SOVEREIGN: To what extent does EVOR's diverse knowledge base evolution improve LLM performance on out-of-distribution cod

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 29, 2026

Abstract

ChatGPT is an artificial intelligence (AI)-based conversational large language model (LLM). The potential applications of LLMs in health care education, research, and practice could be promising if the associated valid concerns are proactively examined and addressed. The current systematic review aimed to investigate the utility of ChatGPT in health care education, research, and practice and to highlight its potential limitations. Using the PRIMSA guidelines, a systematic search was conducted to retrieve English records in PubMed/MEDLINE and Google Scholar (published research or preprints) tha

1 Introduction

Analysis of: ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Research goal: To what extent does EVOR's diverse knowledge base evolution improve LLM performance on out-of-distribution code generation challenges?.

2 Methodology

Multi-query arXiv search (4 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score $<$ 6.5.

3 Results

7 papers retrieved. 10 claims extracted, 10 verified. Tribunal: 8.0/10 → APPROVE (revision_round=0). Policy: AUTO_APPROVE.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

5 Extracted Claims

Claim	Verified	Confidence
ChatGPT is an artificial intelligence (AI)-based conversational large language model (LLM).	✓	0.22
The potential applications of LLMs in health care education, research, and practice could be promising if the associated	✓	0.43
A total of 60 records were eligible for inclusion in the systematic review.	✓	0.24
Benefits of ChatGPT were cited in 51/60 (85.0%) records.	✓	0.23
Concerns regarding ChatGPT use were stated in 58/60 (96.7%) records.	✓	0.25
Benefits of ChatGPT include improved scientific writing and enhancing research equity and versatility.	✓	0.22
Benefits of ChatGPT include utility in health care research such as efficient analysis of datasets, code generation, lit	✓	0.38
Benefits of ChatGPT include benefits in health care practice such as streamlining the workflow, cost saving, documentati	✓	0.35
Benefits of ChatGPT include benefits in health care education including improved personalized learning and the focus on	✓	0.36
Concerns regarding ChatGPT use include ethical, copyright, transparency, and legal issues.	✓	0.20

References

- https://doi.org/10.3390/healthcare11060887
- https://doi.org/10.1145/3639372
- https://doi.org/10.48550/arxiv.2312.10997