

# Machine Learning Methods for Systematic Reviews:

## A Rapid Scoping Review

Stephanie Roth, MLIS<sup>1</sup> and Alex Wermer-Colan, PhD<sup>2</sup>

1. Medical Librarian, Lewis B. Flinn Medical Library, ChristianaCare

2. Academic Director, Loretta C. Duckworth Scholars Studio, Temple University Libraries

### Abstract

**Objective.** At the forefront of machine learning research since its inception has been natural language processing, also known as text mining, referring to a wide range of statistical processes for analyzing textual data and retrieving information. In medical fields, text mining has made valuable contributions in unexpected ways, not least by synthesizing data from disparate biomedical studies. This rapid scoping review examines how machine learning methods for text mining can be implemented at the intersection of these disparate fields to improve the workflow and process of conducting systematic reviews in medical research and related academic disciplines. **Methods.** The primary research question that this investigation asked, “what impact does the use of machine learning have on the methods used by systematic review teams to carry out the systematic review process, such as the precision of search strategies, unbiased article selection or data abstraction and/or analysis for systematic reviews and other comprehensive review types of similar methodology?” A literature search was conducted by a medical librarian utilizing multiple databases, a grey literature search and handsearching of the literature. The search was completed on December 4, 2020. Handsearching was done on an ongoing basis with an end date of April 14, 2023. **Results.** The search yielded 23,190 studies after duplicates were removed. As a result, 117 studies (1.70%) met eligibility criteria for inclusion in this rapid scoping review. **Conclusions.** There are several techniques and/or types of machine learning methods in development or that have already been fully developed to assist with the systematic review stages. Combined with human intelligence, these machine learning methods and tools provide promise for making the systematic review process more efficient, saving valuable time for systematic review authors, and increasing the speed in which evidence can be created and placed in the hands of decision makers and the public.

### Methods

Machine learning refers to a wide range of computational methods involving the optimization of statistical and analytical processes towards enhanced pattern recognition and classification of common features across diverse datasets. At the forefront of machine learning has been experimentation and research involving data mining, especially text mining. Machine learning methods have shown promising applications in the field of information retrieval for identifying keywords, topics, and stylistic patterns across a body of texts.

In the last decade, sophisticated methods of machine learning have become increasingly possible at scale thanks to innovations in graphical processing units and related computing hardware.

The significance of these methodological evolutions for the field of information science and several other fields remains under-examined today. While systematic reviews are a growing practice in library and information science, evidence on the usefulness of machine learning

methods for the improvement of automated searches and filtration of resources and other review methods is lagging recent advancements in the field. Furthermore, assessment of tools and their applications in real-world scenarios remains minimal, with library practitioners needing more guidance on what resources can enable more efficient searching and reviews.

This rapid scoping review, then, seeks to address a potential blind spot in the current review literature for the wide variety of machine learning approaches and methods that have been applied to the systematic review process. To this end, this paper sets out to examine the impact of machine learning and related language processing algorithms, with a focus on what impact these techniques can have on improving the efficiency of human workflows.

Completing a systematic review is no small endeavor, typically taking months of planning and over a year to complete. This was felt firsthand by the review team in undertaking this rapid scoping review. While the timeframe does not appear to be rapid, it was necessary for the review to become rapid to avoid spanning several more years due to the broad nature of the research question. To this end, this review investigates the effectiveness and impact of machine learning for each main stage of the systematic review process. The intended audience is librarians/information professionals who often are the ones guiding researchers through these stages; and to those who develop or are interested in developing future tools/software for systematic reviews.

This rapid scoping review is categorized by the general stages of the systematic review:

### **Protocol/Planning Stage**

This stage involves investigating the feasibility of the review, gathering a team, and formulating/preregistering a protocol outlining the detailed methods the review will follow, including the predetermined inclusion/exclusion criteria for study selection. Working with a librarian is important at the onset of the review and during the protocol development stage.

### **Search Stage**

This stage of the review involves working with a co-authored librarian or information specialist who will create, test, and provide the team with a comprehensive search strategy and translate this search strategy across several databases and grey literature sources. Additional methods may be employed such as citation chaining. Review team members with the most subject matter expertise will hand search the literature.

### **Screening Stage**

#### ***Title/Abstract Screening***

This stage of the review involves screening the title/abstracts found in the search results for relevancy. This process is conducted by two independent and blinded reviewers. Studies are filtered by Yes/No, no reasons are recorded for exclusion.

#### ***Full Text Screening***

This stage of the review involves screening only the full texts of the included studies (Yes responses) from the title/abstract phase. Full texts of these studies are gathered, read, and checked against the predefined inclusion/exclusion criteria. This process is also conducted by

two independent and blinded reviewers. Studies are filtered by Yes/No, reasons for exclusion are provided at this phase.

### **Data Extraction Stage**

This stage of the review involves extracting the data from the included studies from the results of the full text review stage, characteristics of studies are recorded and other important data that will inform the review and statistical analysis if a meta-analysis is required.

### **Appraisal/Synthesis and Analysis Stage**

#### ***Critical Appraisal***

This stage of the review involves analyzing the included studies for risk of bias and/or assessing the quality of the study designs and methods.

#### ***Synthesis/Writing***

This stage of the review involves synthesizing the evidence from the included studies. Conclusions are drawn from the evidence and gaps are explored for further research. The written portion of the review can be completed without including a meta-analysis or other statistical analysis.

