

# Blended RAG Robustness with Hybrid Retrievers Across Scientific, Legal, and Technical Domains

Assignee Research

June 8, 2026

## Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the robustness of Blended RAG with hybrid retrievers vary across different domains (e.g., scientific, legal, and technical) as measured by accuracy on the NQ (Natural Questions) benchmark. 15 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: MetaGen Blended RAG: Unlocking Zero-Shot Precision for Specialized Domain Question-Answering. Research question: How does the robustness of Blended RAG with hybrid retrievers vary across different domains (e.g., scientific, legal, and technical) as measured by accuracy on the NQ (Natural Questions) benchmark?.

## 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 4.2/10.

## 3 Results

13 papers retrieved. 15 claims extracted; 0 independently verified. Quality review score: 4.2/10.

## 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 |
|--|----------|------------|
| The baseline retrieval accuracy using plain k-NN semantic search without metadata was approximately 74.9%.                 | ×        | 0.07       |
| Integrating lexical (BM25) search with semantic (k-NN) search into a hybrid retrieval framework improved the accuracy to   | ×        | 0.07       |
| Incorporating existing metadata fields such as Medical Subject Headings (MeSH terms) improved retrieval accuracy to 78.8   | ×        | 0.04       |
| Enriched metadata fields derived through NLP techniques and LLM-generated semantic annotations improved retrieval accuracy | ×        | 0.08       |
| Combining enriched metadata with existing metadata improved retrieval accuracy to 79.7%.                                   | ×        | 0.04       |
| Using boosted queries with only enriched metadata improved accuracy to 80.3%.  | ×        | 0.06       |
| Employing boosted queries with only existing metadata achieved 80.6% accuracy.   | ×        | 0.04       |
| The highest accuracy of 82.1% was attained by applying boosted hybrid retrieval that leveraged both existing and enriched  | ×        | 0.06       |
| The final configuration represents a substantial absolute gain of approximately 7.2 percentage points over the hybrid se   | ×        | 0.02       |
| The final configuration represents an approximately 10-point gain above the initial semantic (k-NN) baseline.              | ×        | 0.01       |
| Targeted metadata enrichment combined with hybrid query boosting substantially enhances retrieval precision.               | ×        | 0.07       |
| The approach is complementary to any model that is already fine-tuned and can further enhance the search accuracy.         | ×        | 0.09       |
| Most real-world document corpora do not have any organic metadata tagged to them.  | ×        | 0.06       |
| Generation of metadata can be very noisy and thereby drown the relevant parts with irrelevant keywords.                    | ×        | 0.02       |
| The investigation focuses on evaluating both retrieval and generative accuracy in the context of the PubMedQA dataset.     | ×        | 0.07       |

## References

- <http://arxiv.org/abs/2404.07220v2>
- <http://arxiv.org/abs/2402.13178v2>
- <http://arxiv.org/abs/2505.18247v3>