

Comparison of AI-Enhanced Tools for Automating Scientific Literature Reviews

Grys M¹, Casciano R², Pieniążek I¹
¹Certara, Cracow office, Poland; ²Certara, USA

Background and Objectives

- Artificial Intelligence (AI) is transforming scientific literature review (LR) by accelerating and automating the review process.
- This study compares four commercially available AI-enhanced LR tools across various stages of the review process

Methods

Four AI-assisted LR tools (T1, T2, T3, T4) were evaluated between March and August 2024 for their performance in literature search, abstract and full-text screening, and data extraction using two live projects, one systematic and one targeted.

Results

Search and abstracts selection

- All the evaluated tools varied in functionality, each with its own advantages and limitations. However, they are most advanced in the abstract screening stage, while other features, such as data extraction, are still under development.
- The tools assessed included one developed over a decade ago and three introduced in recent years.
- Three tools utilized publicly available large language models (LLMs) with internal adjustments, while one employed a proprietary LLM (T3).
- One utilized non-generative AI (T1), and two used generative AI (T3, T4).
- One tool enabled concept-based AI-assisted searching (T1). This tool also enabled AI to automatically create a search strategy based on the research question.
- Three tools offered AI-driven abstract re-ranking, prioritizing relevant abstracts (T1, T2, T3).
- Two tools offered AI-assisted abstract screening, with AI acting as a second reviewer after training (T1, T2).
- One tool demonstrated a nearly tenfold lower false-negative rate than the others (T1).
- One tool automatically extracted all PICOS elements from abstracts and provided live AI performance statistics, expediting the identification of relevant papers (T1).
- Another tool categorized abstracts by answering yes/no questions, significantly reducing screening time (T2).


certara.com

Overview of the AI tools performance in abstracts selection


Functionality	T1	T2	T3	T4
AI abstracts categorization	H	M	N	N
AI abstracts extraction	H	N	N	N
AI-enhanced abstracts selection (re-ranking)	H	M	M	N
AI-screener (AI replacing one of the reviewers)	H	M	N	N
Autonomous abstract selection by AI	N	N	N	N
Deep-dive abstracts AI analysis	H	N	N	N

Legend

- T1, T2 etc. – AI tools
- L – low performance
- M – medium performance
- H – high performance
- N – no option



While AI holds great potential to automate the review process, it should complement, not replace, human reviewers to maintain accuracy and reliability



Want to learn more?
<< Scan Here

Full text selection, Data extraction

- One of the tools (T1) enabled automatic matching of the imported PDF file to references by using AI for PDF content extraction.
- AI-supported full-text selection is not currently available in any of the analyzed tools. Work on this feature is still ongoing.
- Three tools supported AI-driven data extraction from PDFs, with non-generative AI (T1) outperforming generative AI (T3, T4) in accuracy.
- AI answered questions (prompts) prepared by the reviewer based on data available in the full text.
- Importantly, reviewers maintained control over data selection and extraction at every stage. AI-assisted table extraction and critical appraisal were under development in all tools.
- One of the tools (T1) offered a free trial version that allows conducting a single literature review.

Overview of the AI tools performance in data extraction

Functionality	T1	T2	T3	T4
AI-powered extraction model	H	N	M	M
Highlighting extracted data in the publication	H	N	H	H
Extraction from multiple languages	H	N	N	H
Extraction from tables	N	N	N	N
Extraction accuracy	H	N	L	L

Discussion

- AI-enhanced LR tools effectively streamline targeted reviews, identifying key publications rapidly.
- AI functionalities have the potential to reduce abstract selection time by more than 50% and data extraction time by 70–80%. Further advancements, such as automated network meta-analysis (NMA) and critical appraisal, could further accelerate the literature review process.
- Caution is advised in systematic literature reviews (SLRs) to ensure compliance with regulations.
- Official guidelines/regulations should be developed to define the possibilities of using AI in literature reviews conducted for official documents, e.g., the publication of a systematic review where selection is partially performed by AI.