research evidence, point toward a new model of academic search in 2026.

The old model was approximately:

research question → researcher constructs query → database retrieves records → researcher screens → researcher reads → researcher synthesizes.

The emerging model increasingly looks like:

research objective → AI interprets or decomposes objective → multiple lexical, semantic, and citation searches → machine reranking and relevance assessment →

interactive document interrogation → iterative evidence expansion → grounded synthesis → source verification → researcher judgment.

Scholar Labs implements much of the retrieval side of this transformation. Consensus and Elicit extend through synthesis and agentic work. EBSCO, ProQuest, and JSTOR are inserting AI selectively into existing licensed database workflows. Scite, Consensus, Elicit, EBSCO, and Clarivate are also preparing for an architecture in which academic information services operate behind external AI interfaces. [59]

The role of the researcher changes in parallel. Query construction becomes less dominant as a manual activity. Evaluation becomes more important. You need to ask what corpus the system searched, how it interpreted your question, which retrieval methods it used, whether it generated additional searches, how it ranked candidates, how many documents informed the answer, whether the citations support the generated claims, and whether another run retrieves the same evidence. [60]

This is why Tay's LinkedIn observation is such a useful organizing principle for academic AI literacy. The relevant question is no longer whether an AI search interface feels impressive. The relevant questions concern mechanism, evidence, limitations, and methodological consequences. [2]

Zhao's figure captures the first phase of this change by mapping AI before, during, and after search. Tay's four-quadrant framework captures another dimension by separating quick versus deep retrieval and ranked lists versus synthesized answers. Taken together, the two frameworks form a stronger model: Zhao tells us where AI intervenes in the research workflow; Tay tells us how much searching and synthesis the system delegates to AI. [61]

August 2026 adds a third dimension: where the research interface itself resides. MCP, APIs, and research agents allow the interface to move outside the database. A future literature search may begin in ChatGPT, Claude, another research agent, an institutional research environment, or a specialized application, while several scholarly services work behind the scenes. Scite's August MCP integration, Consensus's MCP and agentic direction, Elicit's API-accessible agent, EBSCO Bridge, and Clarivate Nexus Connect all point toward this model. [62]

The resulting competitive question for academic databases becomes less "Who has the best AI chatbot?" and more "Whose scholarly corpus, retrieval system, evidence provenance, licensing framework, and machine-access layer can support trustworthy AI-mediated research?" The OpenScholar study reinforces the underlying technical reason: retrieval infrastructure and citation-grounded evidence remain central even when language models become more capable. [29]

For research and teaching, this leads to a specific interpretation of 2026. Boolean search has not disappeared. Ranked results have not disappeared. Database expertise has not disappeared. Each now operates inside a larger system containing semantic retrieval, LLM relevance judgment, RAG, document interaction, deep search, and

agents. The researcher increasingly chooses which level of automation fits the epistemic requirements of the task. [61]

For exploratory discovery, higher automation can save substantial effort. For reading support, grounded document interaction offers clear value. For multi-source synthesis, agents can accelerate comparison and orientation. For systematic evidence work, transparency, reproducibility, corpus coverage, recall, and independent verification remain central methodological requirements. Zhao's recommendation to treat current database AI as guided starting points and Tay's evidence on run-to-run instability make this distinction especially important. [42]

The central change in academic search during 2026 is therefore not the replacement of databases by AI. It is the reallocation of research tasks between the researcher, the retrieval system, the generative model, and the research agent. August 2026 provides unusually clear evidence of this transition: EBSCO integrates conversational and document AI into a conventional aggregator; Google Scholar combines deep AI retrieval with query-focused reading; Consensus links retrieval, citation networks, synthesis, and agents; Elicit expands research agents across heterogeneous evidence; Scite lets general AI systems call a scholarly citation service; and OpenAlex exposes semantic scholarly retrieval directly to software and agents. [63]

For AI literacy in higher education, this means the teaching target needs to move beyond prompt writing. Academic AI fluency increasingly requires understanding retrieval architecture, corpus boundaries, semantic matching, query decomposition, reranking, RAG, citation grounding, agentic tool use, reproducibility, and the difference between finding evidence and synthesizing evidence. The sources examined here make one point especially clear: in 2026, evaluating the research process behind an AI answer matters at least as much as evaluating the answer itself. [64]

---

[1] [6] [7] [11] [16] [21] [22] [23] [25] [26] [32] [35] [37] [40] [43] [47] [48] [49] [54] [57] [59] How Is AI Transforming Academic Search? | Katina Magazine

https://katinamagazine.org/content/article/resource-advisor/2026/how-is-ai-transforming-academic-search

[2] [20] [41] [44] [53] [60] [64] If you like my Katina pieces on AI academic search, I can also highly recommend this Katina piece by Aster Zhao of HKUST covering what aggregators such as EBSCO, ProQuest and JSTOR are actually doing… | Aaron Tay

https://www.linkedin.com/posts/aarontay_if-you-like-my-katina-pieces-on-ai-academic-share-7504502586104860672-FOMo/

[3] [18] [19] [28] [30] [36] [39] [42] [45] [51] [52] [55] [56] [58] [61] Do The New AI Features in These Familiar Databases Actually Help Researchers? | Katina Magazine

https://katinamagazine.org/content/article/resource-advisor/2026/new-ai-features-ebsco-proquest-jstor

[4] [8] Introducing Elicit Research Agent: Rigor that powers decisions - Elicit

https://elicit.com/blog/introducing-elicit-research-agent

[5] [13] Google Scholar Blog: Quick Read: see how a paper answers your question

https://scholar.googleblog.com/2026/08/quick-read-see-how-paper-answers-your.html?utm_source=chatgpt.com

[9] [46] Consensus Outperforms Frontier Models in Scientific Search - Consensus: AI Search Engine for Research

https://consensus.app/home/blog/consensus-outperforms-frontier-models-in-scientific-search/?utm_source=chatgpt.com

[10] [27] Semantic Search – Querying | OpenAlex Help Center

https://help.openalex.org/api/semantic-search/?utm_source=chatgpt.com

[12] AI Insights

https://connect.ebsco.com/s/article/Generative-AI-Insights-Beta-Summaries?language=en_US&utm_source=chatgpt.com

[14] [31] [62] August 2026 Release Notes

https://scite.ai/blog/august-2026-release-notes

[15] [63] EBSCO Information Services Introduces New AI-Assisted Research Features in EBSCOhost - News | EBSCO

https://about.ebsco.com/news-center/press-releases/ebsco-introduces-new-ai-assisted-research-features?utm_source=chatgpt.com

[17] [24] [38] Google Scholar Blog: 2026

https://scholar.googleblog.com/2026/?utm_source=chatgpt.com

[29] Synthesizing scientific literature with retrieval-augmented language models | Nature

https://www.nature.com/articles/s41586-025-10072-4

[33] [50] What's Changed in Consensus (Summer '26) — Consensus: AI for Research

https://consensus.app/home/blog/what-has-changed-in-consensus-summer-26/?utm_source=chatgpt.com

[34] New Features and Usage-Based Pricing - OpenAlex blog

https://blog.openalex.org/openalex-api-new-features-and-usage-based-pricing/?utm_source=chatgpt.com