

To what extent does domain-specific fine-tuning of the retriever and generator components in RAG-end2end reduce

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: To what extent does domain-specific fine-tuning of the retriever and generator components in RAG-end2end reduce hallucination rates compared to in-domain prompting for 7B and 70B models. 5 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: THaMES: An End-to-End Tool for Hallucination Mitigation and Evaluation in Large Language Models. Research question: To what extent does domain-specific fine-tuning of the retriever and generator components in RAG-end2end reduce hallucination rates compared to in-domain prompting for 7B and 70B models?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.3/10.

3 Results

16 papers retrieved. 5 claims extracted; 0 independently verified. Quality review score: 3.3/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
THaMES framework consists of three main components: Testset generation, Baseline metric evaluation, and Mitigation strat	×	0.06
The QA-set generation process in THaMES includes seven steps: Knowledge Base Processing, Ground-Truth Weighted Sampling,	×	0.08
THaMES is designed to process various corpora including political news articles, academic papers, and Wikipedia articles	×	0.09
THaMES is compatible with various file formats including PDF, TXT, and CSV.	×	0.03
THaMES utilizes the VectorStoreIndex module provided by LlamaIndex to build a knowledge base out of the raw corpus.	×	0.04

References

- <http://arxiv.org/abs/2409.11353v3>
- <http://arxiv.org/abs/2601.19927v1>
- <http://arxiv.org/abs/2210.02627v1>