On the Use of Generative AI for Literature Reviews: An Exploration of Tools and Techniques

Peter Mozelius[1,2] and Niklas Humble[2,3]
[1]Mid Sweden University, Sweden
[2]University of Gävle, Sweden
[3]Uppsala University, Sweden

peter.mozelius@miun.se
niklas.humble@it.uu.se

Abstract: To carry out a literature review often involves hard and tedious work. There is a tradition of using facilitating tools, that extended to the AI field in 2018 when iris.ai appeared. Today, in the emerging field of Generative AI tools based on Large Language Models, there has been rapid development of new literature search tools and approaches. This study has the aim of exploring this vast array of Generative AI tools, in a literature study where some of the found tools were used to facilitate the selection of relevant publication. Three research questions guided this study: RQ1) "What Generative AI tools can be found in literature?", RQ2) "Which of these tools could be of use in the literature review conducted in this study, and how?", and RQ3) "What are the ethical aspects of using Generative AI tools in literature studies?" The approach has been a scoping review, built around a search that combined the keywords: "AI supported", "AI generated", "AI based" and "Literature review". An initial result set was filtered with inclusion exclusion criteria in a strive for an interesting quality answer to the research questions. However, most publications that passed the filtering lacked any potential to contribute to answer the research questions. The most interesting finding in the first search was a hint about the new feature 'Scopus AI'. A new search with the Scopus AI tool resulted in a small but very relevant set of publications. These publications were analysed in a deductive inductive thematic analysis, and primary sorted into the categories of: 'Generative AI Tools', 'Supportive AI Techniques', and 'Ethical Issues'. Findings indicate that there is a wide variety of tools that can facilitate the skimming process of a literature, and to provide adequate summaries of retrieved publication. However, authors recommendation is to keep the tools on the facilitating support level, and that the main analysis and conclusion should be human conducted. With this, rather traditional approach, researchers will have clearly less ethical issues to consider. Finally, the ethical aspects of Generative AI tools in research ought to be investigated more in detail, in a separate future study.

Keywords: Literature review, Research methods, Generative AI, AI supported research, AI tools

1. Introduction

The idea of automatising literature reviews is older that the current spring of Generative AI (Jonnalagadda, Goyal and Huffman, 2015; Marshall, 2016). However, the recent breakthrough for Generative Artificial Intelligence (GenAI) has opened new possibilities for faster and more efficient literature review processes. GenAI tools are based on deep learning models to generate human-like content in response to prompted instructions or questions (Lim et al., 2023). On the other hand, the generation of human-like content raises ethical consideration, and as highlighted by Buriak et al. (2023), care and cautiousness seem like relevant recommendations.

This study had a focus on how GenAI has changed the landscape for literature reviews. The research approach was a scoping literature review that partly was experimentally carried out with a GenAI tools, but with the main thematic analysis conducted traditionally. The aim of the study was to explore and discuss tool, techniques for GenAI supported literature reviews. Moreover, the ethical aspects of AI generated literature reviews have been analysed and discussed. The tree research questions guided the study were: RQ1) "What Generative AI tools can be found in literature?", RQ2) "Which of these tools could be of use in the literature review conducted in this study, and how?" and RQ3) "What are the ethical aspects of using Generative AI tools in literature studies?"

2. Method

The research approach for this study was an experimental scoping literature review to retrieve a set of publications that have a potential to answer the research questions. In a strive for peer-reviewed quality publications only the authors' choice was the Elsevier Scopus research database. The keywords presented in the abstract were combined to the following search string: (TITLE-ABS-KEY ("Literature review") AND TITLE-ABS-KEY ("AI supported") OR TITLE-ABS-KEY ("AI generated") OR TITLE-ABS-KEY ("AI based")). This gave a result set with 365 publications, where all were published during the last five years, involving AI, and of high quality. Initially, this looked promising, but the problem was that there was a very weak focus on literature reviews with generative AI tools and techniques. After a first skimming, most publications lacked the real potential to contribute to answer the research questions.

However, in between some of the result listings there was a hint about the new feature 'Scopus AI'. A new search was carried out with the search question depicted in Figure 1 here below.