#### ***Meta-analysis/Analysis***

This stage of the review is only for those that require a meta-analysis or statistical analysis. Not every review requires one. It is recommended to work with a biostatistician at the onset of the review to determine if one is necessary.

Since a rapid scoping review, PRISMA-ScR was utilized for the reporting of this review. A protocol was registered in the Open Science Framework as a preregistration (<https://osf.io/j8ydg/>).

### **Selection Criteria**

The following criteria had to be met for a study to meet inclusion criteria for this rapid scoping review: research methods studies, a focus on machine learning, text analysis and/or automation, a focus on all or one or more stage(s) of the systematic review process, and the use of a machine learning application to assist with any or all stages of the review process. This review takes into consideration the overall landscape of machine learning and text analysis in systematic reviews, especially in terms of emerging trends and methods, while also being attentive to the barriers to facilitation and widespread adoption when user-friendly tools are not readily available.

Systematic reviews including all evidence syntheses were excluded since this review focuses on papers about the methods used in systematic reviews. Studies about updating a systematic review were excluded since this review to examines machine learning in the context of doing a new systematic review. Editorials, book chapters and similar works were also excluded.

### **Search Methods**

To identify studies to include or consider for this rapid scoping review a medical librarian (SR), developed detailed search strategies for each database. The search was developed for PubMed

(NLM) and was translated to Embase (Elsevier), Scopus (Elsevier), LISTA (EbscoHost) and the Social Science Premium Collection (ProQuest).

An attempt to locate grey literature was carried out using a Google search and scanning the SuRe Info web resource (<https://sites.google.com/york.ac.uk/sureinfo/home>).

A handsearch was conducted by scanning reference lists and the following journals or conference proceedings of *BMC Systematic Reviews*, *Journal of Clinical Epidemiology*, *International Conference on Evaluation and Assessment in Software Engineering*, *Journal of Biomedical Informatics*, *JAMIA*, *AMIA*, *Research Synthesis Methods*, *BMC Bioinformatics*, *Expert Systems with Applications*, *ESMAR Conf* and *MLA vConference*.

The search includes a date restriction of 2003 to present. The date restriction is justified due to the slow growth of machine learning overall and the use of machine learning in systematic reviews is a present-day advancement. Although this is not an exact cut-off date, the authors did not see a reason to search further back in the literature, as there were no potential harms or risks to persons due to the nature of this rapid scoping reviews focus on research methods.

The full systematic review search was completed on December 4, 2020, and formal handsearching was done on an ongoing basis with an end date of April 14, 2023.

While this review began before PRISMA-S for searching was published, the search methods are reported with the exclusion of a formal peer review of the search strategies.

Details of the search are provided in the Supplementary Materials and are available in the Temple University institutional repository, TUScholarShare.

Original Search Results by Database:

- PubMed (NLM) (10,695 Results)
- Embase (Elsevier) (981 Results)
- Scopus (Elsevier) (5,831 Results)
- LISTA (EbscoHost) (2,921 Results)
- Social Science Premium Collection (ProQuest) (2,589 Results)

The search resulted in 23,190 studies (including 147 from grey literature sources and 38 from hand searching). The 4,440 duplicate studies were found and omitted using Endnote X.7 for the deduplication of records and 18,750 references were eligible to screen.

