

# Hybrid vs. Sparse Retrievers in Blended RAG on Large-Scale TriviaQA

Assignee Research

June 9, 2026

## Abstract

This report synthesises findings from 7 peer-reviewed papers addressing the following research question: How does the recall@K performance of hybrid retrievers in Blended RAG compare to sparse-only retrievers on the TriviaQA benchmark when the document corpus exceeds 1M entries. 11 claims were extracted from source literature; 8 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: A Survey on Retrieval-Augmented Text Generation for Large Language Models. Research question: How does the recall@K performance of hybrid retrievers in Blended RAG compare to sparse-only retrievers on the TriviaQA benchmark when the document corpus exceeds 1M entries?.

## 2 Methodology

Systematic literature search across multiple databases yielded 7 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.3/10.

## 3 Results

7 papers retrieved. 11 claims extracted; 8 independently verified. Quality review score: 7.3/10.

## 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
Retrieval-Augmented Generation (RAG) merges information retrieval (IR) techniques with deep learning advancements.	✓	0.31
RAG addresses the static limitations of large language models (LLMs) by enabling the dynamic integration of up-to-date e	✓	0.29
The surveyed methodology focuses primarily on the text domain.	×	0.11
RAG provides a cost-effective solution to the generation of plausible but possibly incorrect responses by LLMs.	✓	0.28
RAG enhances the accuracy and reliability of LLM outputs through the use of real-world data.	✓	0.17
The article organizes the RAG paradigm into four categories: pre-retrieval, retrieval, post-retrieval, and generation.	✓	0.30
The article offers a detailed perspective on RAG from the retrieval viewpoint.	×	0.14
The article outlines RAG's mechanics and discusses the field's progression through the analysis of significant studies.	✓	0.26
The article introduces evaluation methods for RAG.	✓	0.19
The article addresses challenges faced in RAG and proposes future research directions.	×	0.14
The study aims to consolidate existing research on RAG and clarify its technological underpinnings.	✓	0.21

## References

- <https://doi.org/10.1145/3805774>

- <https://doi.org/10.48550/arxiv.2308.07107>
- <https://doi.org/10.48550/arxiv.2402.09906>