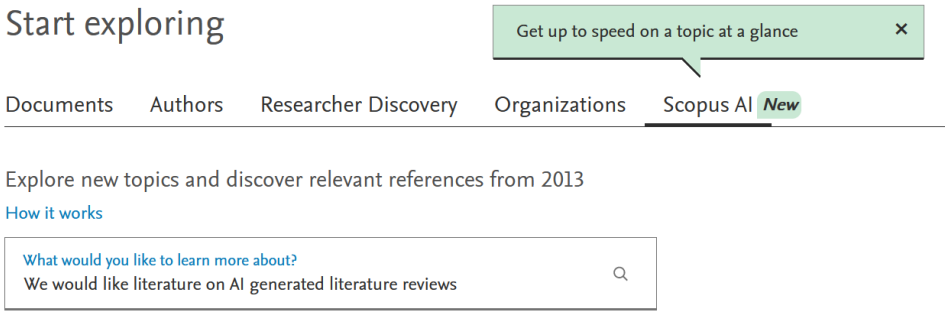

Figure 1: The second search with the Scopus AI tool

The result set from the Scopus AI search was highly relevant with a result set containing 10 publications from 2021 – 2023. Moreover, there was an additional result set with 5 'Foundational papers'. The 11 publications that passed the inclusion and exclusion criteria in Table 1 below are listed in Table 2 in the beginning of the 'Results and discussion' chapter.

Table 1: Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Published in 2019–2023	Published before 2019
Publications in English or Swedish	Other languages
Peer-reviewed	Not peer-reviewed
Direct contribution to the RQ	No contribution to answering the RQ
	Duplicate
	Only abstract available

Main findings in the 11 publications regarding 'Tools', 'Techniques' and Ethics' were extracted by the Scopus AI feature as depicted in Figure 2 below. The summary provided by Scopus AI could be seen as an alternative to the human abstract and keywords screening that is a part of most literature reviews. Results from the AI generated screening have later been cross-checked by the authors in the main analysis, and far from all summaries were 100% correct.

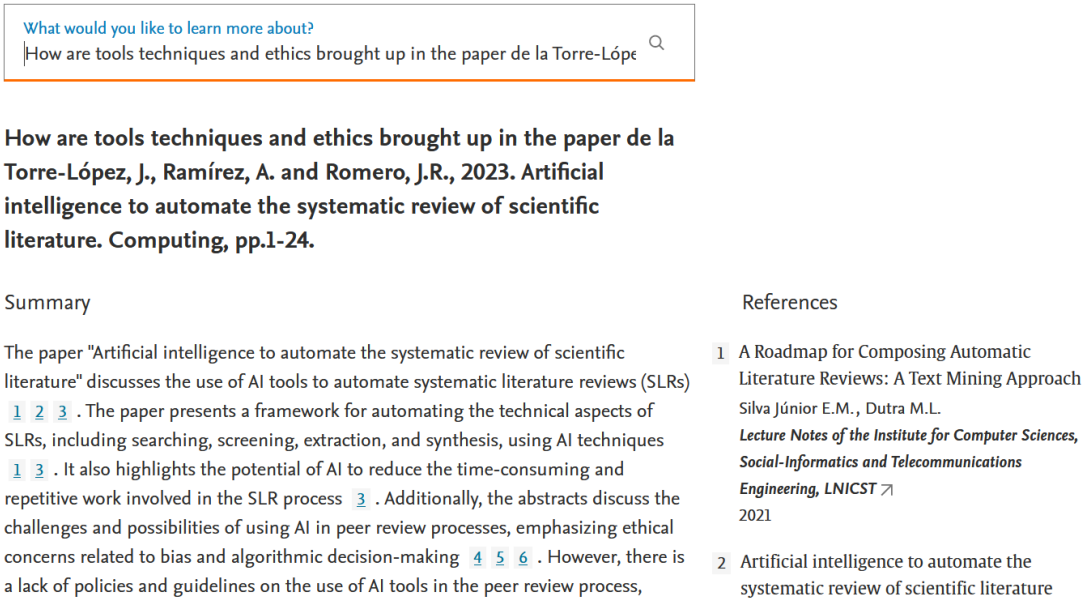

Figure 2: An automatised abstract screening with Scopus AI

After discussions among the authors, the final, more detailed and thorough analysis of the publications, were conducted as a traditional thematic analysis. Main reasons for the choice of a traditional and manual main analysis were quality and accountability. The tested GenAI feature looks promising but should not be seen as a mature and well-evaluated tool.

