

RAG를 활용한 사용자
맞춤 AI 뉴스레터

목차

-
1. 프로젝트 소개 목표 설정
 2. 워크플로우
 3. 실험설계
 4. 데이터 크롤링
 5. RAG
 6. 모델 성능 측정
 7. 프롬프트 엔지니어링
 - LLM을 이용한 키워드 추출
 - 요약문 출력을 위한 프롬프트
 8. 마무리

프로젝트 소개

문제 정의

- 기존 검색 엔진을 활용하여 특정 도메인(AI)에 대해 키워드 나열이나 질문형 쿼리로 검색을 시도했을 때, 관련된 정확한 답변이나 유용한 정보를 얻는 데 어려움
- 검색 결과는 사용자 의도를 충분히 반영하지 못하거나, 필요한 정보를 정확히 제공하지 못하는 경우가 빈번하게 발생 이는 사용자가 원하는 정보를 얻기 위해 추가적인 탐색 시간을 필요로 함.
- 이러한 문제를 해결해보고자 **사용자 맞춤형 RAG(Retrieval-Augmented Generation) 기반 뉴스레터 시스템**을 설계 및 구현하고자 함

프로젝트 목표

대주제

- 한국에서는 네이버 사용자들이 아직도 빈번한만큼 네이버 자체 검색엔진을 RAG를 통해 업그레이드

소주제

- 라그의 활용 방법과 전체 work flow, 발전 가능한 방향들을 한 번 경험해보자



뉴스

블로그

카페

이미지

지식iN

인플루언서

동영상



입력하신 검색어가 길어 '최근 ai~게 되고,' 까지만 검색된 결과입니다. [도움말 보기](#)

옵션 ▾

• **관련도순** • 최신순 | 모바일 메인 언론사 ☐

'최근 ai 업계에서 비디오 생성 모델이...'에 대한 검색결과가 없습니다.

- 단어의 철자가 정확한지 확인해 보세요.
- 한글을 영어로 혹은 영어를 한글로 입력했는지 확인해 보세요.
- 검색어의 단어 수를 줄이거나, 보다 일반적인 검색어로 다시 검색해 보세요.
- 두 단어 이상의 검색어인 경우, 띄어쓰기를 확인해 보세요. [네이버 맞춤법 검사기](#)
- 검색 옵션을 변경해서 다시 검색해 보세요.



AI로 사진 속 물체 지우기

AI 지우개 CLOVA X



매일 달라지는 타임 특가

설맞이 첫 세일! 10% 쿠폰 어택 받기 >



네이버에서 넷플릭스를

지금 멤버십 시작하기



아무리 추워도 여행은 못 참지

겨울에 가기 좋은 국내 여행지

Query : 최근 ai 업계에서 비디오 생성 모델이 뭐가 있을까? 그리고 그 모델들의 성능 비교는 어떻게 되고, 앞으로의 전망은 어때?

워크 플로우

사용자 쿼리 검색



검색 API 호출



임베딩 모델 생성



벡터 스토어 검색



결과 평가



쿼리에서 키워드
추출



텍스트 분할



벡터 스토어 생성



응답 생성

RAG 워크플로우

Brainstorming - 실험설계

확장성 : 더욱 넓은 범위의 도메인은
RAG가 검색(Retrieval)할 수 있는 DB의
크기 영역

신뢰도 : 검색 자체의 신뢰도는 DB를
구축할 내용물의 영역

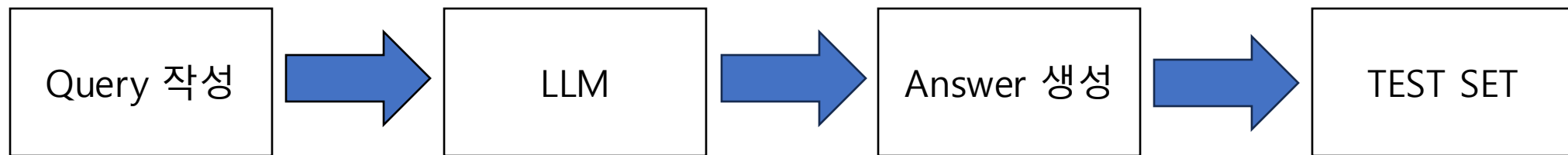
정확도 : 사용자 query에 대한 RAG의
출력물 평가는 사용자의 영역

사용자의 측면에서 정확도를 측정할 때, 자원의 한계를 극복할 정성평가의 대안의 필요성

실험설계 - 어떻게 평가할 것인가

Query 역생성을 통한 TEST SET 구축

기존



커스텀한 RAG의 벡터 데이터베이스와 LLM이 활용하는 데이터베이스 간의 극심한 격차

-> 벡터 데이터베이스 내부의 Retrieval 능력을 측정하는데 어려움

IBM "양자컴퓨터 상용화 핵심은 오류 수정...2029년 완성 목표"

입력 2025.01.22. 오후 10:56 



 1

 4









| 현재 양자컴퓨터-슈퍼컴퓨터 융합 방식 활발... "3년 내 '양자 우월성' 발생할 것"

"양자컴퓨터가 상용화하려면 데이터 계산을 오류 없이 빠르게 처리해야 합니다. 이에 기업이 양자컴퓨터 오류 현상 방지를 위해 연구하고 있습니다. IBM은 2029년 오류 수정(error correction) 기능을 완벽히 갖춘 양자컴퓨터 개발을 목표로 뒀습니다."

