INTRODUCTION

- Literature reviews have many applications in health economics and outcomes research. However, they are limited in breadth and depth by the amount of time reviewers spend and are prone to human error and biases. Natural language processing (NLP) aims to address these issues.
- Axtria reviewed the use of NLP in literature reviews, assessed its benefits and detriments, administered Axtria’s own test case, and developed recommendations for future researchers.

METHODS

- To identify use cases and information on the use of NLP in literature reviews, Axtria searched medical literature databases like PubMed, Science Direct, and Google Scholar; conference abstract lists; and other gray literature. The identified relevant studies are summarized herein.
- NLP was further implemented to conduct screening. Experts in systematic literature review were then consulted regarding the application of NLP to established literature review processes.

TRADITIONAL LITERATURE REVIEW

- In life sciences, literature reviews serve as the foundation for many information-oriented processes. They guide researchers and pharmaceutical companies by providing insights into unmet patient needs and current markets.
- Literature reviews are essential for evaluators such as regulatory bodies and health technology assessment (HTA) bodies, preventing unintentional duplication of research and identifying gaps in existing knowledge.

CHALLENGES WITH TRADITIONAL LITERATURE REVIEWS

- Screening requires significant human labor and introduces risks. Manual data extraction is time-consuming and prone to human error.
- TLRs grapple with limitations in scope and depth, which are intrinsically tied to the time and effort invested by reviewers.
- SLRs are often not updated frequently due to their labor-intensive nature and can be impractical for topics with rapidly evolving scopes.
- SLRs are essential for compiling comprehensive evidence. Automation and NLP allow researchers to focus on critical appraisal and synthesis.

NATURAL LANGUAGE PROCESSING AND ITS APPLICATION IN LITERATURE REVIEW

- Fortunately, advances in artificial intelligence (AI) offer potential solutions to these challenges. NLP offers advantages such as prioritizing articles based on relevance and subtopics. However, its implementation requires technical expertise, and there’s a risk of misclassifying studies due to rigid criteria. It’s crucial to scrutinize discrepancies between AI and human decisions. Furthermore, third-party AI services may not always adhere to established guidelines like PRISMA.
- The key advantages and disadvantages of implementing NLP methods in literature reviews, either alone or in combination with human direction, are summarized below.

PREVIOUS APPLICATIONS AND AXTRIA’S USE CASE

- A systematic review analyzed the potential of NLPs to assist with the SLR process using 3 LVM. GPT-4 achieved a decision match rate of 71.0%, outperforming AZU1 (92.0%) and ModelBone (87.7%).
- Another study validated the AI classifier tool used by the DistillerSR platform, and found an 87.5% match rate with human decisions.
- A pilot study of the cancer-specific, NL-based information retrieval software, Axtria, achieved 95.9% accuracy in identifying relevant articles.
- In Axtria’s use case, GPT-4 attained a 100% accuracy rate in screening a sample of 80 articles for relevance by therapeutic area.

APPLICATIONS OF LITERATURE REVIEW

- Targeted literature reviews (TLRs) vary in adherence to standardized processes, scope, and expertise. Systematic literature reviews (SLRs) are conducted using inclusion/exclusion criteria and databases like PubMed, Embase, Cochrane Library, etc. SLRs are considered more powerful due to their systematic andrepeatable nature, while TLRs are valued because they are fast and require minimal value.

DISCLOSURES

- CM & DH are employees of Axtria Inc., NJ, USA.
- This study was funded by Axtria Inc.

REFERENCES

1. Herrin R, Brink A, Barbaresi DC, Kondili G. A comparative analysis of large literature searches in evidence-based literature review. Poster presentation at ISPOR Annual European Meeting; 13 Nov 2023; Copenhagen, Denmark. Accessed 63x117 0x2272 to 3456x2536.
6. GPT-3.5. Search performed on: 51.0% 67.3% 71.1% 100.0% 87.5% 55.9% 100.0%. A recent study validating DistillerSR’s AI classifier for systematic reviews estimated time savings at 55% of human working hours. Another study using PubMed BERT demonstrated a 45.9% decrease in screening time per abstract.

CONCLUSIONS

- There are many promising developments in NL applications for literature review/synthesis, which can improve the depth and breadth of literature reviews while reducing human labor and the risk of bias and error.
- However, since the NLP model’s accuracy and reliability vary, literature reviewers should implement NLP cautiously, giving precise instructions and sufficient training, verifying NLP decisions, and following practice guidelines where possible.
- Last, optimal NLP usage in literature review/synthesis necessitates a certain degree of technical expertise in conjunction with a strong understanding of scientific and medical writing. Therefore, Axtria recommends consultation with technical and clinical experts when implementing NLP in literature reviews.

POSTER PRESENTED AT

ISPOR US, Atlanta, Georgia, USA; 5 May – 8 May 2024.