

# The Utility of Artificial Intelligence for Systematic Reviews and Boolean Query Formulation and Translation

Leonardo Alaniz, BBA\*†; Cindy Vu, BS\*; Miles J. Pfaff, MD, MHS†‡

The formulation of precise search strings and syntax translation between databases is pivotal to a successful systematic review. By using the subscription-based ChatGPT,<sup>1</sup> researchers can harness natural language processing to generate refined search strings and Boolean queries, streamlining the process and enhancing the accuracy of the systematic review. Although there are few publications addressing ChatGPT's utility as a general research aid,<sup>2-4</sup> there are no reports on how ChatGPT-4 can be leveraged as a collaborative tool to collectively build a search string and translate Boolean queries into other database languages.

We conducted a mock systematic review describing the utility of augmented and virtual reality systems for preoperative planning in craniosynostosis surgery. We compiled a list of medical subject heading [MeSH] terms and keywords relevant to the research question.

Next, we incorporated ChatGPT-4 as a Boolean query aid to formulate an optimal PubMed search string (Table 1). Its first task was to analyze our study aim and provide additional databases to query for relevant articles. It then assisted us in refining our list of relevant MeSH terms and keywords for PubMed. Although it successfully added terms and enhanced our term list, researchers should still cross-reference MeSH terms with the database given ChatGPT-4's training cutoff date of 2021. (See table 1, Supplemental Digital Content 1, which shows a comprehensive list of original and modified MeSH terms and keywords. <http://links.lww.com/PRSGO/C836>.)

Its next task was to formulate a Boolean query for PubMed, which was comprehensive and yielded 766 results. After instructing it to minimize the number of irrelevant articles, ChatGPT-4 changed certain MeSH terms to keywords and added field tags, yielding 435 results.

*From the \*School of Medicine, University of California Irvine, Irvine, Calif.; †Department of Plastic Surgery, University of California Irvine Medical Center, Orange, Calif.; and ‡Pediatric Plastic Surgery, Children's Hospital of California, Orange, Calif.*

*Received for publication June 30, 2023; accepted September 6, 2023.*

*Copyright © 2023 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of The American Society of Plastic Surgeons. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.*

*Plast Reconstr Surg Glob Open 2023; 11:e5339; doi: 10.1097/GOX.0000000000005339; Published online 30 October 2023.*

ChatGPT-4's final task was to translate the PubMed Boolean query into a search string compatible with other databases (Fig. 1). All translated search strings were reliable and produced relevant results. Google Scholar, as ChatGPT-4 warned, was the only exception, as it does not support search strings.

To confirm the generalizability across diverse topics, we implemented our process on existing systematic reviews published in *Plastic and Reconstructive Surgery* within the reconstructive and cosmetic sections. (See table 2, Supplemental Digital Content 2, which shows five validation examples. <http://links.lww.com/PRSGO/C837>.) Our process was successfully able to recreate comprehensive search strategies and Boolean query translations that aligned with the author-generated strategies.

Although the power of large language models (LLM's) can allow human researchers to focus their efforts on study methodology and data analysis, it is essential to proceed cautiously. Many LLM's training datasets have a cut-off date (eg, 2021) and can therefore result in incomplete information. As such, researchers should always verify AI-assisted work and references to ensure both accuracy and relevance.<sup>5</sup> At all times, the use of AI tools must be disclosed to ensure transparency, and they cannot be authors of any produced content.

Although ChatGPT should not replace researchers, we have demonstrated how it can enhance and streamline the process of conducting thorough literature searches. Ultimately, the collaborative effort between AI and human expertise has the potential to augment research output.