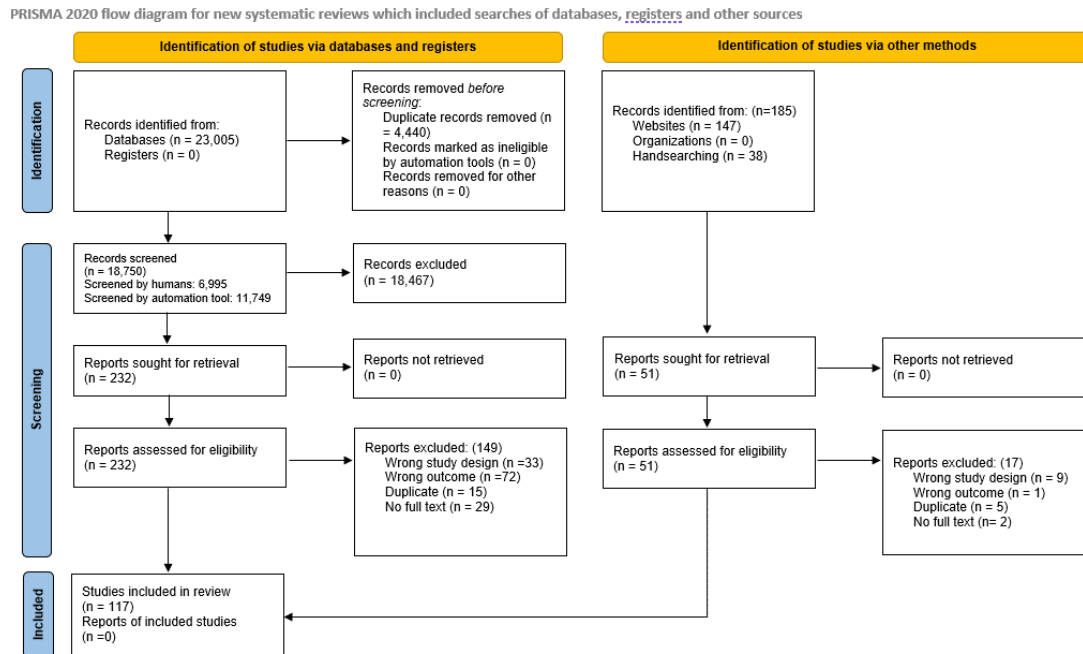
## Study Selection

Titles and abstracts were screened independently and blinded by two reviewers (SR, AWC) to identify studies that potentially met inclusion criteria. Endnote X.7 was used to manage references and remove duplicates before importing into Abstrackr to ensure blinded screening among reviewers. Abstrackr was selected as a screening tool because it uses machine learning to rank studies by relevance. The two blinded reviewers only screened the first 6,995 titles and abstracts. This threshold was decided upon since there were no longer relevant articles. Any disagreements between reviewers of those that were screened, were resolved by a third reviewer (JP) who served as a tiebreaker.

The full text of 283 potentially eligible studies were then reviewed for eligibility by two independent reviewers (SR, AWC). There were no disagreements identified between reviewers.

DeepL was used to translate one non-English language study. It was determined that 166 studies were excluded for the following reasons: wrong study type, wrong outcome, duplicate study, and the lack of availability of a full text. It was determined that 117 studies (1.70%) met eligibility criteria and were included in the final analysis. Details of the study selection process are detailed in the PRISMA Flow Diagram (Figure 1).

Figure 1: PRISMA Flow Diagram



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: <http://www.prisma-statement.org/>

## Data Extraction

Data was extracted from each of the included studies using a custom data extraction form. Two review authors extracted data using Excel spreadsheets, there were no discrepancies between reviewers. Characteristics of included studies are available in the online supplementary materials

## Results

This rapid scoping review categorizes the results based on the stage of the systematic review and in order of the review process. A category was also created for multiple review stages; this category includes papers that included multiple tools/software and/or methods for more than one stage of the review or a single tool/software and/or method that could be used for multiple stages of the review process. Results by review stage:

### Protocol/Planning Stage

No studies were identified for this stage of the review.

## Search Stage

Utilization of machine learning tools for systematic review searching can have its own set of time savings advantages when employed correctly with the involvement of a librarian or information expert. Several studies evaluated new or existing machine learning tools to support the search stage of the systematic review (studies cited in the online supplementary materials). These tools include RCT tagger, Litsearchr, WordNet, Robot Search, Zettair, SLR.qub tool (Systematic Literature Review – query builder), TerMine, Leximancer and Paperfetcher. A tool included in this category is the automated deduplication tool, Deduklick.

Several studies addressed user-friendliness, an important aspect for adoption of machine learning for systematic reviews. While Grames reported on the user-friendliness of the R package, Litsearchr, another caveat to implementation identified was the need for basic experience in using the R coding language. One barrier to the adoption of machine learning in systematic reviews is not having basic knowledge of how-to code. Most researchers conducting systematic reviews will not have this specialized skill set. While not expected, having a basic to advanced understanding in how to code can offer several benefits. One benefit of utilizing code is that there is no cost, as once free systematic review tools move to become proprietary (e.g., Covidence). The availability and cost of tools included in this rapid scoping review are reported in the online supplementary materials.

## Screening Stage

Screening titles/abstracts is one of the most laborious and time-consuming stages of the systematic review process and it is here when systematic review teams are truly tested to see if they have both the capacity and the time to complete a full systematic review. By design, machine learning is a suitable method for ameliorating this problem. While there is plenty of literature in this area, many are computer science-based studies describing complex algorithms or methods that will likely not result in the development of a public-facing tool in the short term. Hamel et al., provides some guidance and a set of recommendations for implementing machine learning in the title/abstract screening stage of the systematic review.<sup>1</sup>

Less than 20 studies in this category referenced an existing tool available for public use. Several examples of publicly available tools include: DistillerSR,<sup>2</sup> Rayyan, Colandr, Abstrackr, EPPI-Reviewer, ASReview, RobotAnalyst, SWIFT Review, MetaMap, RapidMiner, and SyRF. In some cases, a tool in development was mentioned that either could not be found or no longer exists and/or never made it to production for public use such as StArt,<sup>3,4</sup> Revis<sup>5</sup> and TWISTER.<sup>6</sup> It is not certain if they were later merged with another tool, if they are still in the development stage, or if they never were intended for public use.

A main barrier to adoption of machine learning in systematic reviews is user-friendliness. Only 11 studies<sup>7-17</sup> addressed user-friendliness when evaluating their tool or process.

A few projects in existence to assist with the screening stage of the review are not yet suitable for the public or for those without advanced coding skills.<sup>18-20</sup>

The use of machine learning has shown to reduce the time and human effort devoted to the screening stage of a systematic review;<sup>7,21-40</sup> however, the availability of their data was absent, making it hard to replicate. Extensive research is being done in this area, but data and/or computer scientists need to work together and across disciplines to learn from prior work. Data

transparency allows others to build upon their work. However, there seems to be more proprietary tools in this category, making it difficult.