2.1 The Manual Data Analysis

The remaining set of publications were analysed in a deductive inductive thematic analysis, and primary sorted into the pre-defined categories of: 'Generative AI Tools', 'Supportive AI Techniques', and 'Ethical Issues'. Themes that were decided in a minor initial literature search as a pre-phase for the six-phase process for thematic analyses described by Braun and Clarke (2006). In several publications on creative use of GenAI in various fields, there themes of AI tools, AI techniques and AI ethics were brought up (Flick and Worrall, 2022; Mijwil et al., 2023; Pant et al, 2023). Themes that the authors found interesting to bring into the exploration of AI generated literature reviews, and to use as a sieve for the thematic analysis described below.

In the initial phase of the thematic analysis the authors' focus was on getting familiar with the publications, in an iterative reading, rereading and close reading of the selected the publications. In the second phase, the systematic analysis was carried out when authors created preliminary codes and sub-codes and compared them to the AI generated codes in Table 2. In the following third phase, codes and subcodes were sorted into the pre-defined themes, combined with annotations for new emerging themes. Phase four consisted of a collaborative revision involving a critical discussion between the authors about the themes, and potentially emerging new themes or sub-themes. After the fifth phase where the analysis was fine-tuned and cross-checked, results were with the thematic analysis idea of emphasising findings that contribute to the answer the research questions. (Braun & Clarke, 2006)

3. Results and Discussion

The initial AI supported filtering and skimming steps resulted in the 11 publications listed below.

Table 2: The final result set with publications in chronological order

Publication	Tools	Techniques	Ethics
Marshall, I.J. and Wallace, B.C., 2019. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Systematic reviews*, *8*, pp.1-10.	Based on the provided abstract, there is no direct information about the specific tools, techniques, and ethics discussed in the paper "Toward systematic review automation: a practical guide to using machine learning tools in research synthesis"	Based on the provided abstract, there is no direct information about the specific tools, techniques, and ethics discussed in the paper "Toward systematic review automation: a practical guide to using machine learning tools in research synthesis"	Based on the provided abstract, there is no direct information about the specific tools, techniques, and ethics discussed in the paper "Toward systematic review automation: a practical guide to using machine learning tools in research synthesis"
van Dinter, R., Tekinerdogan, B. and Catal, C., 2021. Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology*, *136*, p.106589.	Tools based on machine learning techniques, such as SVMs and Naïve Bayes. More a description of deep learning techniques than of tools	Selection of Primary Studies, Natural Language Processing, Text mining. The proposed modified SLR process includes automation and bibliometrics to improve the efficiency and effectiveness of literature reviews	Publication bias, the phenomenon that authors are more likely to publish positive results than the negative results of their research
da Silva Júnior, E.M. and Dutra, M.L., 2021. A roadmap toward the automatic composition of systematic literature reviews. *Iberoamerican Journal of Science Measurement and Communication*.	MySLR and Porifera RobotReviewer	Searching, screening, extraction, and synthesis A framework for automatically creating systematic literature reviews by combining and placing existing techniques in stages where they possess the greatest potential to be useful	The specific ethical considerations in the automatic composition of systematic literature reviews are not explicitly addressed
Wagner, G., Lukyanenko, R. and Paré, G., 2022.	ASReview, Litbaskets, RevMan, RobotReviewer	Step 1: Problem formulation, Step 2:	Transparency and replicability.

Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, *37*(2), pp.209-226.	WebPlotDigitizer and Graph2Data for extracting data from statistical plots	Literature search, Step 3: Screening for inclusion, Step 4: Quality assessment, Step 5: Data extraction, Step 6: Data analysis and interpretation	Contributions require human interpretation and insightful syntheses, as well as novel explanation and theory building.
Müller, H., Pachnanda, S., Pahl, F. and Rosenqvist, C., 2022. The application of artificial intelligence on different types of literature reviews-A comparative study. In *2022 International Conference on Applied Artificial Intelligence (ICAPAI)* (pp. 1-7). IEEE.	ChatGPT	Searching, screening, extraction, and synthesis, and proposing a pipeline for automatic creation of literature reviews	
Burger, B., Kanbach, D.K., Kraus, S., Breier, M. and Corvello, V., 2023. On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management*, *26*(7), pp.233-241.	ChatGPT Python, including frameworks like Pandas, PyTorch, or TensorFlow	(1) The use of AI in the preparation phase (2) Use of AI in the research initiation phase (3) Use of AI in the data analysis phase	The researcher is always fully responsible for the results they get from AI models. The researcher must always keep these points in mind and exercise caution when working with AI, as the field is constantly evolving and new challenges are always emerging.
Tlili A, Huang R, Mustafa MY, Zhao J, Bozkurt A, Xu L, Wang H, Salha S, Altinay F, Affouneh S, Burgos D. Speaking of transparency: Are all Artificial Intelligence (AI) literature reviews in education transparent?. Journal of Applied Learning and Teaching. 2023;6(2).	More of Artificial Intelligence in Education than Artificial Intelligence supported Literature reviews	PRISMA criteria and PRISMA techniques to increase tranparency	The use of AI in education raises ethical challenges related to transparency
Carabantes, D., González-Geraldo, J.L. and Jover, G., 2023. ChatGPT could be the reviewer of your next scientific paper. Evidence on the limits of AI-assisted academic reviews. *Profesional de la información/Information Professional*, *32*(5).	ChatGPT	The use of ChatGPT in academic reviews with ChatGPT as a tool to support literature review processes for undergraduate students Idea generation, summarizing literature, and manuscript preparation	Ethical concerns such as plagiarism and misinformation Limitations of ChatGPT in enabling a reliable and transparent generation of content, as well as its incompatibility with current ethical and legal obligations of academic publishing
de la Torre-López, J., Ramírez, A. and Romero, J.R., 2023. Artificial intelligence to automate the systematic review of scientific literature. *Computing*, pp.1-24.	Respondents commonly used tools such as **Covidence** and **RevMan**, with reported time savings and increased accuracy, but also encountered barriers such as lack of knowledge and mismatch to workflow	Searching, screening, extraction and synthesis Peer-review process Emphasising ethical concerns	*Ethical concerns, particularly associated with bias and the potential unintended consequences of algorithmic bias * The lack of policies on the use of AI tools in the peer review process * The need for transparent procedures to protect the integrity of the peer review
Ismail, F., Tan, E., Rudolph, J., Crawford, J. and Tan, S., 2023. Artificial intelligence in	ChatGPT, GPT-4, Bard, Bing Chat, Claude, or Ernie, DALL-E, GitHub Copilot,	A method for creating a systematic literature review to facilitate an analysis of	There is a need to investigate the potential benefits and challenges of Generative AI tools to

higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning and Teaching*, *6*(2).	GPT-4 plugins, Midjourney, Runway, and Synthesia	studies on generative AI and higher education A systematic approach to craft out and analyse the metadata of articles published on specific types of generative AI	ensure its ethical, effective, and responsible use
Guo, E., Gupta, M., Deng, J., Park, Y.J., Paget, M. and Naugler, C., 2024. Automated Paper Screening for Clinical Reviews Using Large Language Models: Data Analysis Study. *Journal of Medical Internet Research*, *26*(1), p.e48996.	**ChatGPT** and **GPT-4 APIs**	Plain language prompts to the publicly available OpenAI GPT application programming interface (API) Screening titles and abstracts in clinical reviews	

The results from the thematic analysis are presented and discussed theme by theme here below, with a focus on findings that add value to the answering of the research questions.

3.1 Generative AI Tools

The first initial finding of the built in Scopus AI tool clearly facilitated the inception phase of this literature review. To get the earlier huge and unfocused result set filtered with a coherent focus on publications bringing up AI generated literature reviews was helpful and saved hours of hard work. Regarding the AI supported abstract skimming this could have been done manually, and the AI generated summaries had to be checked since they are not always correct. The summaries worked relatively well for longer and structured abstracts, but for short and abstract abstracts the results were poor or none. An example is the 112 words abstract in the article Marshall and Wallace (2019), where the abstract summary consisted of "Based on the provided abstract, there is no direct information about the specific tools, techniques, and ethics discussed in the paper". In a human screening a scarce abstract can be followed up by skimming the results and conclusion sections, but maybe that abstract screening is a doubtful technique for both humans and GenAI tools in literature reviews.

