

Multi-Stage Retrieval with Hierarchical Filtering Enhances LLM Faithfulness in Religious QA

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the integration of a multi-stage retrieval system with hierarchical filtering impact the faithfulness and accuracy of 7B vs 70B LLMs on religious QA benchmarks compared to single-pass RAG. 10 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: FARSIQA: Faithful and Advanced RAG System for Islamic Question Answering. Research question: How does the integration of a multi-stage retrieval system with hierarchical filtering impact the faithfulness and accuracy of 7B vs 70B LLMs on religious QA benchmarks compared to single-pass RAG pipelines?.

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 5.2/10.

3 Results

13 papers retrieved. 10 claims extracted; 2 independently verified. Quality review score: 5.2/10.

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
FARSIQA achieves an 'Answer Correctness' score of 74.3% on the IslamicPCQA benchmark as evaluated via LLM-as-Judge on mu	✓	0.17
FARSIQA achieves a Negative Rejection score of 97.0%.	×	0.10
FARSIQA's Negative Rejection score represents a 40-point improvement over standard baselines.	×	0.13
The FARSIQA knowledge base was curated from eleven reputable online Persian Islamic encyclopedias, including WikiShia, W	×	0.08
The collection of Islamic Encyclopedias resulted in approximately 431,000 unique documents.	×	0.03
Data from authoritative Q&A websites IslamQuest.net and porseman.com was incorporated into the knowledge base.	×	0.04
The Religious Q&A Platforms corpus added approximately 304,000 question-answer pairs to the knowledge base.	×	0.04
The total curated knowledge base contains over one million documents.	×	0.08
FARSIQA is built upon the FAIR-RAG architecture, which stands for Faithful, Adaptive, Iterative Refinement framework for	✓	0.24
For encyclopedic articles, the text was split by paragraphs to maintain semantic cohesion during preprocessing.	×	0.03

References

- <http://arxiv.org/abs/2510.25621v1>
- <http://arxiv.org/abs/2504.01346v4>

- <http://arxiv.org/abs/2404.14464v1>