### **Data Extraction**

Machine learning can have time saving advantages for the data extraction stage of the systematic review that involves manually reading and extracting relevant texts from the included and chart the data. Duc An Bui, et al<sup>41</sup> utilized a PDF text classification tool to see how it helped with the data extraction stage of the systematic review, focusing on PDFBox (<https://PDFbox/pdfbox.apache.org/>) in combination with an annotation tool called GATE. Two studies<sup>42,43</sup> referenced the tool, EXACT (<https://bio-nlp.org/EXACT/>), a tool designed to assist in the data extraction of clinical trials from a trials registry, clinicaltrials.gov. Both studies report that the data extracted was accurate and had a significant reduction on workload.

Torres & Cruzes introduced a tool called, Textum which used machine learning to assist researchers in analyzing specific parts of a paper. It estimated to have an overall 80% reduction in the time spent analyzing the texts of a traditional review.<sup>44</sup> However, today, this tool could not be located online, and it is not clear if it ever existed for public use. Other studies noted machine learning methods might help with the data extraction stage of the systematic review, but they have not yet resulted in a tool for the public.<sup>45,46</sup> However, only one has shared data from their study, which is a barrier to further developing machine learning for this stage of the review.<sup>45</sup> One limitation to adoption of these tools is that only one study in this category mentioned being user-friendly. The AFLEX-tag tool was reported to have a user-centered design.<sup>47</sup> Unfortunately, this tool does not seem to be publicly available.

### **Appraisal/Synthesis and Analysis Stage**

The appraisal, synthesis, and analysis stages of the systematic review are areas where machine learning can aid, but likely not be a replacement for, human input. Several tools in this category were based on coding algorithms,<sup>48</sup> with the exception of RobotReviewer (<https://www.robotreviewer.net/>), a tool publicly available that could assist with reducing the time spent on the risk of bias assessment stage, but not serve as a complete replacement for manual risk of bias assessment.<sup>48-50</sup>

A couple of R packages were also designed to help with the final analysis stage of the systematic review, such as Robumeta<sup>51</sup> and PublicationBias<sup>52</sup> which can assist with the sensitivity analysis for publication bias in systematic reviews.

Lingo3G is a machine learning tool that could support scoping reviews utilizing clustering to generate themes and/or a set of codes across several studies more rapidly than manual methods for synthesizing and coding the literature.<sup>53</sup> Marshall et al. (2015)<sup>48</sup> described a novel method for using support vector machines (SVM) to help automate the risk of bias assessment of clinical trials in systematic reviews. The author's goal was to pair it with another tool in the pipeline to semi-automate the screening of abstracts.

Millard et al. reported a tool that was found to reduce the amount of time required by human reviewers for the risk of bias assessment.<sup>54</sup> On average it was found that more than 33% of research articles could be labeled with higher certainty than that of a human reviewer. While the use of machine learning to assist with the risk of bias assessment was like the method introduced

by Marshall et al.,<sup>48,49</sup> one difference according to Millard et al.<sup>54</sup> is that their team tested its method using full text articles rather than only titles and abstracts.

### Multiple Review Stages

In a case study by Clark et al., multiple tools were used to complete a full small-scale systematic review in just two weeks (2weekSR) for multiple stages of the review.<sup>55</sup> One of the tools they used for the 2weekSR, was the RobotReviewer (<https://www.robotreviewer.net/>), a semi-automation tool that uses machine learning to help with the risk of bias assessment of randomized controlled trials. A suite of automation tools used for the completion of a 2weekSR included the Systematic Review Accelerator,<sup>55,56</sup> which is designed to speed up each stage of the systematic review process. Recently the 2weekSR was tested on large scale systematic reviews,<sup>57</sup> completed within a few weeks. While review team members for the 2weekSR had protected research time to work on these projects, the findings are still demonstrative of the time reduction benefit of using machine learning.

Similarly, Haddaway et al. evaluates the use of partial automation using computational methods to assist in the facilitation of conducting a mapping review.<sup>58</sup> Since mapping reviews (and similarly scoping reviews) often assess a greater volume of literature than a traditional systematic review, partial automation can offer several workload advantages for the review team for various stages of the review. Lagopoulos & Tsoumakas, explored similar advantages with the hybrid machine learning tool, Elastic (<https://www.elastic.co/what-is/elasticsearch-machine-learning>). Elastic had assisted with the preparation, retrieval, and appraisal stages of the systematic review. Utilizing this technology was called a technology assisted review (TAR). This hybrid approach is one that doesn't involve creating a Boolean query by information experts. It relies instead upon initial machine learning retrieval methods, inter-review ranking and intra-review ranking.

When Altana, A. J. et al.<sup>59</sup> examined multiple machine learning tools for multiple systematic review stages, one noteworthy tool that stood out was Swift-Review (<https://www.sciome.com/swift-review/>). Swift-Review uses statistical text mining and machine learning methods to help with search refinement to mine for relevant terms and with literature prioritization to help rank order documents for manual screening. Interestingly, Altana reports a low uptake in utilizing tools like Swift-Review in systematic reviews despite the advantages they may provide. Barriers to adoptability were usability, licensing, the steep learning curve, lack of support, mismatch to the workflow and the lack of time needed to assess or evaluate a new tool.

It is important for researchers to evaluate how much time should be devoted to assessing, evaluating, and implementing new tools. Further complicating this effort is the fact that free tools often are not sustainable with little support for users, especially when they are created by an individual or as a side hobby. This can often make the effort for review teams to implement machine learning tools more time intensive which is a barrier to the time reduction they normally offer. Out of the studies in this category, only two studies by Fabbri et al.<sup>3,4</sup> address the user-friendliness of the machine learning tool for systematic reviews, StArt ([http://lapes.dc.ufscar.br/tools/start\\_tool](http://lapes.dc.ufscar.br/tools/start_tool)). It would be helpful if future studies adopt this approach.