한국IBM 표창희 아시아-태평양 지역 퀀텀 엔터프라이즈 영업 총괄상무는 최근 여의도 한국IBM 사옥에서 진행한 기술 세션에서 양자컴퓨터 개발 현황과 자사 목표를 이같이 밝혔다. 빠른 연산 처리를 오류 없이 할 수 있는 오류 수정 기술이 절실하다는 설명이다.

양자컴퓨터는 양자물리학 기반으로 대규모 연산 처리를 할 수 있는 새로운 형태의 컴퓨터다. 일반 슈퍼컴퓨터는 연산 처리를 '비트' 단위로 계산한다. 0 또는 1로만 사용해 출력값을 내놓는 이 진법 형태다. 양자컴퓨터는 '큐비트' 단위로 처리한다. 큐비트는 0과 1 사이 무한한 값을 동시에 처리할 수 있다. 이를 '양자얽힘' 또는 '중첩'이라 부른다.

Query 생성

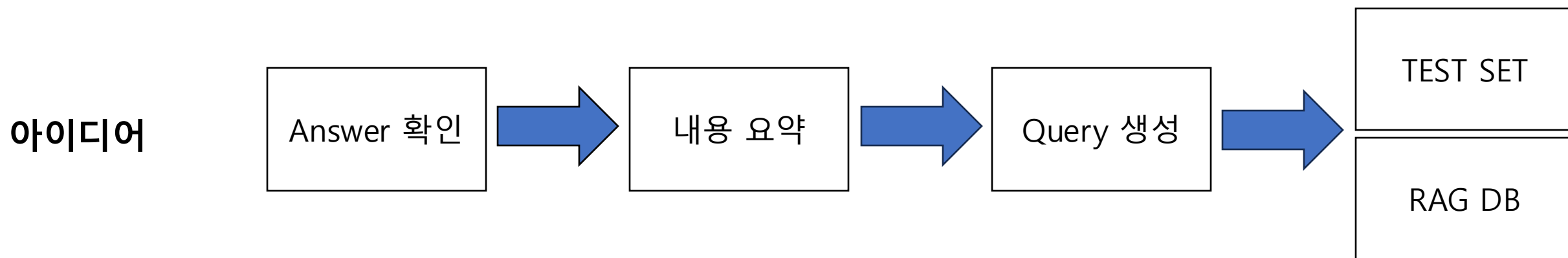
양자컴퓨터가 상용화되기 위해서는 어떤 기술들이 필요해? 장기적인 전망이나 실현 가능성은?

양자컴퓨터가 기존의 컴퓨터보다 훨씬 빠른 속도로 데이터를 처리할 수 있는 것은 이해가 되는데, 그래서 왜 오류 수정 기술이 왜 상용화의 핵심이라는거야?

그래서, 양자컴퓨터의 상용화를 위해서, IBM이나 다른 기업들이 이 오류 수정 기술을 어떻게 연구하고 있어?

