

# RAG 챗봇 데모

자율형 IoT 센터  
박지상

# 간단한 소개

- 박지상
- 정보 + 데이터 과학 전공 (2025.12 학사 졸업)
- 삼성SDI 인턴 (2025.06 - 2025.08)
  - RAG기반 챗봇 개발

# RAG

- Retrieval(검색) + Augmented(증강) + Generation(생성)
- “오픈북 테스트”
  - 기존 LLM: 책을 덮고 암기한 내용으로만 시험 봄 (틀릴 확률 높음).
  - RAG: 질문이 나오면 교과서(데이터)를 펼쳐서 정답 부분을 찾은 뒤, 보고 답을 씀 (정확도 높음).

# 원리

- 벡터 검색(Vector Search): 질문과 의미가 비슷한 데이터를 자석처럼 끌어당김
- 특정 정보 검색에 최적:
  - 질문: "환자 ID 12345의 혈압은?"
  - 결과: 여러개 데이터 중 '12345'와 관련된 딱 1개를 정확히 찾음

# 강점

파일 업로드

CSV 또는 Excel 파일을 업로드하세요.

Drag and drop file here.  
Limit: 200MB per file - CSC\_91.xlsx

[Browse files](#)

Heart\_Disease\_Pred...  
13,690

총 270개의 행을 읽었습니다.

총 270개의 조작으로 나누었습니다.

최종 저장된 데이터 수: 270개

**내 파일과 대화하기 (AnyData Chatbot)**

Heart\_Disease\_Prediction.csv 분석 완료! (270개의 데이터)

안녕하세요! CSV나 Excel 파일을 업로드해주시면 내용을 분석해 드릴게요.

나이가 29세인 환자의 “혈압(BP)”과 콜레스테롤 수치는 얼마야?

안녕하세요. 문의하신 29세 환자의 혈압(BP)은 130, 콜레스테롤 수치는 204입니다.  
참고 대치치: 28, 48, 45 등

- 질문: "나이가 29세인 환자의 혈압(BP)과 콜레스테롤 수치는 얼마야?"

216	29	1	2	130	204	0	2	202	0	0	1	0	3 Absence
-----	----	---	---	-----	-----	---	---	-----	---	---	---	---	-----------

# 약점

- 문제 상황: "환자들의 평균 나이는?"
- 실패 원인 (Sampling Bias):
  - RAG는 한 번에 소량의 조각(예: 5개)만 읽을 수 있음.
  - 전체를 다 보지 못하고, 검색된 5명만 보고 평균을 계산함.
- 결과: 실제 평균 대신 샘플 평균을 정답인 양 답변
- RAG는
  - 검색 도구 O
  - 통계 도구 X

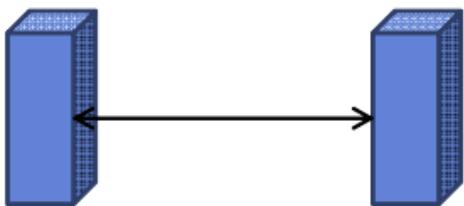
 환자들의 평균 나이는 몇살이야?

 환자들의 평균 나이를 직접적으로 계산할 수는 없지만, 제공된 데이터에 따르면 환자들의 나이는 42세에서 71세 사이입니다. 좀 더 정확한 평균 나이를 알고 싶으시다면, 전체 환자 나이 데이터를 제공해주시면 계산해 드릴 수 있습니다.

참고 데이터: 63, 67, 69 등

# 현재 모델의 한계

- 질문이 들어올시 두 문서 연결  
하지 못해 답변 실패



문서A: 철수는 A아파트에 산다

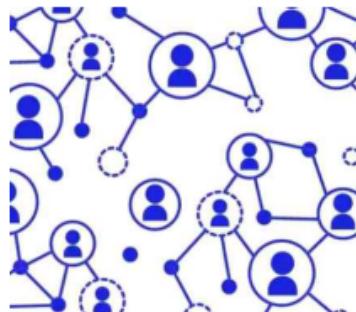
문서B: A 아파트는 작년에 침수 피해를 입었다  
사용자 질문: 철수네 집은 수해안전지대?

문서A: 철수 존재

문서B: 철수라는 단어 x

질문에는 A아파트라는 단어 x

- 하이브리드 RAG (Graph RAG)
- Vector DB(의미 검색) + Graph DB(관계 추론) = GraphRAG
- 숲(전체 구조)과 나무(세부 정보)  
를 동시에 파악.



# 해결책 - GraphRAG

- 현재의 Vector RAG는 나무(단일 정보)는 잘 보지만 숲(전체 관계/통계)은 못 본다.
- 해결: GraphRAG (지식 그래프) 도입.
- 효과:
  - 데이터 간의 연결 고리(Relationship) 파악.
  - "평균", "관계", "추론"이 필요한 질문까지 커버 가능.
- Vector RAG (현재): 비정형 문서/텍스트 분석 기반 마련.
- Graph RAG (관계): 시설물 간 연계성 및 인프라 네트워크 분석.
- Spatial RAG (공간): GIS 데이터 연계를 통한 위치 기반 분석.
- Agent (통합): 이 3가지를 상황에 맞춰 호출하고, 결과를 시각화하는 자율 AI 모듈 구축.

# 감사합니다

자율형IoT센터  
박지상