Miranda et al., explores the development of a new tool with screenshots of how it is applied to assist with several stages of the systematic review (i.e., the search, article selection, and data extraction). However, this tool doesn't appear to be available to the general public.<sup>60</sup>

## **Discussion**

This rapid scoping review examines the use of machine learning and related language processing algorithms and its impact on improving the efficiency of human workflows that researchers have been developing to reduce the amount of time necessary to complete each stage of the systematic review when compared to manual methods. Completing a systematic review that is done adequately is a time-consuming task and takes a minimum of a year or more depending on several factors.

A few limitations of this review include not exploring machine learning research outside of systematic reviews, which may have led to omitting papers still potentially relevant to the systematic review process. While not a formal exclusion of the paper, the search did not explore all computer science and computational method sources, databases and/or journals.

This rapid scoping review provides important insights into the state of machine learning developments to reduce the time and human labor spent on conducting a new systematic review from start to finish. Included in this review were the barriers to adoption of machine learning for researchers and the lack or reproducibility of machine learning data for those developing new software or tools.

## **Conclusion**

Testing and providing more robust insights into the implementation of these machine learning methods and/or tools, such as demonstrating and rating the user-friendliness for the general user, are especially important, but are not widely demonstrated in the current literature. The results of this review may help computer scientists and/or programmers to eliminate research waste by identifying what methods or tools are already being developed across several disciplines. Perhaps new collaborations will result in building new tools that can universally address the current inefficiencies of conducting a manual systematic review. Librarians and information specialists will be able to find new ways to partner with researchers to supplement laborious tasks with machine learning.

While machine learning can assist with the systematic review process at various stages, it is still an emerging field for wide-scale application and must rely upon human input to be successfully implemented. Those who are developing tools or machine learning algorithms should work to make sure their research is clear and transparent to allow others to build upon their work. Future machine learning tools will need to be built with the end-user in mind for ease of use and widespread adoption. The barriers to adoption of machine learning should be considered and addressed during development.

Based on the current research, the use of machine learning to make the systematic review process more efficient remains favorable. Nevertheless, the researcher will need to be able to evaluate this new technology and select those that are best suited for their systematic review team's needs. Combined with human intelligence, machine learning looks promising for making the systematic review process more efficient, saving time for the review team, and increasing the speed in

which evidence is created. However, more experimental research studies with reproducible and open datasets are needed in the future to prove its effectiveness.

While this is an evolving area, currently machine learning is not a replacement for human effort. Those with the most systematic review methodological expertise, librarians, information specialists, statisticians, are all still essential in the overall design and implementation of the systematic review. Machine learning has the potential now to accelerate the rate of review completion for researchers and/or librarians and other information experts who invest time in learning and adopting these new tools.

## Acknowledgements

Many thanks to Ania Korsunska for her assistance in screening titles/abstracts and to Jenny Pierce, MS for her assistance in serving as a tiebreaker.

## Data Availability Statement

Search strategies and the deduplicated citations are deposited in TUScholarShare available at: <https://scholarshare.temple.edu/handle/20.500.12613/4637>

Appendices and supplementary materials such as the data extraction template, search strategies, and characteristics of included studies can be found at: <https://osf.io/x84t5/>

Ms. Roth may be contacted at [stephanie.roth@christianacare.org](mailto:stephanie.roth@christianacare.org).

## References

1. Hamel, C., Hersi, M., Kelly, S. E., Tricco, A. C., Straus, S., Wells, G., . . . Hutton, B. (2021, December 20). Guidance for using artificial intelligence for title and abstract screening while conducting knowledge syntheses. *BMC Medical Research Methodology*, 21(1), 285. <https://doi.org/10.1186/s12874-021-01451-2> PubMed
2. Gartlehner, G., Wagner, G., Lux, L., Affengruber, L., Dobrescu, A., Kaminski-Hartenthaler, A., & Viswanathan, M. (2019, November 15). Assessing the accuracy of machine-assisted abstract screening with DistillerAI: A user study. *Systematic Reviews*, 8(1), 277. <https://doi.org/10.1186/s13643-019-1221-3> PubMed
3. Fabbri, S., Hernandez, E., Di Thommazo, A., Belgamo, A., Zamboni, A., & Silva, C. (2013). Using information visualization and text mining to facilitate the conduction of systematic literature reviews. In J. Cordeiro, L. A. Maciaszek, & J. Filipe (Eds.), *Enterprise Information Systems* (Vol. 141, pp. 243–256). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-40654-6\\_15](https://doi.org/10.1007/978-3-642-40654-6_15)
4. Fabbri, S., Silva, C., Hernandez, E., Octaviano, F., Di Thommazo, A., & Belgamo, A. (2016). Improvements in the StArt tool to better support the systematic review process. *Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering*, 1–5. <https://doi.org/10.1145/2915970.2916013>
5. Felizardo, K. R., Salleh, N., Martins, R. M., Mendes, E., MacDonell, S. G., & Maldonado, J. C. (2011). Using visual text mining to support the study selection activity in systematic literature reviews. *2011 International Symposium on Empirical Software Engineering and Measurement*, 77–86. <https://doi.org/10.1109/ESEM.2011.16>