3.1.1 Before and after ChatGPT

The oldest publications that were published before the release of ChatGPT, bring up more of general deep learning techniques (Marshall and Wallace, 2019; van Dinter, Tekinerdogan and Catal, 2021). In the study by Marshall and Wallace (2019), there are discussions on two prototype platforms for Identifying reports of randomized controlled trials. The first one is ExaCT (Kiritchenko et al., 2010) and the second one, developed by the authors is RobotReviewer (Marshall et al., 2017). Other tools brought up in this publication were Thalia [for literature exploration], Abstracker, EPPI reviewer and RobotAnalyst [for screening], and NaCTeM [for text mining] (Przybyła et al., 2018; Marshall and Wallace, 2019). The study by van Dinter, Tekinerdogan and Catal (2021), discusses machine learning and text mining on a more abstract level, referencing to more specific tool descriptions in the publications of O'Mara-Eves (2015), Jonnalagadda, Goyal and Huffman (2015), Thomas, McNaught and Ananiadou (2011), Feng, Chiam and Lo (2017), and Beller et al. (2018).

Moreover, the pre-ChatGPT publications bring up tools such as MySLR, Porifera and RobotReviewer (da Silva Júnior and Dutra, 2021), and ASReview, Litbaskets, RevMan, RobotReviewer and WebPlotDigitizer (Wagner, G., Lukyanenko and Paré, 2022). Here the idea is to use different tools for the different steps i a literature review process, and to use the tools in the steps where they have the greatest potential to be useful. Some tools have a strong specialisation, with WebPlotDigitizer as an example for extracting data from statistical plots. In the study by da Silva Júnior and Dutra (2021), a literature review is divided into the steps of searching, screening, extraction, and synthesis. Compared to the division presented by Wagner, G., Lukyanenko and Paré (2022), with the six steps of 1) Problem formulation, 2) Literature search, 3) Screening for inclusion, 4) Quality assessment, 5) Data extraction, and 6) Data analysis and interpretation.

Later, in the publication from late 2022 and onwards ChatGPT is the most frequently occurring tool (Müller et al., 2022; Burger et al., 2023; Carabantes, González-Geraldo and Jover, 2023 Guo et al, 2024). Now with the idea of using the same GenAI tool for all the steps of searching, screening, extraction, and synthesis in a literature review (Müller et al., 2022). There are also ideas of complementing ChatGPT with programming and scripts, where the Python programming language with specialised extensions such as Pandas, PyTorch, or TensorFlow seems like an interesting option (Burger et al., 2023). GenAI can be applied to more modalities than text only,

and today, in 2024 the array of Gen AI tools involves a lot more than ChatGPT. There are also several versions of ChatGPT and an application programming interface and plug-ins for ChatGPT (Guo et al., 2024). Some examples of GenAI tools for other modalities than text are DALL-E and Runway [image generation], and Synthesia and Runway [video generation]. However, for literature reviews the main modality is text and how to extract and analyse text, where the current alternative tools to ChatGPT are Bard, Bing Chat, Claude and Copilot. (Ismail et al., 2023)

3.2 Supportive AI Techniques

One theme in the collected material is that AI techniques shows promise in supporting humans in narrowed and specialised tasks of literature reviews, such as abstract screening and paper selection. Prior research notes that the screening and selection phase of a literature review is one of the more explored and prominent of AI assisted literature reviews (Marshall and Wallace, 2019; da Silva Júnior and Dutra, 2021; de la Torre-López, Ramírez, and Romero, 2023). At the same time, it is stressed that humans should be kept in the loop even at these stages since machine learning algorithms mainly estimate the probability of inclusion/exclusion and therefore rather expedite tasks than automate them (Marshall and Wallace, 2019; Müller, Pachnanda, Pahl, and Rosenqvist, 2022). A related challenge is that the used AI techniques rank papers according to their probability of being relevant for inclusion, low ranked papers can still be relevant and therefore potentially be overlocked, then how do we make sure that the reviewer does not stop the manual screening and leave the rest to the AI too early (Marshall and Wallace, 2019)? Another potential problem that has been highlighted in previous research is that some AI techniques for screening and selecting may be more proficient in excluding irrelevant papers than what they are in identifying potentially relevant papers (Guo, Gupta, Deng, Park, Paget, and Naugler, 2024).

