

광주 2팀

신현학  
권해림  
김서린  
이두호

# New Chats

AI와 검색의 만남,  
최신 정보를 빠르게



# 목차

1. 프로젝트 개요 및 목표
2. 문제 정의 및 해결방안
3. 주요 기능 및 기술적 특징
4. API 활용 방법
5. 프로젝트 임팩트
6. 향후 개발 계획



# 프로젝트 개요 및 목표

## 프로젝트 개요

- 사용자의 질문에 대해 네이버 및 구글의 최신 정보를 검색하고, 가장 관련성 높은 정보를 기반으로 AI가 답변을 생성하는 시스템.
- 검색과 AI 생성 모델을 결합한 RAG(Retrieval-Augmented Generation) 기반 어시스턴트.

## 프로젝트 목표

- 최신 정보 제공: 실시간 검색 데이터를 활용하여 최신 정보 기반 답변 생성.
- 정보 신뢰성 확보: 검증된 출처(네이버, 구글)의 데이터를 활용하여 신뢰성 있는 정보를 제공.
- 사용자 경험 향상: 자연스러운 대화와 유용한 정보를 제공하는 AI 어시스턴트를 구현.

# 문제 정의 및 해결방안

## 문제 정의

- AI 모델의 한계: 기존 AI 모델은 최신 정보와 실시간 데이터 반영에 한계가 있음.
- 검색의 어려움: 사용자가 직접 검색해야 하는 번거로움과 정보 과잉 문제.
- 정보 신뢰성: 검색 결과의 정확성과 관련성을 보장하기 어려움.

## 해결방안

- RAG 기반 시스템: 검색 엔진(네이버, 구글)과 AI 모델을 결합하여 실시간 정보 제공.
- 벡터 데이터베이스: 검색 결과를 분석하고 가장 관련성이 높은 내용을 추출하여 답변 생성.
- 다중 소스 활용: 네이버와 구글을 통합해 데이터 소스를 다양화.