6. Kreiner, K., Hayn, D., & Schreier, G. (n.d.). Twister: a tool for reducing screening time in systematic literature reviews. 5.
7. Matwin, S., Kouznetsov, A., Inkpen, D., Frunza, O., & O'Blenis, P. (2010, July-August). A new algorithm for reducing the workload of experts in performing systematic reviews. *Journal of the American Medical Informatics Association : JAMIA*, 17(4), 446–453. <https://doi.org/10.1136/jamia.2010.004325> PubMed
8. Cleo, G., Scott, A. M., Islam, F., Julien, B., & Beller, E. (2019, June 20). Usability and acceptability of four systematic review automation software packages: A mixed method design. *Systematic Reviews*, 8(1), 145. <https://doi.org/10.1186/s13643-019-1069-6> PubMed
9. Gates, A., Guitard, S., Pillay, J., Elliott, S. A., Dyson, M. P., Newton, A. S., & Hartling, L. (2019, November 15). Performance and usability of machine learning for screening in systematic reviews: A comparative evaluation of three tools. *Systematic Reviews*, 8(1), 278. <https://doi.org/10.1186/s13643-019-1222-2> PubMed
10. Hamel, C., Kelly, S. E., Thavorn, K., Rice, D. B., Wells, G. A., & Hutton, B. (2020, October 15). An evaluation of DistillerSR's machine learning-based prioritization tool for title/abstract screening - impact on reviewer-relevant outcomes. *BMC Medical Research Methodology*, 20(1), 256. <https://doi.org/10.1186/s12874-020-01129-1> PubMed
11. Hindriks, S., & van de Schoot, R. (n.d.). A study on the user experience of the ASReview software tool for experienced and unexperienced users. 42.
12. Liao, J., Ananiadou, S., Currie, L. G., Howard, B. E., Rice, A., Sena, S. E., . . . Macleod, M. R. (2018). Automation of citation screening in pre-clinical systematic reviews [Preprint]. *Neuroscience*. <https://doi.org/10.1101/280131>
13. Olofsson, H., Brolund, A., Hellberg, C., Silverstein, R., Stenström, K., Österberg, M., & Dagerhamn, J. (2017, September). Can abstract screening workload be reduced using text mining? User experiences of the tool Rayyan. *Research Synthesis Methods*, 8(3), 275–280. <https://doi.org/10.1002/jrsm.1237> PubMed
14. Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016, December 5). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210. <https://doi.org/10.1186/s13643-016-0384-4> PubMed
15. Przybyła, P., Brockmeier, A. J., Kontonatsios, G., Le Pogam, M. A., McNaught, J., von Elm, E., . . . Ananiadou, S. (2018, September). Prioritising references for systematic reviews with RobotAnalyst: A user study. *Research Synthesis Methods*, 9(3), 470–488. <https://doi.org/10.1002/jrsm.1311> PubMed
16. Tan, M. C. (2018). Colandr. *Journal of the Canadian Health Libraries Association / Journal de l'Association Des Bibliothèques de La Santé Du Canada*, 39(2), 85–88. <https://doi.org/10.29173/jchla29369>
17. van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdem, F., . . . Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, 3(2), 125. <https://doi.org/10.1038/s42256-020-00287-7>

18. Westgate, M. J. (2019, December). revtools: An R package to support article screening for evidence synthesis. *Research Synthesis Methods*, 10(4), 606–614. <https://doi.org/10.1002/jrsm.1374> PubMed
19. Yu, Z., Kraft, N. A., & Menzies, T. (2018). Finding better active learners for faster literature reviews. *Empirical Software Engineering*, 23(6), 3161–3186. <https://doi.org/10.1007/s10664-017-9587-0>
20. Yu, Z., & Menzies, T. (2019). FAST2: An intelligent assistant for finding relevant papers. *Expert Systems with Applications*, 120, 57–71. <https://doi.org/10.1016/j.eswa.2018.11.021>
21. García Adeva, J. J., Pikatza Atxa, J. M., Ubeda Carrillo, M., & Ansuategi Zengotitabengoa, E. (2014). Automatic text classification to support systematic reviews in medicine. *Expert Systems with Applications*, 41(4), 1498–1508. <https://doi.org/10.1016/j.eswa.2013.08.047>
22. Bekhuis, T., & Demner-Fushman, D. (n.d.). Towards automating the initial screening phase of a systematic review. 5.
23. Bekhuis, T., Tseytlin, E., Mitchell, K. J., & Demner-Fushman, D. (2014, January 27). Feature engineering and a proposed decision-support system for systematic reviewers of medical evidence. *PLoS One*, 9(1), e86277. <https://doi.org/10.1371/journal.pone.0086277> PubMed
24. Frunza, O., Inkpen, D., Matwin, S., Klement, W., & O’Blenis, P. (2011, January). Exploiting the systematic review protocol for classification of medical abstracts. *Artificial Intelligence in Medicine*, 51(1), 17–25. <https://doi.org/10.1016/j.artmed.2010.10.005> PubMed
25. Hashimoto, K., Kontonatsios, G., Miwa, M., & Ananiadou, S. (2016, August). Topic detection using paragraph vectors to support active learning in systematic reviews. *Journal of Biomedical Informatics*, 62, 59–65. <https://doi.org/10.1016/j.jbi.2016.06.001> PubMed
26. Jonnalagadda, S., & Petitti, D. (2013). A new iterative method to reduce workload in systematic review process. *International Journal of Computational Biology and Drug Design*, 6(1/2), 5–17. <https://doi.org/10.1504/IJCBDD.2013.052198> PubMed
27. Kim, S., & Choi, J. (2012, March). Improving the performance of text categorization models used for the selection of high quality articles. *Healthcare Informatics Research*, 18(1), 18–28. <https://doi.org/10.4258/hir.2012.18.1.18> PubMed
28. Kim, S., & Choi, J. (2014, February). An SVM-based high-quality article classifier for systematic reviews. *Journal of Biomedical Informatics*, 47, 153–159. <https://doi.org/10.1016/j.jbi.2013.10.005> PubMed
29. Kouznetsov, A., Matwin, S., Inkpen, D., Razavi, A. H., Frunza, O., Sehatkar, M., . . . O’Blenis, P. (2009). Classifying biomedical abstracts using committees of classifiers and collective ranking techniques. In Y. Gao & N. Japkowicz (Eds.), *Advances in Artificial Intelligence* (Vol. 5549, pp. 224–228). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-01818-3\\_29](https://doi.org/10.1007/978-3-642-01818-3_29)
30. A., Nie, J.-Y., Thomas, J., Hong, Q. N., & Pluye, P. (2018). Discriminating between empirical studies and nonempirical works using automated text classification. *Research Synthesis Methods*, 9(4), 587–601. <https://doi.org/10.1002/jrsm.1317>

