

# Diversity Metrics in Retriever Portfolio Selection Predict Out-of-Domain Generalization

Assignee Research

June 9, 2026

## Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: Can diversity metrics derived from retriever portfolio selection on AmbiEval predict generalization performance on out-of-domain benchmarks like NaturalQuestions without additional fine-tuning. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: BEIR: A Heterogenous Benchmark for Zero-shot Evaluation of Information Retrieval Models. Research question: Can diversity metrics derived from retriever portfolio selection on AmbiEval predict generalization performance on out-of-domain benchmarks like NaturalQuestions without additional fine-tuning?.

## 2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.3/10.

## 3 Results

16 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 6.3/10.

## 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.

## References

- <http://arxiv.org/abs/2605.31176v1>
- <http://arxiv.org/abs/2104.08663v4>
- <http://arxiv.org/abs/2510.03260v1>