

Reference Services and Instruction

Rebecca Graff, Col. Ed.

Haven't We Been Here Before?

The Rapid Proliferation of AI-Assisted Research Discovery Tools

Anna Marie Johnson

Coming soon to a reference desk or consultation near you: LitMaps, Keenious, Scite, Elicit, Consensus, Connected Papers, Research Rabbit...It seems that weekly I receive an email notice or a tip from a colleague about a "new" research tool that will help me "do research better and faster."

Artificial intelligence (AI)-assisted research tools are rapidly proliferating and are already starting to be mentioned by students and faculty in reference interactions, but a closer look reveals several things.

- Almost all of these tools represent different interfaces to the same set of data. Most draw from a common underlying dataset, typically some combination of the Semantic Scholar database and PubMed, occasionally supplemented by OpenAlex, Crossref, or arXiv.
- Almost all of them have a similar origin story:

Young computer science, mathematics, or data science graduate students struggled to use traditional search engines and databases to uncover the research papers they knew "had to be out there;" these young researchers developed a quick prototype with some open-source dataset (Wikipedia!), and people loved it. This discovery tool attracted attention and imagination and now they've hired some software engineers and have a board of famous computer science advisors.

These products each have their own spin, many of them interesting and potentially laudable. Scite tries to categorize citations as supporting or contrasting. Consensus aims to provide the scientific consensus on a research question. Research Rabbit provides the kinds of citation visualizations that previously one could only get in Web of Science. Elicit is positioning itself to assist with systematic reviews. LitMaps allows some customization of the search strategy.

But we've been here before, haven't we? (At least many of us librarians of a certain vintage.) Remember Ask Jeeves, Alta Vista, Lycos, and WebCrawler? They were all tools for searching the contents of the internet before Google came along. (Fun fact, Google was first known as BackRub—yikes!) Clearly, not all of those tools are still with us, and it seems likely that the current batch of AI-assisted tools will likely experience some similar fates; however, this could take some time to sort out. What should librarians be doing in the meantime? How do we address these tools in reference interactions, and what types of strategies do we need toward these tools broadly?

1. Educating ourselves

Educating ourselves about these tools is by far the most important task. Fortunately, there are resources to help. Ithaka is keeping an updated list on AI tools of all varieties (<https://sr.ithaka.org/our-work/generative-ai-product-tracker/>), with a section for Discovery Tools. This list has been the most helpful to me in getting a sense of the AI-assisted discovery landscape. In addition, in-depth tool comparisons are starting to appear in the literature.¹ Keep an eye out for those or, better yet, write one yourself!

These tools sometimes get confused with generative AI tools that also claim to do research, and sometimes the line is blurred. Discovery tools such as Elicit, Consensus, and Scite were built to do similar searches as traditional library tools such as Scopus or Web of Science. Generative AI is an add-on to many of these tools, and it tends to be retrieval-augmented generation (RAG). These tools will not produce as many fabricated citations as the generative AI tools such as ChatGPT because they are grounded in structured scholarly databases.

Another potential avenue to consider is the rise of Model Context Protocol (MCP) connectors from traditional library content vendors (e.g., Ebsco, ProQuest, or Oxford University Press) into large language models (LLMs) such as ChatGPT and Gemini. In this scenario, users would not need to leave the LLM environment but would still be able to (theoretically at least) access paywalled, high-quality content. This development could eliminate the need for AI-assisted research tools such as Elicit and Scite, or it could simply complicate the situation further because Scite has an MCP tool now itself. All the more reason why we need to be aware as these tools continue to rapidly develop.

2. Educating our users

If you haven't already, you will soon begin to hear about these tools from your library users via reference services. It will be important to regard these interactions as teachable moments. If users present a citation that they can't find, I gently inquire what tool they used to obtain it. In the same way that I often emphasize that Google Scholar lacks transparency in the sources being searched, I will highlight that some of these tools are similarly opaque. In addition, they can still hallucinate citations, so cautioning patrons is also an important point to convey.

For some users, suggesting these tools might even be appropriate, particularly if they are currently using Google Gemini and ChatGPT. Those AI products have modes they call "deep research," but students and faculty should be wary of using them for that purpose. In my experience, Gemini and ChatGPT scrape open websites to find information, which can result in a potpourri of sources with varying levels of scholarly verification, everything from Wikipedia to blogs to open-source, peer-reviewed journals. Even when I have asked Google Gemini to cite only a particular kind of source (e.g., peer-reviewed journals), it cannot reliably do so because it simply doesn't have access to journals behind paywalls. Even though the aforementioned AI discovery tools are often based on the same data (e.g., Semantic Scholar, PubMed), they restrict their databases to more scholarly resources. This is an important distinction to make with the students and faculty who likely won't be paying as much attention. Think of it as using the right tool for the right job. Moving a patron from ChatGPT to a tool such as LitMaps might be an easier sell than insisting that they should be using a library-licensed vendor database that requires them to think in keywords.

3. Engaging with vendors and products

I realize that we all feel pressed for time, but those of us working in reference and instruction need to prioritize talking with the companies and entities behind these products. We need to get clear answers on what journals are being included, for example. Scite is one product that has [made its list open](#). Most of the other products are quite vague on this point. Just saying your tool is mining the Semantic Scholar database is not enough. What is included in that database? We need more than just publisher icons. In addition, who is addressing the environmental impact of AI use? Who is looking out for privacy? The companies and entities that have solid answers to these questions are the tools we want to try and put some weight behind.

4. Continuing to argue for high-quality open data that covers more than just the sciences

Semantic Scholar and OpenAlex are part of an effort to make bibliographic data open, freely available, and more easily searched. Both build on the venerable traditions of PubMed, ERIC, Agricola, and other US government-funded bibliographic databases. As librarians, we should applaud these efforts while also calling for inclusion of literature from other parts of the world, particularly the Global South. In addition, these tools have all begun with the scientific literature in mind, not the humanities. Inclusion of humanities citation metadata would be welcome.

In many ways, despite all the uncertainty, this is an exciting time. These emerging discovery tools are the embodiment of the sentiment articulated by Roy Tennant in 2001: "Only librarians like to search, everyone else likes to find."² Getting students and faculty to the scholarly literature they want and need in a more expedient manner should be something we in the reference and instruction community celebrate. Digging into the nuts and bolts of how these tools work, teaching students and faculty about these new tools, and looking out for our constituents' best interests has always been and will continue to be our role.

References

1. Jeanine Mary Williamson and Peter Fernandez, "'Through the Looking Glass: Envisioning New Library Technologies' Academic Search Using Artificial Intelligence Tools," *Library High Tech News* 42, no. 4 (2025): 1–5. <https://10.1108/LHTN-01-2025-0014>.
2. Roy Tennant, "Digital Libraries-Cross Database Search: One-Stop Shopping," *Library Journal* 126, no. 17 (2001): 29.