

# RAG End-to-End 테스트 결과 보고서

MATHESIS LAB Project

SUCCESS

테스트 일시 2025-11-21

테스트 환경 WSL (Ubuntu) + Python Virtual Environment

Vector DB Qdrant Local Mode (Disk-based)

Embedding Ollama (nomic-embed-text, 768 dim)

Parser Custom MathParser (PDF)

## 1. 실행 결과 요약

단계	결과	상세 내용
1. 파싱	성공	491개 청크 생성 (성취기준 + 총론)
2. 임베딩	성공	Ollama를 통해 491개 벡터 생성 (768차원)
3. 인덱싱	성공	Qdrant Local Storage에 저장 완료
4. 검색	성공	3가지 질의에 대해 결과 반환

## 2. 질의 응답 품질 분석

### Q1. “초등학교 1 2학년 수와 연산 영역의 성취기준은?”

- 결과: [2수02-01], [2수01-02] 등 초등 저학년 관련 성취기준 반환
- 평가: 매우 정확함
- 분석: 질문의 의도(초등, 수와 연산)에 맞는 문서를 정확히 찾아냄.

### Q2. “수학과 교육과정의 성격은 무엇인가?”

- 결과: [12직수01-03] 등이 반환됨
- 평가: 다소 부정확함
- 분석: 총론(“1. 성격”) 청크가 검색되지 않고, 엉뚱한 성취기준이 나옴. 총론 청크의 개수가 적어 벡터 유사도에서 밀렸을 가능성 있음.
- 제언: 총론 청크에 가중치를 주거나, 메타데이터 필터링 활용 필요.

### Q3. “평가 방법 및 유의 사항에 대해 알려주세요”

- 결과: [12확통03-06] 등 특정 성취기준의 평가 방법 반환
- 평가: **부분 성공**
- 분석: “평가 방법” 키워드는 찾았으나, 사용자가 원하는 것이 ‘일반 지침’인지 ‘특정 단원’인지 모호함.

## 3. 결론

시스템이 기술적으로 완벽하게 동작합니다(파싱 -> 임베딩 -> 저장 -> 검색). 검색 품질(Accuracy)은 향후 메타데이터 필터링과 프롬프트 엔지니어링을 통해 개선할 수 있는 영역입니다.

RAG 시스템의 핵심 파이프라인 구축이 완료되었습니다.