실험설계 - 어떻게 평가할 것인가

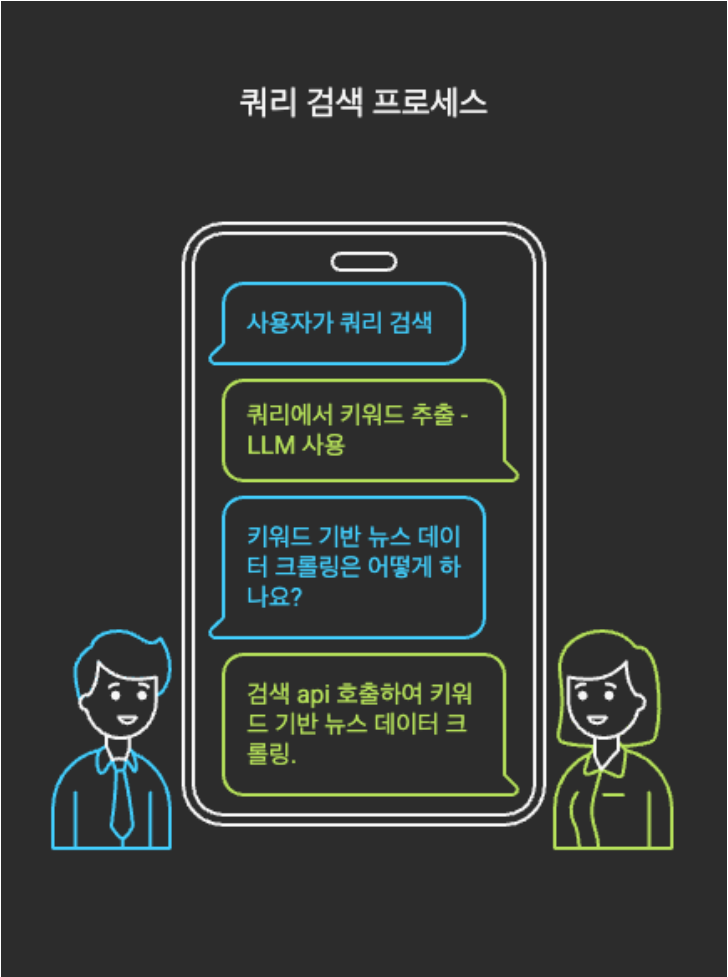
Query 역생성을 통한 TEST SET 구축



정량적인 지표를 통해 정성적인 평가를 간접적으로 파악 - 주관적 요소 감소
사용자 query 대해 RAG 알고리즘이 유의미한 정보를 제공할 수 있는지 평가하는데 유효
기계적인 평가가 가능

-> 정성 평가를 전부 대체할 순 없다는 한계점 존재
쿼리 생성 자체의 주관성을 간과할 수 없음

데이터 크롤링



▼ RAG 벡터 DB 구축

AI 딥러닝 머신러닝.txt 2202.8KB

1. AI 산업 기술 - 500

글로벌 AI 전망 추세.txt 1977.0KB

2. AI 딥러닝 머신러닝 - 500

AI의 사회적 영향.txt 1760.9KB

3. 자율주행 AI - 500

AI 챗봇

Title	URL	Content
제목 1	https://n.news...	본문 내용 1
제목 2	https://n.news...	본문 내용 2

4. AI 추천 - 300

AI 창작

AI 자율주행.txt 2109.2KB

8. AI 보안 - 300

AI 의료.txt 1252.2KB

9. AI 의료 - 300

AI 윤리.txt 1283.9KB

10. AI 창작 - 300

11. AI 로봇 공학 - 300

총 1,012개 데이터 확보

	Title	URL	Content
0	한발 늦게 'AI 원주민' 되겠다는 카카오...성능 발표는 '소심하게'	https://n.news.naver.com/mnews/article/138/000...	[디지털데일리 이진한 기자] 지난 22일 자체 개발한 통합 인공지능(AI) 플랫폼...
1	엔비디아가 포토샵을 '생성 AI 혁명'으로 이끈다	https://n.news.naver.com/mnews/article/262/000...	[박원익의 유익한 IT] "인공지능 '아이폰 모멘트' 시작됐다" ● 거대언어모델이...
2	오픈AI, 생산성 관리 '프로젝트' 기능 공개...향후 '실마스'서 AI 에이전트 등장 관심	https://n.news.naver.com/mnews/article/092/000...	생산성-자율성 증가 기대감...업계 "남은 5일간 차세대 모델-오피레이터 선보일 듯" ...
3	구글-삼성전자 'AI 동맹군' XR 전쟁서 위력 발휘할 것	https://n.news.naver.com/mnews/article/009/000...	산재이 굶타 구글 아시아-태평양 총괄사장 올해 'AI 에이전트' 경쟁 치열 내년 빅...
4	국방 AI 혁신...합성데이터에 주목 [각스]	https://n.news.naver.com/mnews/article/015/000...	KIDA 국방데이터연구단, 국방 데이터 활용 논의 쟁쟁에이아이, 합성데이터로 전시 ...