Miles J. Pfaff, MD, MHS  
Department of Plastic Surgery  
University of California, Irvine  
200 S. Manchester Ave, Suite 650  
Orange, CA 92868  
E-mail: [mpfaff@hs.uci.edu](mailto:mpfaff@hs.uci.edu)

## DISCLOSURE

*The authors have no financial interest to declare in relation to the content of this article.*

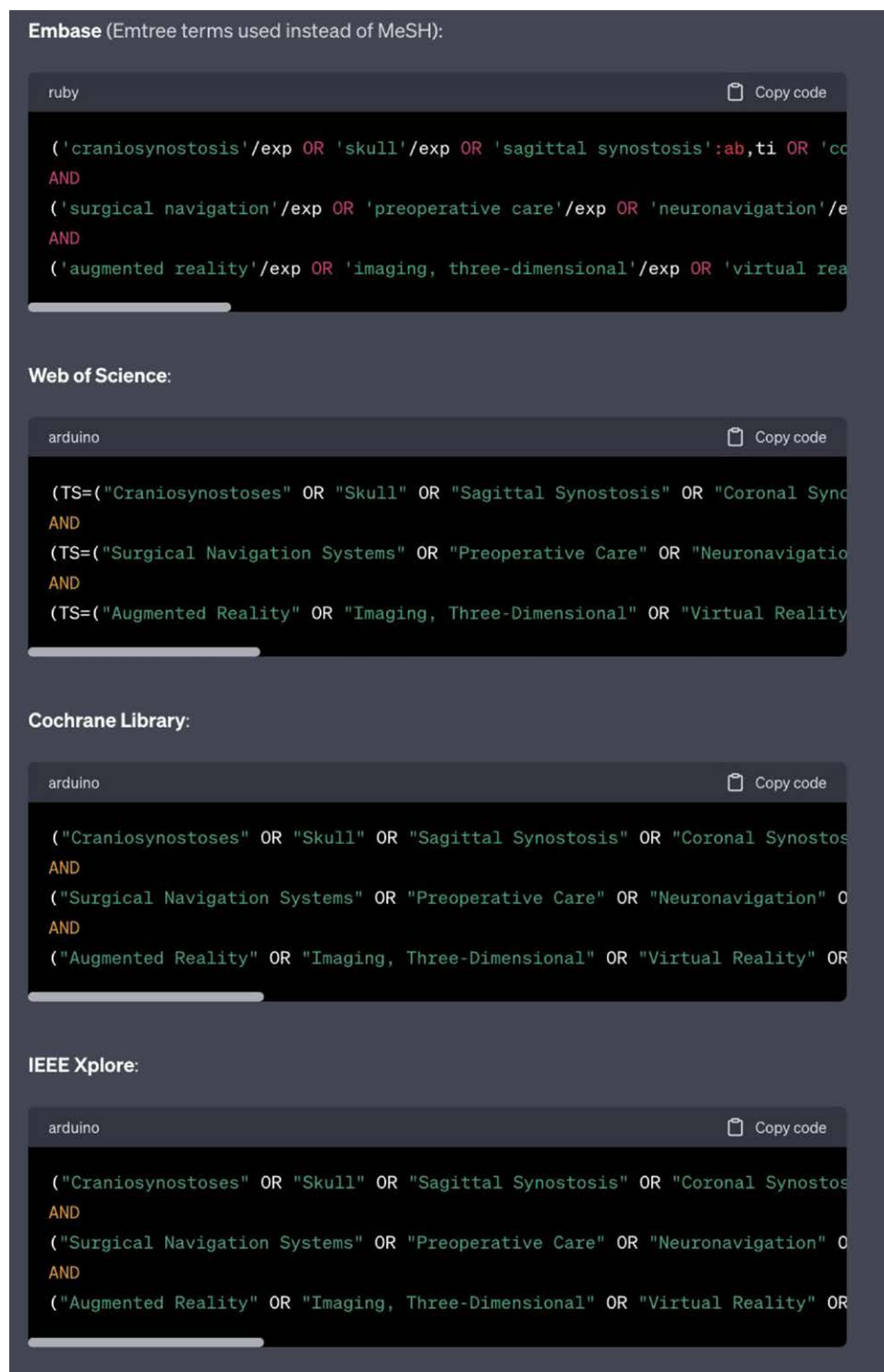
## REFERENCES

1. OpenAI. GPT-4 language model. March 14, 2023. Available at <https://openai.com/research/gpt-4/>. Accessed June 11, 2023.
2. Wang S, Scells H, Koopman B, et al. Can ChatGPT write a good Boolean query for systematic review literature search? Available

Related Digital Media are available in the full-text version of the article on [www.PRSGlobalOpen.com](http://www.PRSGlobalOpen.com).