# 주요 기능 및 기술적 특징

## 제작한 RAG의 주요 기능

### 1 사용자가 질문 입력

- 사용자의 메시지를 받아 처리 시작

### 2 키워드 추출 및 검색 시도

- 네이버 키워드 추출 → 뉴스 검색
- 네이버 결과가 없을 경우 구글 검색 실행

### 3 검색 결과 처리

- 결과를 문서화하고 벡터 DB로 변환

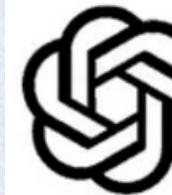
### 4 RAG 기반 답변 생성

- 검색 결과를 활용해 사용자 질문에 최적화된 답변 생성

### 5 응답 반환

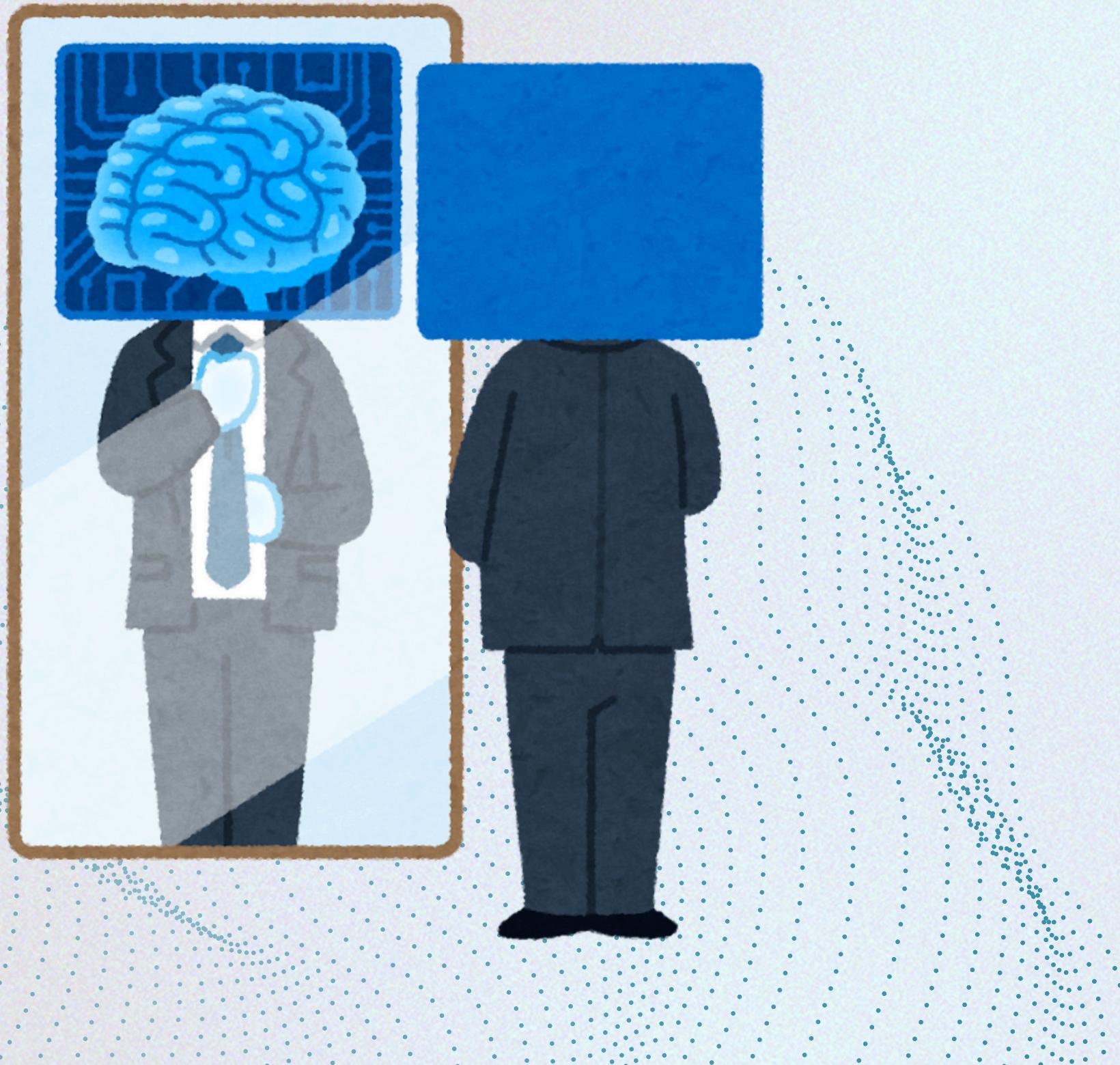
- JSON 형식으로 결과와 함께 응답 제공

+



**OpenAI Assistant API**  
with Retriever Tool

# API 활용 방법



## 1. 네이버 검색 API

- 사용 목적: 뉴스 검색.
- 활용 방식: 사용자 메시지에서 추출된 키워드를 사용해 HTTP 요청으로 검색.

## 2. Google SerpAPI

- 사용 목적: 검색 엔진에서 웹 데이터를 가져오기.
- 활용 방식: 사용자의 질문을 기반으로 질의를 재구성하여 API 호출.

## 3. Upstage API

- 사용 목적: AI 모델 호출 및 임베딩 생성.
- 활용 방식: 검색어 추출, 질의 확장, 벡터화 등 다양한 작업에 사용.

# 프로젝트 임팩트

## 1. 사용자 경험 개선

- 사용자가 직접 검색하지 않아도 원하는 정보를 빠르게 얻을 수 있음.
- 질문에 대한 명확하고 신뢰성 있는 답변 제공.

## 2. 효율적인 정보 활용

- 최신 뉴스와 실시간 데이터 활용으로 사용자 요구에 즉각 대응.
- 관련성이 높은 검색 결과를 기반으로 한 유의미한 답변 생성.

## 3. 기술적 확장성

- 새로운 검색 엔진 및 데이터 소스를 손쉽게 추가 가능.
- 벡터 DB와 RAG 구조를 활용하여 다른 도메인에 쉽게 적용 가능.

# 향후 개발 계획

## 1. 추가 데이터 소스 통합

- 다양한 뉴스 API 및 전문 데이터베이스 연동.
- 사용자 맞춤형 추천 기능 추가.

## 2. 사용자 인터페이스 개선

- 웹 및 모바일 UI 개발로 접근성 강화.
- 음성 입력 및 출력 기능 추가.

## 3. 모델 성능 향상

- 대규모 언어 모델(LLM)의 업데이트 및 최신 임베딩 기술 적용.
- 검색 결과 순위 산정 알고리즘 고도화.

## 4. 글로벌 확장

- 다국어 지원 및 글로벌 검색 엔진 통합.
- 지역별 맞춤형 검색 엔진과 협력.