AI 로봇 공학.txt 1754.9KB

15. 글로벌 AI 전망 추세 - 300

AI 교육 기술.txt 848.4KB

실제 데이터 예시

AI IoT 사물인터넷.txt 1145.8KB

기본 모델 구축

RAG_DB

임베딩 모델



Adapter Models for

intfloat/multilingual-e5-large-instruct

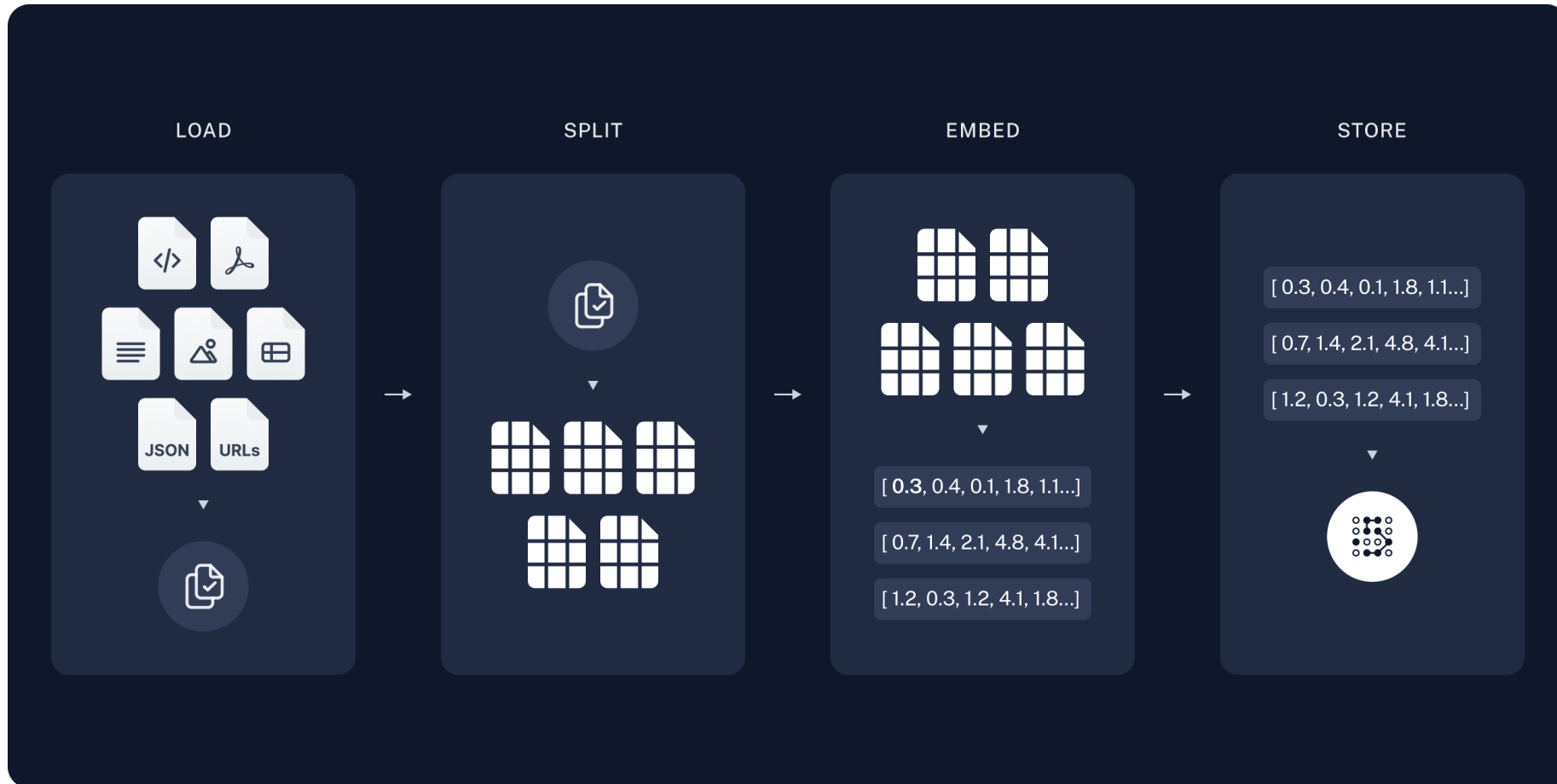
huggingface.co/models

생성 모델

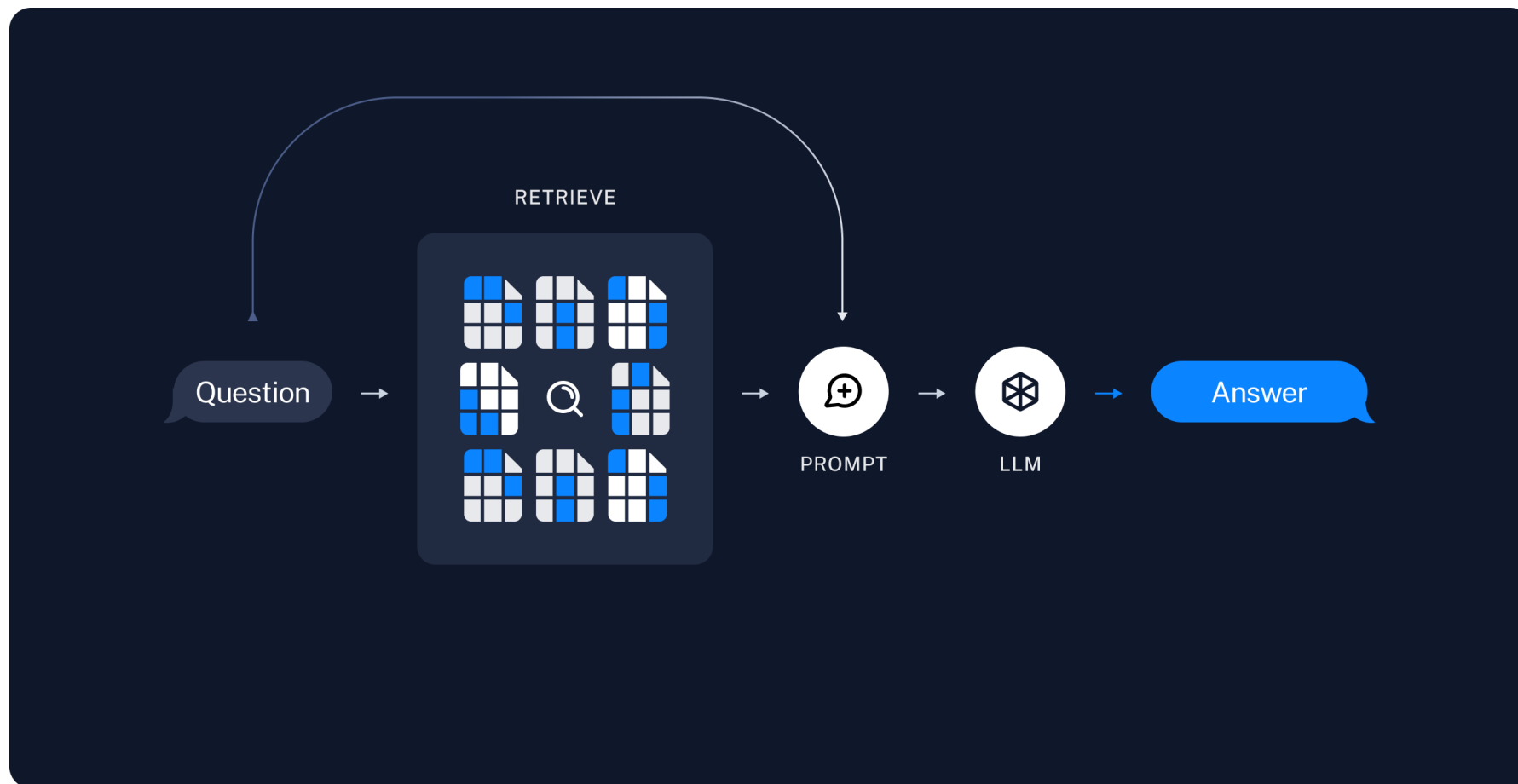


Gemini

RAG (Retrieval-Augmented Generation)



RAG (Retrieval-Augmented Generation)



모델 성능 측정에 사용한 평가지표 - 요약

Metrics	Score
BERTScore Recall	0.6523
BERTScore Precision	0.7093
BERTScore F1	0.6792
BLEU Score	0.0142
METEOR Score	0.0940
ROUGE-1 Recall	0.0816
ROUGE-1 Precision	0.3562
ROUGE-1 F1	0.1270
ROUGE-2 Recall	0.0306
ROUGE-2 Precision	0.1377
ROUGE-2 F1	0.0485
ROUGE-L Recall	0.0792
ROUGE-L Precision	0.3473
ROUGE-L F1	0.1234

```
108 df = pd.read_csv("/content/drive/MyDrive/Mini_AiFFELTHON/data/test.csv")
109
110 # test 셋의 'query'를 네이버 검색 엔진에서 실행하고, 'search' 변수에 담은 뒤, RAG 생성물과 비교
111 search_results = []
112 for query in df['query']:
113     news_data = fetch_naver_news(query)
114     search_results.append(news_data)
115
116 df['search'] = search_results
117
118 df.head()
```

... API 호출 횟수: 1

네이버 뉴스 링크 발견: <https://n.news.naver.com/mnews/article/056/0011875605?sid=101>

타이틀: [성공예감] CES 2025, 비약한 중국과 심심한 한국 - 송이라 기자 (서울경제신문)

기사 본문 (네이버 뉴스 경로):

===== 인터뷰 자료의 저작권은 KBS 라디오에 있습니다. 전문 게재나 인터뷰 인용 보도 시, 아래와 같이 채널명과 정확한 프로그램명을 밝혀주시기를 바랍니다. =====

API 호출 횟수: 2
API 호출 횟수: 3
API 호출 횟수: 4
API 호출 횟수: 5
API 호출 횟수: 6
API 호출 횟수: 7
API 호출 횟수: 8
API 호출 횟수: 9
API 호출 횟수: 10
API 호출 횟수: 11
API 호출 횟수: 12
API 호출 횟수: 13
API 호출 횟수: 14
API 호출 횟수: 15
API 호출 횟수: 16
API 호출 횟수: 17
API 호출 횟수: 18
API 호출 횟수: 19
API 호출 횟수: 20
API 호출 횟수: 21
API 호출 횟수: 22

