

CA2079

動向レビュー

## 生成AIを用いた文献調査ツール

やだしけんたろう  
矢田竣太郎\*

### 1. はじめに

人工知能 (AI) 技術が人々の生活や日々の業務に取り入れられて久しく、調査や研究の場面でも有用なツールが多数登場している。特に生成 AI が登場してから動きは加速しており、図書館関係者でもその語を見聞きしない日がないほどだろうと思われる。本稿では、図書館現場でも使われうる、生成 AI を用いた文献調査ツールに着目して、現状と展望を述べる。特に文献調査は研究支援業務で重要なため、大学図書館の文脈に論の重きを置くことにしたい。また、筆者は言語処理技術を主たる手法に据える図書館情報学研究者を自認するとともに、生成 AI の実体ないし基盤である大規模言語モデル (LLM) の安全性を検討する研究者グループ (LLM-jp 安全性検討ワーキンググループ<sup>(1)</sup>) に所属している。本稿と同じ目的の既報<sup>(2)</sup>がある中で、本稿は技術と安全の観点を付加できたらと思う。

### 2. 技術的背景

AI 技術そのものは検索エンジンにすら古くから組み込まれており、目新しいものではない。検索キーワードをシステム側で自動で拡張したり<sup>(3)</sup>、文献とクエリの類似度を生成 AI 並みの言語モデルで判定したり<sup>(4)</sup>、文献をさまざまな観点であらかじめラベル付けしておく<sup>(5)</sup>といった、言語処理ベースの技術の実用には多くの実績があったのである。

しかし生成 AI を用いたものは、今までにないレベ

ルの脚光を浴びているように感じられる。従来技術が文章の類似度判定やラベル付けといった言語理解タスクに基づいていたのに対し、文章そのものを産生する生成 AI の採用によって、自然文での対話をユーザーインタフェース (UI) にできたことが、その大きな要因と考えられる。「箇条書きでまとめて」、「短く要約して」、「表にして」といった指示に、それまでの質問応答システム<sup>(6)</sup>の比ではない高い品質で、生成 AI は従うことができる。生成 AI 以前の AI 調査ツールはキーワードをベースとした従来の検索エンジンと UI 上は大差なく、使用に一定のスキルが求められたことは否めない。自然文で検索できても、画面上の検索結果を読み解くという暗黙の IT スキルは必要だった。それが生成 AI を用いる場合、検索結果をも AI が要約した文章として表示される。ユーザはツールの使い方を学習する必要がないというわけである (図 1)。

生成 AI は基本的に、入力テキストを別のテキストへと変換あるいはその続きを出力する LLM という機械学習モデルを、質問応答に即するように訓練したものである。モデルは本来、訓練した際の情報にしか触れていないので、再学習させない限り新しい情報に基づいた応答は困難であるが、「大規模」である LLM を再学習してモデル更新をするのは多大なコストがかかるので避けたい。ここで LLM は、入力されたテキストに新しい情報が含まれる場合、それをあたかもその場で学習したかのように応答できる (文脈内学習<sup>(7)</sup>) ことから、Retrieval-augmented Generation (RAG)<sup>(8)</sup> という手法が考案された。ユーザが入力する質問文に、関連文書の中身を結合することで、その文書の内容を踏まえた応答をさせる。これにより、学習時に持ち得なかった新情報や、特定のドメインに特化した応答を、より高い精度で出力できるようになった。「検

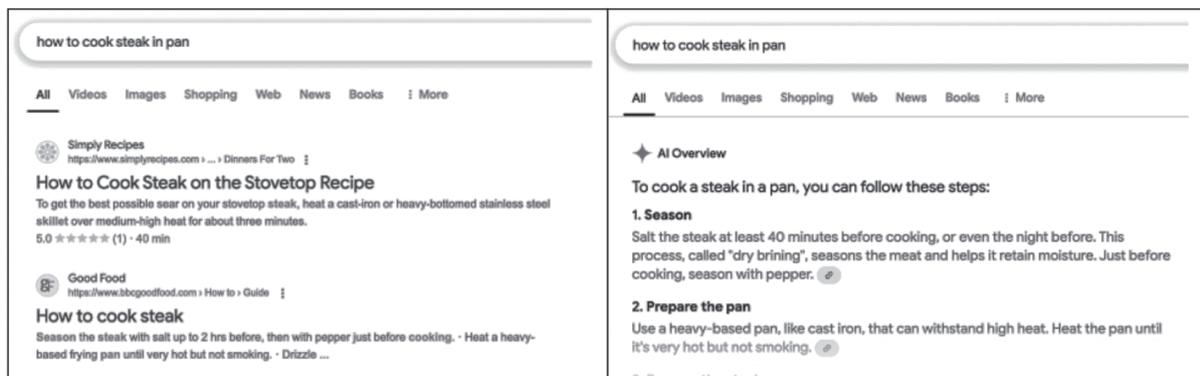


図1 Google の検索結果 UI の比較。左は従来の表示方式、右は生成 AI によるサマリ表示。従来方式は Web の仕組みや URL の理解を前提としているが、AI サマリは識字力のみ依存する。