3.2.1 Reducing the workload

Another theme that emerged from the analysis is that supportive AI techniques could reduce the workload in literature reviews. Included papers note that the human workload can be reduced in literature reviews, especially in the operational stages (da Silva Júnior and Dutra, 2021), and that many tasks that are perceived as repetitive and costly by humans could be automated (de la Torre-López, Ramírez, and Romero, 2023). Machine learning (ML) techniques are highlighted as especially interesting for supporting paper screening, which is perceived by many as a burdensome task (de la Torre-López, Ramírez, and Romero, 2023). Prior research further stresses the importance of designing AI systems for literature reviews with both human and machine in mind, since full automation, end-to-end automation, will only be applicable in a subset of the literature reviews that are needed (Müller, Pachnanda, Pahl, and Rosenqvist, 2022).

3.3 Ethical Issues

A theme relating to ethical issues is the needs that the AI systems serve. A problem that has been noted in the collected material is that AI systems or prototypes that have been developed by academic groups to serve a purpose often does not age well, with errors, slow interfaces, and broken links (Marshall and Wallace, 2019). It is often difficult to obtain academic grants for maintaining these types of systems and commercial companies are not likely to spend resources on adapting a technique or method specialised for research if there is not a substantial user group or economic interest (Marshall and Wallace, 2019). Papers further highlight that different types of literature reviews often have different needs and therefore require different designs of the AI system. Applying AI techniques for literature reviews have for example been studied in computing and medicine (de la Torre-López, Ramírez, and Romero, 2023), although more research is needed to expand and explore application in these fields also (Guo, Gupta, Deng, Park, Paget, and Naugler, 2024). It could however be problematic to apply the same techniques in other fields, since terminology and types of research differ between every field of research (de la Torre-López, Ramírez, and Romero, 2023). Further, the design of AI tools should take into consideration the objective of literature reviews (Müller, Pachnanda, Pahl, and Rosenqvist, 2022). For example, is the literature review qualitative or quantitative, interpretivist or a positivist, narrow or broad, and so on (Müller, Pachnanda, Pahl, and Rosenqvist, 2022).

3.3.1 Assessment bias

A recurrent theme in the collected material is the risk of biases in the AI algorithms. Papers points to, for example, the risk of assessment biases in the selection, extraction and classification phases of conducting literature reviews (Marshall and Wallace, 2019; de la Torre-López, Ramírez, and Romero, 2023). It is further highlighted that classification is strongly reliant on the content or vocabulary of the paper and the field that it is a part of, which could result in excluding papers that apply a different or new terminology, but still relevant (de la Torre-López, Ramírez, and Romero, 2023). Prior research therefore recommends that the input from AI

systems in literature reviews are viewed as suggestions and not the final assessment (Marshall and Wallace, 2019). However, it should also be noted that AI could be used to lower the risk of human biases and errors, for example through supporting a selection of papers that could be viewed as more consistent and objective (Guo, Gupta, Deng, Park, Paget, and Naugler, 2024).

4. Concluding Discussion

AI supported built-in filtering in research databases seems like a promising feature, but the filtering might be biased and leave out some publication. This is a new technique and as pointed out at the Scopus web portal "Scopus AI is a newly-developed feature. The quality of results may vary." Some days during this study the feature worked excellent, but others there were repeated gateway time-outs caused by the Scopus AI tool. Moreover, the whole web portal collapsed sometimes after searches with the built-in AI tool.