모델 성능 측정에 사용한 평가지표 - RAGAS

평가항목	이름	설명
생성 (Generation)	Faithfulness	생성된 응답이 검색된 컨텍스트에 기반했는지를 평가
생성 (Generation)	Answer Relevancy	생성된 응답이 사용자 질문(프롬프트)과 얼마나 관련이 있는지를 평가
검색 (Retrieval)	Context Recall	검색된 문서 중 얼마나 많은 문서가 실제로 관련 있는지를 평가 (검색된 정보의 정확도)
검색 (Retrieval)	Context Precision	검색된 문서 중 실제로 관련 문서가 얼마나 포함되었는지를 평가 (필요한 정보를 얼마나 잘 찾는지)

모델 성능 측정에 사용한 평가지표 - RAGAS

$$\text{Faithfulness score} = \frac{|\text{Number of claims in the generated answer that can be inferred from given context}|}{|\text{Total number of claims in the generated answer}|}$$

$$\text{answer relevancy} = \frac{1}{N} \sum_{i=1}^N \cos(E_{g_i}, E_o)$$

$$\text{answer relevancy} = \frac{1}{N} \sum_{i=1}^N \frac{E_{g_i} \cdot E_o}{\|E_{g_i}\| \|E_o\|}$$

$$\text{context recall} = \frac{|\text{GT sentences that can be attributed to context}|}{|\text{Number of sentences in GT}|}$$

$$\text{Context Precision@K} = \frac{\sum_{k=1}^K (\text{Precision@k} \times v_k)}{\text{Total number of relevant items in the top } K \text{ results}}$$

$$\text{Precision@k} = \frac{\text{true positives@k}}{(\text{true positives@k} + \text{false positives@k})}$$

모델 성능 측정에 사용한 평가지표 - RAGAS

평가항목	이름	TEST SET 평균 수치
생성 (Generation)	Faithfulness	0.914094
생성 (Generation)	Answer Relevancy	0.417772
검색 (Retrieval)	Context Recall	0.894703
검색 (Retrieval)	Context Precision	0.960784