**Table 1. ChatGPT-4 Prompts and Query Formulation Results**

Step	Prompt	Response	PubMed Yield
1	<p>We are conducting a systematic review and meta-analysis describing the utility of augmented and virtual systems for preoperative planning in craniostylosis surgery. We are first going to start with a search on PubMed and then proceed to other databases. We have two questions: To ensure we include as most relevant articles as possible and as few irrelevant articles as possible, which other databases (aside from PubMed), can we conduct a literature search on?</p> <p><i>[provided our original MeSH terms and keywords]</i></p> <p>What other MeSH Terms or keywords available on PubMed, can we add to include a more comprehensive search and ensure we don't miss any key articles? What MeSH Terms or keywords from the list above should we remove to ensure we obtain as few irrelevant articles as possible?</p>	<p>Certainly! Your existing search strategy seems quite comprehensive already. Here are some additional suggestions for other databases, as well as potentially relevant MeSH terms and keywords.</p> <ol style="list-style-type: none"> <li>1. Other databases you might want to consider include: Embase, Web of Science, Cochrane Library, IEEE Xplore, ACM Digital Library, CINAHL, Scopus, and Google Scholar</li> <li>2. MeSH terms and keywords: <ul style="list-style-type: none"> <li>-For the Anatomical Area/Procedure, consider adding MeSH terms and keywords like: <ul style="list-style-type: none"> <li>• MeSH terms: "Skull"[Mesh], "Cranial Sutures"[Mesh]</li> <li>• Keywords: suturoctomy, craniofacial anomalies</li> </ul> </li> <li>-For preoperative planning: <ul style="list-style-type: none"> <li>• MeSH terms: "Preoperative Care"[Mesh], "Neuronavigation"[Mesh], "Computer Simulation"[Mesh]</li> <li>• Keywords: preoperative simulation, surgical simulation, preoperative assessment, surgical rehearsal</li> </ul> </li> <li>-For augmented reality and virtual reality: <ul style="list-style-type: none"> <li>• MeSH Terms: "Computer Simulation"[Mesh], "Imaging, Three-Dimensional"[Mesh]</li> <li>• Keywords: Computer-assisted surgery, mixed reality, haptic feedback, 3D simulation, 3D printing</li> </ul> </li> </ul> </li> </ol>	—
2	<p>Based on the following list of MeSH terms and key words, can you create a Boolean Query (ie, search string) for PubMed to ensure we maximize the number of relevant articles and minimize the number of irrelevant articles? Remember, this is for a systematic review describing the utility of augmented and virtual systems for preoperative planning in craniostylosis surgery. In other words, we are interested in assessing how virtual or augmented reality can help with planning for craniostylosis. Please use Boolean operators, search field tags, proximity searching, and wildcards as needed to optimize the search.</p> <p><i>[We then provided updated list of MeSH Terms and Key Words with modifications from Step 1.]</i></p>	<p>Sure, I can help create a search strategy using Boolean operators, search field tags, proximity searching, and wildcards.