\* 筑波大学図書館情報メディア系

索できる生成 AI」と呼ばれるものは、この RAG を組み込んでいる。当然、入力テキストに関連した文書を見つける検索部分の実装が重要であり、入力テキストから検索エンジン向けのキーワードを生成してから検索、入力テキストと意味的類似度が近い文書を言語処理技術で直接検出、といった方式がある。次節で紹介するような商用ツールでは、検索を繰り返して品質を向上しているような挙動もよく見られるが、その内部実装は当然非公開である。この点、学術文献データベース Semantic Scholar を提供する Ai2 社は RAG を用いた文献検索エンジンのオープン実装 OpenScholar を公開<sup>(9)</sup>しており、業界スタンダードと思われるシステム構成と採用技術を窺い知ることができる。

### 3. 代表的なツール

執筆時点で注目を集めているツールを、主たる用途に着目して大きく 2 つに分類し、紹介する。なお、ほとんどのツールは英語での使用を前提としており、日本語でも一応の動作はするものの、公式にサポートされていないので、有用性に劣る点は注意が必要である。

#### 3.1 検索・探索の支援

前述のような、検索結果を取り込んで生成 AI が要約したり、その内容について質問応答（フォローアップ）したりできるツールが、商用レベルで多数リリースされている。

- ・ Elicit<sup>(10)</sup> : 「Do clinical language models compromise patient privacy?」といった自然な疑問文で研究主題 (Research Question : RQ) を入力すると、関連性の高い文献を検索したうえで、表形式にまとめて表示する。自動構築される表には各論文の要約や、核となる主張、調査・実験論文の場合はそのデータ統計などがカラム (列) として用いられる。全文がアクセス可能な論文に対し、生成 AI で要約する作業を繰り返すことで実現しているとみられる。
- ・ Consensus<sup>(11)</sup> : RQ を入力すると、それが支持されるかどうかに基づいて論文を整理する。検索結果の何%が RQ を支持 (Yes) するかがグラフィカルに表示され、生成 AI が論文内容をもとに RQ に回答するレポートも出力される。ただしデフォルトではトップ 10 記事のみを分析するので、実用上の注意が必要だろう。
- ・ Scite Assistant<sup>(12)</sup> : Scite は文献が引用されている文脈に応じて文献検索できるサービスである。文献は肯定的にも否定的にも引用されうることから、本文

中でどのように引用されているかに基づいて学術論文を分類した引用文脈データベースを Scite 社は構築している。Scite Assistant はその検索結果の要約や質問応答に生成 AI を用いたものであり、文献のアカデミックな評判を調べるのに適している。

- ・ Perplexity (Academic Focus)<sup>(13)</sup> : ChatGPT に先んじて公開されていた検索エンジンで、LLM を用いた「対話形式の検索」を世に提案した最初期のサービスである。Perplexity は検索対象のドメインを絞り込む Focus という機能を提供しており、この選択肢の中に学術文献に特化した Academic がある。
- ・ Gemini Deep Research<sup>(14)</sup> : 2024 年 12 月に Google 社の生成 AI である Gemini に追加された機能で、学術文献に特化して調査し、複数ページものレポートを出力できる。RQ を入力すると、Gemini はまず検索の計画を立て、ユーザに確認を求める。検索計画はチャット上で変更を指示できるほか、検索が実行されている間はこういったサイトを調査しているかリアルタイムで確認できる。Google 社らしい連携として、生成されたレポートは Google ドキュメントとして編集できる。
- ・ GPTs で公開されている研究用途の “GPT” : OpenAI 社の ChatGPT<sup>(15)</sup> は、ユーザが独自のプロンプトや RAG システムを設定したカスタム版の ChatGPT を公開できる GPTs というプラットフォームを提供しており、この中に文献調査向けに調整されたものがいくつか見られる。個人も作成可能なので悪意あるものも流通しうるが、対話の中で何がどのように検索されるかは事前に表示されるので、プライバシーや機密の漏洩をユーザは防止できる。なお、ChatGPT 本体にも Web 検索できる機能が追加されている。