Low Faithfulness

query		generated_answer	answer	faithfulness
24	AI 기술이 다양한 산업에서 어떻게 활용되고 있으며, AI 인프라의 주요 요구 사항과 핵심 요소는 무엇인가요?	제공된 텍스트만으로는 AI 기술의 다양한 산업 활용 사례와 AI 인프라의 주요 요구 사항 및 핵심 요소를 명확하게 알 수 없습니다. 텍스트는 게임, 노후 인프라 안전성 향상, 그리고 투자 성과와 관련된 AI 활용에 대한 언급만 하고 있을 뿐, 폭넓은 산업적 활용이나 인프라 요구사항에 대한 정보는 포함하고 있지 않습니다.	[기고] 산업 경계를 넘어서는 AI 혁신, 최적의 인프라에서 시작된다\nhttps://n.news.naver.com/mnews/article/092/0002359474?sid=105\n비즈니스 환경이 급속도로 진화하면서 기업들은 끊임없이 생산성 향상과 데이터 활용, 사용자 경험 개선을 추구하고 있다. 이러한 변화 속에서 인공지능(AI)은 기업들의 핵심 과제를 해결할 강력한 도구로 주목받고, 많은 기업은 AI 도입을 통해 경쟁력을 강화하려 노력하고 있다. 산업별 AI 활용 현황과 전망 HS 효성인포메이션시스템 정문준 매니저 다양한 산업 분야에서 AI는 혁신적인 변화를 이끌고 있다. 유통 업계에서는 AI 기반 분석을 통해 재고 관리를 최적화하고 개인화된 고객 서비스를 제공하면서 매출 증대와 고객 만족도 향상을 동시에 달성하고 있다. AI는 고객의 구매 패턴을 분석해 최적의 재고 수준을 유지하고 각 고객에게 맞춤형 상품을 추천함으로써 구매 전환율을 높이는 데 기여한다. 의료 분야에서는 예측 분석과 머신러닝을 활용해 진단 정확도를 높이고 환자 치료 결과를 개선하고 있다. AI 알고리즘은 의료 영상을 분석해 초기 단계의 질병을 발견하고 환자의 의료 기록을 바탕으로 최적의 치료 방법을 제시한다. 이를 통해 의료진은 더 정확하고 효율적인 의사결정을 내릴 수 있다. 금융권에서는 대규모 언어 모델(LLM)을 도입해 실시간 금융 사기 탐지와 자동화된 거래 전략을 구현하고 있다. AI는 수많은 거래 데이터를 실시간으로 분석해 이상 거래를 감지하고 시장 동향을 예측해 최적의 투자 전략 수립을 지원한다. 제조업에서는 AI 기반 예측 유지보수로 설비 다운타임을 최소화하고 생산 효율을 높이고 있다. 센서에서 수집된 데이터를 AI가 실시간으로 분석해 설비 고장을 사전에 예측하고, 최적의 유지보수 시점을 제시함으로써 생산 중단을 방지한다. 이는 생산성 향상과 비용 절감으로 이어진다. AI 워크로드의 특성과 인프라 요구사항 AI 워크로드는 모델 학습과 실행, 유지관리에 있어 특별한 요구사항을 가진다. 예측 분석, 자연어 처리, 이상 징후 감지, 이미지 인식, 추천 시스템 등 각각의 워크로드는 고성능 컴퓨팅 자원과 대용량 스토리지가 필요하다. 특히 딥러닝 모델은 수십에서 수천억 개의 파라미터를 처리해야 해 강력한 GPU 성능이 요구된다. AI 인프라의 핵심 요소는 강력한 컴퓨팅 성능이다. GPU와 같은 가속기는 복잡한 연산을 병렬로 처리해 AI 모델의 학습 및 추론 속도를 크게 향상시킨다. 또한 지속적으로 증가하는 데이터 규모에 맞춰 스토리지 시스템은 높은 확장성과 처리 성능을 갖춰야 한다. 네트워크 인프라도 중요한 요소다. AI 워크로드는 대량의 데이터를 빠르게 이동시켜야 해 고속 네트워크가 필수다. 실시간 처리가 필요한 애플리케이션의 경우 낮은 지연시간이 매우 중요하다. 보안 역시 간과할 수 없다. AI 시스템은 데이터 보호와 규정 준수를 위한 강력한 보안 기능을 제공해야 한다. AI 도입의 현주소와 미래 전망 2023년 맥킨지 설문조사에 따르면 66%의 조직이 AI를 도입하는 탐색 또는 파일럿 단계에 있는 것으로 나타났다. 이는 AI 도입에 여전히 많은 과제가 있음을 시사한다. 하지만 AI 시장은 2027년까지 1천510억 달러 규모로 성장이 예측되며 기업들의 AI 투자도 꾸준히 증가할 것이다. AI 도입의 성공을 위해서는 명확한 전략과 로드맵이 필요하다. 기업은 자사의 비즈니스 목표와 환경에 맞는 AI 활용 사례를 발굴하고 이를 단계적으로 구현해 나가야 한다. AI 시스템의 구축과 운영을 위한 전문 인력 확보와 교육도 중요하다. HS 효성 AI 플랫폼(이미지= HS 효성인포메이션시스템) HS 효성인포메이션시스템은 고성능 컴퓨팅과 AI 워크로드를 위한 통합 AI 플랫폼을 제시하고 있다. 검증된 레퍼런스 아키텍처를 기반으로 데이터센터 AI 인프라의 설계, 배포, 관리를 획기적으로 단순화하고, AI 소프트웨어부터 GPU 서버, 네트워크, 스토리지에 이르는 모든 구성 요소를 단일 벤더를 통해 제공한다. 현재 가장 많이 도입되고 있는 H100 GPU의 경우 최대 128개 노드까지 확장 가능한 최고 성능의 AI 개발 환경을 지원한다. NVMe 플래시 기반의 고성능 스토리지 시스템은 대규모 AI 워크로드를 효율적으로 처리한다. AI는	0.250000