31. Li, D., Wang, Z., Wang, L., Sohn, S., Shen, F., Murad, M. H., & Liu, H. (2017). A text-mining framework for supporting systematic reviews. 17.
32. Li, D., Zafeiriadis, P., & Kanoulas, E. (2020). APS: An active PubMed search system for technology assisted reviews. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2137–2140. <https://doi.org/10.1145/3397271.3401401>
33. Liu, J., Timsina, P., & El-Gayar, O. (2018). A comparative analysis of semi-supervised learning: The case of article selection for medical systematic reviews. *Information Systems Frontiers*, 20(2), 195–207. <https://doi.org/10.1007/s10796-016-9724-0>
34. Miwa, M., Thomas, J., O'Mara-Eves, A., & Ananiadou, S. (2014, October). Reducing systematic review workload through certainty-based screening. *Journal of Biomedical Informatics*, 51, 242–253. <https://doi.org/10.1016/j.jbi.2014.06.005> PubMed
35. Norman, C. R., Leeflang, M. M. G., Porcher, R., & Névéal, A. (2019, October 28). Measuring the impact of screening automation on meta-analyses of diagnostic test accuracy. *Systematic Reviews*, 8(1), 243. <https://doi.org/10.1186/s13643-019-1162-x> PubMed
36. Rizzo, G., Tomassetti, F., Vetrò, A., Ardito, L., Torchiano, M., Morisio, M., & Troncy, R. (2015). Semantic enrichment for recommendation of primary studies in a systematic literature review. *Digital Scholarship in the Humanities*, fqv031. <https://doi.org/10.1093/llc/fqv031>
37. El-Gayar, O., & Liu, J. (n.d.). Active learning for the automation of medical systematic review creation. 12.
38. Tomassetti, F., Rizzo, G., Vetro, A., Ardito, L., Torchiano, M., & Morisio, M. (2011). Linked data approach for selection process automation in systematic reviews. *15th Annual Conference on Evaluation & Assessment in Software Engineering (EASE 2011)*, 31–35. <https://doi.org/10.1049/ic.2011.0004>
39. van Altena, A. J., & Olabarriaga, S. D. (n.d.). Predicting publication inclusion for diagnostic accuracy test reviews using random forests and topic modelling. 9.
40. Wallace, B. C., Noel-Storr, A., Marshall, I. J., Cohen, A. M., Smalheiser, N. R., & Thomas, J. (2017, November 1). Identifying reports of randomized controlled trials (RCTs) via a hybrid machine learning and crowdsourcing approach. *Journal of the American Medical Informatics Association : JAMIA*, 24(6), 1165–1168. <https://doi.org/10.1093/jamia/ocx053> PubMed
41. Bui, D. D. A., Del Fiol, G., & Jonnalagadda, S. (2016, June). PDF text classification to leverage information extraction from publication reports. *Journal of Biomedical Informatics*, 61, 141–148. <https://doi.org/10.1016/j.jbi.2016.03.026> PubMed
42. Kiritchenko, S., de Bruijn, B., Carini, S., Martin, J., & Sim, I. (2010, September 28). ExaCT: Automatic extraction of clinical trial characteristics from journal publications. *BMC Medical Informatics and Decision Making*, 10, 56. [PubMedhttps://doi.org/10.1186/1472-6947-10-56](https://doi.org/10.1186/1472-6947-10-56)
43. Pradhan, R., Hoaglin, D. C., Cornell, M., Liu, W., Wang, V., & Yu, H. (2019, January). Automatic extraction of quantitative data from ClinicalTrials.gov to conduct meta-analyses.