#### 3.2 読解の支援

論文記事を指定すると、その内容について要約したり質問したりできるツールもある。上記に挙げたツールにも同様の機能を有するものがあるが、ここに挙げるのはそれをメインとするものである。

- ・ SciSpace<sup>(16)</sup> : 論文 PDF をアップロードすると、PDF とチャット欄が縦 2 分割で表示され、論文内容について質問しながら本文を読み進められる。PDF ビューワ側で本文や数式を選択すると、その要約や説明を要求できる。SciSpace では文献検索のほか、投稿規定に合わせた文献フォーマットの変更などの

ツールも多数提供されている。

- ・ NotebookLM<sup>(17)</sup> : Web ページを指定したり PDF をアップロードしたりすると、それらの文書群について質問できる Google 社のサービスである。学術文献に特化しているわけではないが、論文に対しても遜色なく動作する。生成 AI (Gemini) の回答には文献の該当箇所に飛べるリンクが埋め込まれており、裏付けをとりやすい。
- ・ Illuminate<sup>(18)</sup> : 執筆時点では Google 社のベータ版サービスだが、論文の URL や PDF を指定すると、その内容を 15 分程度のラジオ番組 (あるいはポッドキャスト) 風にまとめた音声データを出力する。同様の機能は NotebookLM にも搭載されているが、Illuminate は論文に特化し、番組の形式 (話者の人数や専門性など) もより細かく設定できる。

## 4. 展望

### 4.1 ツール発展の方向性

直近で実現の可能性が高いのは、より複雑な調査を自動化することである。ショッピングサイトのような複雑な Web アプリを操作できる AI も各社から発表されている<sup>(19)(20)</sup>。プログラムから呼び出しやすい窓口 (API) が用意されたデータベースだけが上記ツールの検索対象であったが、今後は API 未整備の古いデータベースや Web サイトをも人間がブラウジングするように操作して情報収集できるようになる。PDF 形式の表でしか統計データを公開していないような情報源に思い当たる読者も多いはずだ。サイトトップからページをたどって PDF ファイルを見つけ、該当の表を探すといった一連の操作を自律的に実行するシステムも近いうちに使えるようになる。

もう一つは、研究業務そのものの自動化である。Sakana AI 社による、機械学習分野の研究を自動遂行する AI<sup>(21)</sup> は話題となった。情報学分野では実験もコンピュータ上で行えるため早期に実現したと言える。今後、例えば生体を使った研究でも、コンピュータ経由での実験機器の自動操作が充実すれば展開可能と見られる。フィールドワークなどの対人調査についても将来、AI ロボットとして身体を得た場合には不可能ではない。「人間がその手で実施する研究」が指す範囲が小さくなっていく中で、研究支援のあり方も変わるだろう。

### 4.2 研究支援業務での使いどころ

現在の生成 AI ツールは、言語モデルが保持する知識と、上位の検索結果たかだか数十件に基づいて、回

答や要約を行う。職業研究者やレファレンスに長けた図書館員にとって、そのような調査が何を意味するかは明白である——実務においては無用の長物と思われる読者も少なくないだろう。実際、AI 回答の正否と文献内容の照合に時間が取られるだけに見える。しかしそれは、伝統的な文献データベースをいかに検索し、得られた文献をいかに読み解けばよいか分かるという経験知があってこそその認識だ。検索に使うべきキーワードや、内容を整理するための専門的観点がわからない未知の分野について調べるとき、伝統的なパスファインダー等のレファレンスツールでは成し得なかった形で見つけられるようになったことは喜ばしい。1 パラグラフほどかけて説明した疑問について「それは XX 学の専門用語で YY という現象ですね」という回答を得るためには、社会的資本が必要だった。上述のツールはこの体験を万人にもたらせる点で有意義である。しかし当然、生成 AI による滑らかな回答には「分かった気になる」ことを加速させる向きはあり、初学者ほど使用に注意を要するだろう。複数ツールの結果を比べる、確証バイアス<sup>(22)</sup>を避ける調べ方を学習しておくなど、生成 AI ベースかどうかによらず文献調査の際に留意すべき点は、引き続き押さえておく必要がある。

### 4.3 大学図書館員としての向き合い方

跋扈する生成 AI ツールが高等教育の現場に与える影響は大きく、各大学が使用の指針や注意喚起を発表しているところである。学内の研究支援や情報リテラシー教育を担う大学図書館は、教職員や学生に対してどのような情報提供をすべきだろうか。筆者が提案させていただけるならば、学内での利用が多いツールだけでも、その「安全性」について検証した結果をまとめてみてはどうかと思う。国内外の大学図書館が生成 AI のユーザガイドを作成している例<sup>(23)</sup>は多く見られるが、概説的な使い方案内や一般的な注意喚起に留まっているように見受けられる。ここから一歩進んで、大学図書館員自らツールを使用しながら、次のような点を評価できないだろうか。