	AI와 머신러닝을 통해 기업 비즈니스, 금융 및 제조업에서 실질적인 가치를 제공하는 방법에 대한 구체적인 내용이 없습니다. 따라서 질문에 대한 답변을 드릴 수 없습니다.	제공된 텍스트에는 AI와 머신러닝이 기업 비즈니스, 금융 및 제조업에서 실질적인 가치를 제공하는 방법에 대한 구체적인 내용이 없습니다. 따라서 질문에 대한 답변을 드릴 수 없습니다.	은 AI·머신러닝 리더십 등의 성과를 냈다. 회사 측은 코오롱비티와 종판 계약 후 협력 파트너사가 19개로 늘어났고 기술 지원부터 사업 모델 공동 발굴까지 협업 범위를 확대 중이라고 강조했다. 특히 종판 체제 구축 효과에 대해 이 대표는 "글로벌 정책에 따라 지난해 6월부터 한국을 포함한 여러 국가에서 종판 체제로 전환 됐다"며 "컨설턴트·마케팅·기술지원·커스터머 성공' 담당 인력 등을 전담 조직으로 두고 올해 본격적인 시너지를 낼 준비가 됐다"고 밝혔다. 새해부터는 비즈니스 전략이 AI를 중심으로 방향성을 잡을 예정이다. 회사는 우선 데이터·AI 시대에 신뢰할 수 있는 파트너로 자리매김하는 동시에 기존 투자 보호와 기업 경쟁력 강화를 꾀할 계획으로, 이에 더해 데이터·AI 리더십을 통해 사회 공헌을 확대한다는 구상이다. 이 대표는 이를 뒷받침할 핵심 솔루션으로 '바이야(Viya)' 플랫폼을 소개했다. '바이야'는 AI 모델 개발부터 운영까지 전 과정을 통합 관리하는 SAS의 클라우드 네이티브 플랫폼이다. 파이썬·R 등 오픈소스와 연동돼 데이터 사이언티스트가 데이터 전처리에 쏟아야 했던 시간을 기존 70~80%에서 20~30%로 줄이는 것이 특징이다. 이 대표는 "이 솔루션은 데이터 사이언티스트가 전체 업무 시간의 상당수를 데이터 전처리에 쏟아부어야 하던 과정을 획기적으로 단축해 분석·예측 모델링에 집중할 수 있게 만든다"며 "이러한 획기적인 생산성을 기존 투자 자산을 보호해야 하는 국내 금융·공공·제조 업체에도 적용하겠다"고 밝혔다. SAS 코리아 로고 (사진=지디넷코리아 DB) '바이야'가 국내에서 가시적 성과가 있었는지를 묻는 기자의 질문에 그는 "지난 7월 공식 출시한 뒤 금융권의 클라우드 활용 가속화 추세와 맞물려 '클라우드 네이티브' 환경에 대한 문의가 부쩍 늘었다"며 "방문리뷰 제 완료로 금융권 클라우드 도입이 확산되고 공공에서도 프라이빗 클라우드를 구축하는 사례가 많아 '바이야'가 최적의 선택지가 되고 있다"고 설명했다. 이어 발표에 나선 김근태 CA 본부 상무는 시장 현황과 '바이야'를 통한 고객 사례를 구체적으로 공유했다. 그는 AI 솔루션이 단순한 챗봇에서 나아가 기업 비즈니스에 실질적인 가치를 제공하도록 발전해야 한다며 '바이야'는 거대언어모델(LLM)이 부족한 예측·분석 능력을 보완해 보다 정확한 의사결정 지원을 해줄 수 있다고 설명했다. 이는 SAS가 보유한 고급 분석 기술과 생성형 AI의 결합 덕분이다. 김 상무에 따르면 생성형 AI는 방대한 데이터를 학습해 텍스트 생성과 자연어 처리 능력에서는 탁월하지만 특정 비즈니스 상황에서 발생하는 데이터의 정밀한 해석과 예측에서는 한계가 있다. 이를 보완하기 위해 '바이야'는 LLM이 다루기 어려운 정형 데이터를 분석하고 이를 기반으로 정확한 예측과 인사이트를 제공해 시너지를 낸다. 김 상무는 "데이터 사이언티스트 생산성이 3.5배, 머신러닝 운영(ML Ops) 담당자 생산성이 4.5배 상승했다는 연구 결과가 나왔다"며 "대기업 제조 현장에서 불량 원인 분석 시간을 수개월에서 '원클릭' 수준으로 단축했고 금융권에선 대출 GPT 관련 리와 사기 방지 등에 AI 적용 속도를 높이고 있다"고 전했다. 이어 "빅데이터가 쌓일수록 예측 정확도와 분석 효율이 더 높아지는 구조"라며 "단순히 '챗 GPT' 같은 LLM을 그대로 연결하는 게 아니라 각 기업 환경에 맞춰 필요한 데이터와 파이프라인에 적용함으로써 오답을 줄이고 업무에 실질적으로 활용될 수 있게 한다"고 설명했다.	0.500000

