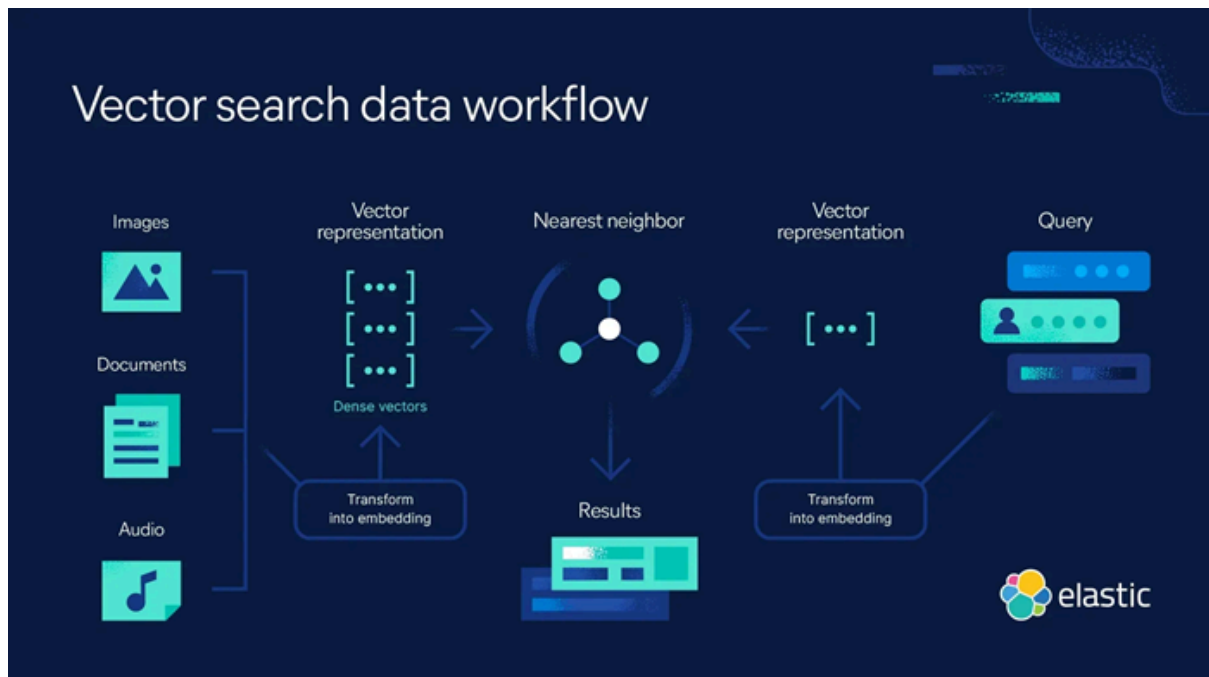


Q. 프롬프트 질문은 RAG 구성에서 어디에 전달되는가?



1. 검색어를 벡터화
2. 검색어 쿼리와 유사한 문서를 벡터DB에서 추출
3. 추출된 문서를 조합하여 최종 답변 생성

RAG 워크플로우에서 사용자의 프롬프트(질문)는 먼저 검색 단계로 전달됩니다.

즉, 입력된 질문은 임베딩 모델을 통해 벡터로 변환되어 벡터 DB에 쿼리됩니다.

RAG 파이프라인은 “사용자 쿼리를 임베딩하여 인덱싱된 문서에 유사도 검색을 수행하고, 가장 유사한 문서를 추출”하는 방식으로 동작합니다.

이후 검색된 문서들과 원래 질문이 함께 LLM에 주어져 답변을 생성하게 됩니다.

Q. RAG 챗봇 사례 중 실제 구현 예시는?

A. 사례1. VOC 형태의 민원 대응 시스템

예시: 인천국제공항공사에서 대외 홈페이지 민원 자동 응대를 위한 시스템 구성

특징: 외부 민원 접수 → 관련 문서 검색 → 자동 응답 생성 및 번역 → 사용자 응대

사례2. 내부 직원용 업무지원 챗봇