- ・ 利用規約やプライバシーポリシーの安全性 : 生成 AI に入力した未発表の研究アイデアが、ツール開発者や提供者に流出しかねないものもある。入力情報の著作権や、第三者への提供に関する規定はまちまちだ。例えば、東京のスタートアップが提供する生成 AI ベースの検索エンジン Felo Search<sup>(24)</sup> は研究利用を謳う新機能<sup>(25)</sup>も発表しているが、本稿執筆時点で同社の規約は送受信されうるユーザの研究データに対して何ら明確な姿勢を示していない。加

Do clinical language models compromise patient privacy?

**Do Clinical Language Models Compromise Patient Privacy?**

The question of whether clinical language models compromise patient privacy is crucial in the era of artificial intelligence in healthcare. These models, including natural language processing (NLP) systems, are often trained on large datasets derived from patient records, clinical notes, and other sensitive medical data. While these systems offer transformative potential for diagnostics, clinical documentation, and decision-making, their use raises concerns about privacy, data security, and ethical compliance.

I will now retrieve recent academic research, case studies, or reviews on this topic to provide detailed evidence. Please hold on.

Talked to sider.ai ^

The following was shared:	Privacy policy
query: "Do clinical language models compromise patient privacy?"	
language: "en"	
year_start: 2018	
year_end: 2025	
sorted: true	
count: 20	

図2 文献調査 GPTs の1つ「Scholar GPT」を用いると、生成 AI が途中で検索を代行するが、その際に生成 AI が用いる検索条件も表示できる。

えて、サービス利用と同時に適用されてしまう利用規約とプライバシーポリシーへのリンクは目立たない位置にある。小規模な会社の生成 AI ツールは法的整備の優先順位が低いこともあり、「新進気鋭」のツールに飛びつくのはリスクがある。

- ・ハルシネーション注意喚起の表示方式：生成 AI を用いた文献調査ツールにおいても、いわゆるハルシネーション（誤情報の出力）が生じうる。ベースとなる LLM は理論的にハルシネーションを避けられ

ないとする研究もある<sup>(26)</sup>。図1でも紹介した Google 検索の AI サマリーが、あろうことか誤情報を最上位に表示しうるのは深刻に懸念されている<sup>(27)</sup>。上述のほとんどのツールでは「LLM は間違えるのでソースを確認せよ」といった注意喚起が各所に表示されるが、その視認性や表示方法はまちまちなので、ユーザの視点から評価することは重要になる。例えば Consensus の場合、同様の注意喚起は、生成 AI による回答レポートの末尾に小さく表示されるボタンを押して初めて表示される。

- ・調査時のバイアスへの対応：調査におけるバイアスの低減や自覚のためにシステマティックレビューのような文献調査手法が確立している。しかし本稿で紹介したツールは検索の網羅性が不透明なものも少なくない。既存の文献調査ガイドラインに照らしながら、実行された検索の詳細が表示されるか（図2）、回答に引用情報が含まれるか、引用元や原文を参照しやすいか（図3）といった観点から評価しなおすと良いだろう。

生成 AI を用いた文献調査ツールの安全性に関するこれらの調査は、学術雑誌・出版社の著作権規定や信頼度を調査する既存の取り組みの延長にあたるとは言えないだろうか。当然、各ツールは頻繁にアップデートされるので、定期的な再評価が必要になるが、主なものに限れば現実的かと思われる。そのような動きがあれば筆者としてもぜひ応援・助力させていただきたい。