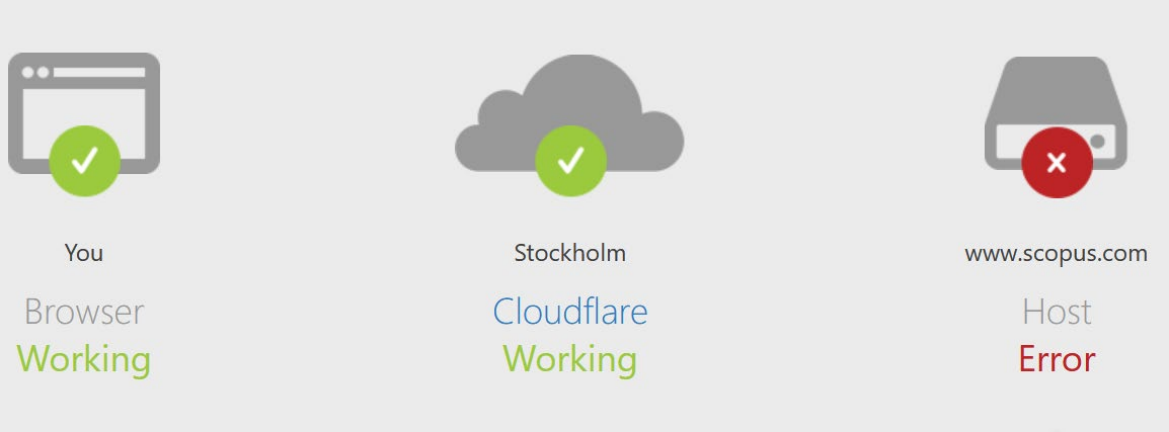

Figure 3: Scopus AI timeouts resulting in HTTP errors.

GenAI is a field that moves forward rapidly and in this literature review some of the found tools might be classified as 'yesterday's tools'. However, the presented findings could be seen as a summary of the rapid development in the field and that the fundamental techniques such as multi-layered neural networks and large language models remains. Finally, several ethical issues were found, and that they need further investigation. In this study, authors decided to use GenAI for the initial search and for abstract screening, but that data extraction and the thematic analysis was carried out manually. With the traditional approach to the most important phases of the literature study, accountability is a non-issue. Finally, the implications of the study are that many of the rapidly developed GenAI tool are far from mature, but literature reviews will definitely be AI supported in the near future.

5. Limitations and Future Research

Generative AI is a moving target and since some of the oldest publications in the result set are publications are from before the release of ChatGPT, the listing of found tools is missing some of the recently released GenAI tools. As an example, the literature review conducted by van Dinter, Tekinerdogan and Catal (2021), discusses some tools and techniques from publications older than 2020. Much has happened in the rapid development of GenAI, and some tools that would be interesting to evaluate in a future study are ChatPDF (Panda, 2023), Elicit (Whitfield and Hofmann, 2023) and ResearchRabbit (Cole and Boutet, 2023). Finally, ethics and accountability need further research.

References

Beller, E., Clark, J., Tsafnat, G., Adams, C., Diehl, H., Lund, H., Ouzzani, M., Thayer, K., Thomas, J., Turner, T. and Xia, J., 2018. Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). *Systematic reviews*, *7*, pp.1-7.

Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), pp.77-101.

Burger, B., Kanbach, D.K., Kraus, S., Breier, M. and Corvello, V., 2023. On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management*, *26*(7), pp.233-241.

Buriak, J.M., Akinwande, D., Artzi, N., Brinker, C.J., Burrows, C., Chan, W.C., Chen, C., Chen, X., Chhowalla, M., Chi, L. and Chueh, W., 2023. Best Practices for Using AI When Writing Scientific Manuscripts: Caution, Care, and Consideration: Creative Science Depends on It. *ACS nano*, *17*(5), pp.4091-4093.

Carabantes, D., González-Geraldo, J.L. and Jover, G., 2023. ChatGPT could be the reviewer of your next scientific paper. Evidence on the limits of AI-assisted academic reviews. *Profesional de la información/Information Professional*, *32*(5).

Cole, V. and Boutet, M., 2023. ResearchRabbit. *The Journal of the Canadian Health Libraries Association, 44*(2), p.43.

da Silva Júnior, E.M. and Dutra, M.L., 2021. A roadmap toward the automatic composition of systematic literature reviews. *Iberoamerican Journal of Science Measurement and Communication*.

de la Torre-López, J., Ramírez, A. and Romero, J.R., 2023. Artificial intelligence to automate the systematic review of scientific literature. *Computing*, pp.1-24.