</p> <p>((("Craniosynostoses"[Mesh] OR "Skull"[Mesh] OR "Cranial Sutures"[Mesh] OR "Sagittal Synostosis"[Tiab] OR "Coronal Synostosis"[Tiab] OR "Metopic Synostosis"[Tiab] OR "Craniosynostosis"[Tiab] OR "Suturoctomy"[Tiab] OR "Craniofacial Anomalies"[Tiab])) AND ("Surgical Navigation Systems"[Mesh] OR "Preoperative Care"[Mesh] OR "Neuronavigation"[Mesh] OR "Computer Simulation"[Mesh] OR "Preoperative Planning"[Tiab] OR "Virtual Surgical Planning"[Tiab] OR "Preoperative Simulation"[Tiab] OR "Surgical Simulation"[Tiab] OR "Preoperative Assessment"[Tiab] OR "Surgical Rehearsal"[Tiab]) AND ("Augmented Reality"[Mesh] OR "Computer Simulation"[Mesh] OR "Imaging, Three-Dimensional"[Mesh] OR "Virtual Reality"[Mesh] OR "Augmented Reality"[Tiab] OR "AR"[Tiab] OR "Computer-Assisted Surgery"[Tiab] OR "Mixed Reality"[Tiab] OR "Haptic Feedback"[Tiab] OR "3D Simulation"[Tiab] OR "3D Printing"[Tiab] OR "Virtual Reality"[Tiab] OR "VR"[Tiab]))</p>	766
3	<p>This search strategy resulted in too many results: 766, many of which are irrelevant. Can you help refine the search strategy to focus our results and minimize irrelevant studies? Perhaps you can try removing and/or changing some of the more general MeSH terms simply into keywords if you think that is a good idea.</p>	<p>Certainly, here's a revised strategy:</p> <p>("Craniosynostoses"[Mesh] OR "Skull"[Mesh] OR "Sagittal Synostosis"[tiab] OR "Coronal Synostosis"[tiab] OR "Metopic Synostosis"[tiab] OR "Craniosynostosis"[tiab] OR "Suturoctomy"[tiab]) AND ("Surgical Navigation Systems"[Mesh] OR "Preoperative Care"[Mesh] OR "Neuronavigation"[Mesh] OR "Computer Simulation"[Mesh] OR "Preoperative Planning"[tiab] OR "Virtual Surgical Planning"[tiab] OR "Preoperative Simulation"[tiab] OR "Surgical Simulation"[tiab] OR "Preoperative Assessment"[tiab] OR "Surgical Rehearsal"[tiab]) AND ("Augmented Reality"[Mesh] OR "Imaging, Three-Dimensional"[Mesh] OR "Virtual Reality"[Mesh] OR "Augmented Reality"[tiab] OR "AR"[tiab] OR "Computer-Assisted Surgery"[tiab] OR "Mixed Reality"[tiab] OR "3D Simulation"[tiab] OR "3D Printing"[tiab] OR "Virtual Reality"[tiab] OR "VR"[tiab])</p>	435
4	<p>Can you translate the following PubMed search string into search strings that can be used for the following databases: Embase, Web of Science, Cochrane Library, IEEE Xplore, ACM Digital Library, Scopus, and Google Scholar. Please be as accurate as possible, and do not add additional terms that were not included in the original search string unless it is really needed.</p>	<p>Please note that although I'm translating your search string, the databases you mentioned may not support MeSH terms or the same level of Boolean logic as PubMed does. Also, it's important to note that each database has its own unique search functionality, which may require adjustments to the syntax. Below are approximations for the search strings in each of the databases.</p> <p><i>[See Figure 1 for example of translations]</i></p>	—



**Fig. 1.** Translation of search string from PubMed to other databases.

- at <https://arxiv.org/abs/2302.03495>. Published online February 9, 2023. Accessed June 6, 2023.
- Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthc Basel Switz.* 2023;11:887.
  - Aydın Ö, Karaarslan E. OpenAI ChatGPT generated literature review: digital twin in healthcare. In: Aydın Ö, ed. *Emerging Computer Technologies 2*. İzmir Akademi Dernegi; 2022:22–31.
  - ElHawary H, Gorgy A, Janis JE. Large language models in academic plastic surgery: the way forward. *Plast Reconstr Surg Glob Open.* 2023;11:e4949.