図3 Elicit では、生成 AI による要約をクリックすると、要約元となる本文該当箇所がポップアップで表示される（元論文情報は筆者が匿名加工した）。

- (1) LLM-jp et al. LLM-jp: A cross-organizational project for the research and development of fully open Japanese LLMs. arXiv. 2024. arXiv:2407.03963. <https://doi.org/10.48550/arXiv.2407.03963>, (accessed 2025-01-17). LLM 勉強会. <https://llm-jp.nii.ac.jp/>, (参照 2025-01-27).
- (2) 森内泰. 文献調査に利用できる生成 AI ツールと、その使われ方. 長崎県立大学論集 (経営学部・地域創造学部). 2023, 57(2), p. 19-30. <http://hdl.handle.net/10561/1967>, (参照 2025-01-17).
- (3) Carpineto, C.; Romano, G. A survey of automatic query expansion in information retrieval. ACM Computing Surveys. 2012, 44(1), p. 1-50.
- (4) Nayak, P. "Understanding searches better than ever before". Google The Keyword. 2019-10-25. <https://blog.google/products/search/search-language-understanding-bert>, (accessed 2025-01-13).
- (5) 井上和仁, 川原綾, 藤井遥子, 大久保三四朗, 加藤真壽美. AIを活用した新しい医薬情報検索(JDream SR プロダクト・レビュー). 薬学図書館. 2022, 67(2), p. 69-73. [https://doi.org/10.11291/jpla.67.2\\_69](https://doi.org/10.11291/jpla.67.2_69), (参照 2025-01-17).
- (6) 2020年ごろの企業ホームページに見られたFAQチャットボットを思い出していただくと良い。
- (7) Dong, Q. et al. "A survey on in-context learning". Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. Florida, 2024-11-12/16. Association for Computational Linguistics, 2024, p. 1107-1128.
- (8) Gao, Y. et al. Retrieval-Augmented Generation for Large Language Models: A survey. arXiv. 2023, arXiv:2312.10997. <https://doi.org/10.48550/arXiv.2312.10997>, (accessed 2025-01-17).
- (9) Asai, A. "Ai2 OpenScholar: Scientific literature synthesis with retrieval-augmented language models". Ai2. 2024-11-19. <https://allenai.org/blog/openscholar>, (accessed 2025-01-13).
- (10) Elicit. <https://elicit.com>, (accessed 2025-01-17).
- (11) Consensus. <https://consensus.app/>, (accessed 2025-01-17).
- (12) scite. <https://scite.ai/assistant>, (accessed 2025-01-17).
- (13) Perplexity. <https://www.perplexity.ai/>, (accessed 2025-01-17).
- (14) Citron, D. "Try Deep Research and our new experimental model in Gemini, your AI assistant". Google The Keyword. 2024-12-11. <https://blog.google/products/gemini/google-gemini-deep-research/>, (accessed 2025-01-17).
- (15) ChatGPT. <https://chatgpt.com/>, (accessed 2025-01-17).
- (16) "Chat with PDF". SCISPACE. <https://typeset.io/chat-pdf>, (accessed 2025-01-17).
- (17) NotebookLM. <https://notebooklm.google.com/>, (accessed 2025-01-17).
- (18) Illuminate. <https://illuminate.google.com/>, (accessed 2025-01-17).
- (19) Pichai, S. et al. "Introducing Gemini 2.0: our new AI model for the agentic era". Google The Keyword. 2024-12-11. <https://blog.google/technology/google-deepmind/google-gemini-ai-update-december-2024/>, (accessed 2025-01-13).
- (20) Perplexity Team. "Shop like a Pro: Perplexity's new AI-powered shopping assistant". Perplexity Blog. 2024-11-18. <https://www.perplexity.ai/hub/blog/shop-like-a-pro>, (accessed 2025-01-13).
- (21) "[AIサイエンティスト]: AIが自ら研究する時代へ". sakana.ai. 2024-08-13. <https://sakana.ai/ai-scientist-jp/>, (参照 2025-01-13).
- (22) 当たりをつけられれば生成 AI に頼るまでもないレベルの事例ではあるが、この心理学用語を見つける際の ChatGPT 使用例を、この小節の論旨に沿ったメタ的な例として紹介する。ChatGPT. <https://chatgpt.com/share/6784b64e-82fc-800d-ba30-027dcfaab94f>, (accessed 2025-01-17).
- (23) Harvard Library. "Artificial Intelligence for Research and Scholarship". Research Guides. <https://guides.library.harvard.edu/c.php?g=1330621&p=9798082>, (accessed 2025-01-13).
- (24) Felo. <https://felo.ai/search>, (accessed 2025-01-17).
- (25) Felo Search Tips Buddy. "How Felo 3.0 Search Agent Can Revolutionize Your Academic Research". Felo Search. 2025-01-11. <https://felo.ai/blog/search-agent-academic-research/>, (accessed 2025-01-14).
- (26) Banerjee, S.; Agarwal, A.; Singla, S. LLMs Will Always Hallucinate, and We Need to Live With This. arXiv. 2024, arXiv:2409.05746. <https://doi.org/10.48550/arXiv.2409.05746>, (accessed 2025-01-17).
- (27) "検索結果の概要を自動生成する「AI Overviews」が「誤情報」を表示、グーグルが認めた問題の中身". WIRED. 2024-06-01. <https://wired.jp/article/google-ai-overview-search-issues/>, (参照 2025-01-13).

[受理: 2025-02-10]

Yada Shuntaro  
Literature Review Tools Using Generative AI