Feng, L., Chiam, Y.K. and Lo, S.K., 2017. Text-mining techniques and tools for systematic literature reviews: A systematic literature review. In *2017 24th asia-pacific software engineering conference (apsec)* (pp. 41-50). IEEE.

Flick, C. and Worrall, K., 2022. The ethics of creative AI. In *The Language of Creative AI: Practices, Aesthetics and Structures* (pp. 73-91). Cham: Springer International Publishing.

Guo, E., Gupta, M., Deng, J., Park, Y.J., Paget, M. and Naugler, C., 2024. Automated Paper Screening for Clinical Reviews Using Large Language Models: Data Analysis Study. *Journal of Medical Internet Research*, *26*(1), p.e48996.

Ismail, F., Tan, E., Rudolph, J., Crawford, J. and Tan, S., 2023. Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning and Teaching*, *6*(2).

Jonnalagadda, S.R., Goyal, P. and Huffman, M.D., 2015. Automating data extraction in systematic reviews: a systematic review. *Systematic reviews*, *4*(1), pp.1-16

Kiritchenko, S., De Bruijn, B., Carini, S., Martin, J. and Sim, I., 2010. ExaCT: automatic extraction of clinical trial characteristics from journal publications. *BMC medical informatics and decision making*, *10*, pp.1-17.

Lim, W.M., Gunasekara, A., Pallant, J.L., Pallant, J.I. and Pechenkina, E., 2023. Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), p.100790.

Marshall, C., 2016. *Tool support for systematic reviews in software engineering* (Doctoral dissertation, Keele University).

Marshall, I.J., Kuiper, J., Banner, E. and Wallace, B.C., 2017. Automating biomedical evidence synthesis: RobotReviewer. In *Proceedings of the conference. Association for Computational Linguistics. Meeting* (Vol. 2017, p. 7). NIH Public Access.

Marshall, I.J. and Wallace, B.C., 2019. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Systematic reviews*, *8*, pp.1-10.

Mijwil, M.M., Hiran, K.K., Doshi, R., Dadhich, M., Al-Mistarehi, A.H. and Bala, I., 2023. ChatGPT and the future of academic integrity in the artificial intelligence era: a new frontier. *Al-Salam Journal for Engineering and Technology*, *2*(2), pp.116-127.

Müller, H., Pachnanda, S., Pahl, F. and Rosenqvist, C., 2022. The application of artificial intelligence on different types of literature reviews-A comparative study. In *2022 International Conference on Applied Artificial Intelligence (ICAPAI)* (pp. 1-7). IEEE.

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M. and Ananiadou, S., 2015. Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic reviews*, *4*, pp.1-22.

Panda, S., 2023. Enhancing PDF interaction for a more engaging user experience in library: Introducing ChatPDF. *IP Indian Journal of Library Science and Information Technology*, *8*(1), pp.20-25.

Pant, A., Hoda, R., Spiegler, S.V., Tantithamthavorn, C. and Turhan, B., 2023. Ethics in the Age of AI: An Analysis of AI Practitioners' Awareness and Challenges. *ACM Transactions on Software Engineering and Methodology*.

Przybyła, P., Brockmeier, A.J., Kontonatsios, G., Le Pogam, M.A., McNaught, J., von Elm, E., Nolan, K. and Ananiadou, S., 2018. Prioritising references for systematic reviews with RobotAnalyst: a user study. *Research synthesis methods*, *9*(3), pp.470-488.

Tlili A, Huang R, Mustafa MY, Zhao J, Bozkurt A, Xu L, Wang H, Salha S, Altinay F, Affouneh S, Burgos D. Speaking of transparency: Are all Artificial Intelligence (AI) literature reviews in education transparent?. Journal of Applied Learning and Teaching. 2023;6(2).

Thomas, J., McNaught, J. and Ananiadou, S., 2011. Applications of text mining within systematic reviews. *Research synthesis methods*, *2*(1), pp.1-14.

Wagner, G., Lukyanenko, R. and Paré, G., 2022. Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, *37*(2), pp.209-226.

van Dinter, R., Tekinerdogan, B. and Catal, C., 2021. Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology*, *136*, p.106589.