Retrieval-Augmented Generation and Zero-Shot Re-Ranking Hallucination Rates in Knowledge-Intensive Tasks

Assignee Research

June 2, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How do retrieval-augmented generation models and zero-shot re-ranking approaches differ in hallucination rates when evaluated on knowledge-intensive reasoning tasks. 15 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: FAIR-RAG: Faithful Adaptive Iterative Refinement for Retrieval-Augmented Generation. Research question: How do retrieval-augmented generation models and zero-shot re-ranking approaches differ in hallucination rates when evaluated on knowledge-intensive reasoning tasks?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.3/10.

3 Results

15 papers retrieved. 15 claims extracted; 1 independently verified. Quality review score: 5.3/10.

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
The Reciprocal Rank Fusion (RRF) algorithm is used to re-rank documents without requiring hyperparameter tuning.	×	0.01
The RRF algorithm produces a single, robustly ranked list of the top-5 most relevant documents as candidate evidence.	×	0.02
An LLM agent evaluates each document's utility with respect to the original user query and discards irrelevant, off-topi	×	0.02
The Structured Evidence Assessment (SEA) module uses a checklist-based methodology to perform granular gap analysis.	×	0.08
The SEA module deconstructs the user's query into a checklist of discrete, required informational components or 'finding	×	0.08
The SEA module systematically audits the collected evidence against the checklist, confirming which findings are support	×	0.08
The evidence is deemed sufficient only if all required findings are confirmed.	×	0.08
FAIR-RAG is designed to robustly handle complex queries by orchestrating a dynamic, multi-stage workflow.	×	0.08
FAIR-RAG begins with an Adaptive Routing module that analyzes query complexity to determine an optimal execution path.	×	0.06
For non-trivial queries, FAIR-RAG initiates a cyclical process designed to progressively build and validate a context.	×	0.06
The core of the FAIR-RAG cycle is an Iterative Refinement loop where LLM agents intelligently decompose information need	×	0.14
Each cycle in FAIR-RAG culminates in a Structured Evidence Assessment (SEA) module, which acts as an analytical gating m	✓	0.22
The SEA module emulates a cognitive workflow by first deconstructing the user's query into a checklist of required findi	×	0.08
The SEA module systematically synthesizes the retrieved evidence against the checklist, verifying what is confirmed and	×	0.07
The identified gaps in the SEA module provide a precise, actionable signal for the next iteration.	×	0.09

References

- http://arxiv.org/abs/2508.05197v2
- http://arxiv.org/abs/2503.16581v1
- http://arxiv.org/abs/2510.22344v1