*Journal of Clinical Epidemiology*, 105, 92–100.  
<https://doi.org/10.1016/j.jclinepi.2018.08.023> PubMed

44. Torres, J. A. S., & Cruzes, D. S. (2013). Automatically locating results to support systematic reviews in software engineering. 15.
45. Cramond, F., O'Mara-Eves, A., Doran-Constant, L., Rice, A. S., Macleod, M., & Thomas, J. (2019, March 7). The development and evaluation of an online application to assist in the extraction of data from graphs for use in systematic reviews. *Wellcome Open Research*, 3, 157. <https://doi.org/10.12688/wellcomeopenres.14738.3> PubMed
46. Basu, T., Kumar, S., Kalyan, A., Jayaswal, P., Goyal, P., Pettifer, S., & Jonnalagadda, S. R. (n.d.). A novel framework to expedite systematic reviews by automatically building information extraction training corpora.
47. Ramezani, M., Kalivarapu, V., Gilbert, S. B., Huffman, S., Cotos, E., & O'Conner, A. (2017). Rapid tagging and reporting for functional language extraction in scientific articles. *Proceedings of the 6th International Workshop on Mining Scientific Publications*, 34–39. <https://doi.org/10.1145/3127526.3127533>
48. Marshall, I. J., Kuiper, J., & Wallace, B. C. (2015, July). Automating risk of bias assessment for clinical trials. *IEEE Journal of Biomedical and Health Informatics*, 19(4), 1406–1412. <https://doi.org/10.1109/JBHI.2015.2431314> PubMed
49. Marshall, I. J., Kuiper, J., & Wallace, B. C. (2016, January). RobotReviewer: Evaluation of a system for automatically assessing bias in clinical trials. *Journal of the American Medical Informatics Association : JAMIA*, 23(1), 193–201. <https://doi.org/10.1093/jamia/ocv044> PubMed
50. Gates, A., Johnson, C., & Hartling, L. (2018, March 12). Technology-assisted title and abstract screening for systematic reviews: A retrospective evaluation of the Abstrackr machine learning tool. *Systematic Reviews*, 7(1), 45. <https://doi.org/10.1186/s13643-018-0707-8> PubMed
51. Pustejovsky, J. E., & Tipton, E. (2022, April). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science*, 23(3), 425–438. <https://doi.org/10.1007/s11121-021-01246-3> PubMed
52. Mathur, M. B., & VanderWeele, T. J. (2020, November). Sensitivity analysis for publication bias in meta-analyses. *Applied Statistics*, 69(5), 1091–1119. <https://doi.org/10.1111/rssc.12440> PubMed
53. Stansfield, C., Thomas, J., & Kavanagh, J. (2013, September). 'Clustering' documents automatically to support scoping reviews of research: A case study. *Research Synthesis Methods*, 4(3), 230–241. <https://doi.org/10.1002/jrsm.1082> PubMed
54. Millard, L. A., Flach, P. A., & Higgins, J. P. (2016, February). Machine learning to assist risk-of-bias assessments in systematic reviews. *International Journal of Epidemiology*, 45(1), 266–277. <https://doi.org/10.1093/ije/dyv306> PubMed
55. Clark, J., Glasziou, P., Del Mar, C., Bannach-Brown, A., Stehlik, P., & Scott, A. M. (2020, May). A full systematic review was completed in 2 weeks using automation tools: A case

- study. *Journal of Clinical Epidemiology*, 121, 81–90.  
<https://doi.org/10.1016/j.jclinepi.2020.01.008> PubMed
56. Clark, J., McFarlane, C., Cleo, G., Ishikawa Ramos, C., & Marshall, S. (2021, May 31). The impact of systematic review automation tools on methodological quality and time taken to complete systematic review tasks: Case study. *JMIR Medical Education*, 7(2), e24418.  
<https://doi.org/10.2196/24418> PubMed
57. Scott, A. M., Glasziou, P., & Clark, J. (2023, May). We extended the 2-week systematic review (2weekSR) methodology to larger, more complex systematic reviews: A case series. *Journal of Clinical Epidemiology*, 157, 112–119.  
<https://doi.org/10.1016/j.jclinepi.2023.03.007> PubMed
58. Haddaway, N. R., Callaghan, M. W., Collins, A. M., Lamb, W. F., Minx, J. C., Thomas, J., & John, D. (2020, November 12). On the use of computer-assistance to facilitate systematic mapping. *Campbell Systematic Reviews*, 16(4), e1129. <https://doi.org/10.1002/cl2.1129> PubMed
59. van Altena, A. J., Spijker, R., & Olabarriaga, S. D. (2019, March). Usage of automation tools in systematic reviews. *Research Synthesis Methods*, 10(1), 72–82.  
<https://doi.org/10.1002/jrsm.1335> PubMed
60. Miranda, J. M., Muñoz, M., Uribe, E., Márquez, J., Uribe, G., & Valtierra, C. (2014). Systematic review tool to support the establishment of a literature review. In Á. Rocha, A. M. Correia, F. B. Tan, & K. A. Stroetmann (Eds.), *New Perspectives in Information Systems and Technologies, Volume 1* (Vol. 275, pp. 171–181). Springer International Publishing.  
[https://doi.org/10.1007/978-3-319-05951-8\\_17](https://doi.org/10.1007/978-3-319-05951-8_17)

---

Copyright (c) 2023 Delaware Academy of Medicine / Delaware Public Health Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.