

# SOVEREIGN: Can a dynamic passage count per hop in iterative retrieval improve accuracy on the MuSiQue multi-hop benchmark

SOVEREIGN Research Kernel

Autonomous draft — Owner review required before publication

May 29, 2026

## Abstract

Large Language Models (LLMs) showcase impressive capabilities but encounter challenges like hallucination, outdated knowledge, and non-transparent, untraceable reasoning processes. Retrieval-Augmented Generation (RAG) has emerged as a promising solution by incorporating knowledge from external databases. This enhances the accuracy and credibility of the generation, particularly for knowledge-intensive tasks, and allows for continuous knowledge updates and integration of domain-specific information. RAG synergistically merges LLMs' intrinsic knowledge with the vast, dynamic repositories of exte

## 1 Introduction

Analysis of: Retrieval-Augmented Generation for Large Language Models: A Survey. Research goal: Can a dynamic passage count per hop in iterative retrieval improve accuracy on the MuSiQue multi-hop benchmark while maintaining comparable inference latency to fixed-passage retrieval?.

## 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

10 papers retrieved. 6 claims extracted, 6 verified. Tribunal: 7.5/10  $\rightarrow$  APPROVE (revision\_round=1). 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
Retrieval-Augmented Generation (RAG) has emerged as a promising solution by incorporating knowledge from external databases	✓	0.34
RAG enhances the accuracy and credibility of the generation, particularly for knowledge-intensive tasks	✓	0.25
RAG allows for continuous knowledge updates and integration of domain-specific information	✓	0.24
RAG synergistically merges LLMs' intrinsic knowledge with the vast, dynamic repositories of external databases	✓	0.32
The progression of RAG paradigms includes Naive RAG, Advanced RAG, and Modular RAG	✓	0.23
RAG frameworks have a tripartite foundation including retrieval, generation and augmentation techniques	✓	0.21

## References

- <https://doi.org/10.18653/v1/2023.findings-emnlp.378>
- <https://doi.org/10.1136/bmj.g7647>
- <https://doi.org/10.48550/arxiv.2312.10997>