

Blended RAG with Hybrid Retrievers vs. Traditional RAG on HotpotQA at Scale

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the performance of Blended RAG with hybrid retrievers compare to traditional RAG methods on the HotpotQA benchmark when scaling the document corpus from 10K to 100K documents. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 1.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Blended RAG: Improving RAG (Retriever-Augmented Generation) Accuracy with Semantic Search and Hybrid Query-Based Retrievers. Research question: How does the performance of Blended RAG with hybrid retrievers compare to traditional RAG methods on the HotpotQA benchmark when scaling the document corpus from 10K to 100K documents?.

2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 1.7/10.

3 Results

13 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 1.7/10.

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.

References

- <http://arxiv.org/abs/2404.07220v2>
- <http://arxiv.org/abs/2510.25621v1>
- <http://arxiv.org/abs/2502.11228v2>