High Faithfulness

	query	generated_answer	answer	faithfulness
0	LG 전자의 AI 기반 스마트홈 기술이 CES 2025에서 어떻게 소개되었는지 알려주세요.	LG전자는 CES 2025에서 '공감지능과 함께 하는 일상의 라이프스 굿'이라는 메시지 아래 AI 기반 스마트홈 기술을 선보였습니다. 주요 내용은 다음과 같습니다.\n\n* AI 홈 존: 생성형 AI를 탑재한 'LG 씽큐 온'과 온디바이스 AI 기반의 콘셉트 제품 등 다양한 AI 홈 허브를 통해 집안 곳곳의 센서로 고객의 말, 행동, 주변 환경을 감지하고, 연결된 IoT 기기를 최적의 상태로 제어하는 모습을 시연했습니다. 예를 들어, 수면 중 고객의 심박수, 호흡, 기침 등을 분석하여 온수를 제안하거나 실내 온도와 습도를 자동 조절하는 기능을 보여주었습니다. 136형 차세대 LG 마이크로 LED는 고객의 목소리를 식별하여 맞춤형 webOS 콘텐츠를 제공하는 기능도 소개되었습니다. 수집된 데이터는 LG 전자의 독자 보안 시스템 'LG 쉴드'로 보호된다고 강조했습니다.\n\n* LG 시그니처	[CES 2025] 노트북 컷더니 AI가 알아서 '척척'... LG 제안 '공감지능'의 미래\n\nhttps://n.news.naver.com/mnews/article/417/0001050110?sid=101\n\nLG 전자, AI 기술 접목 소개인화 AI 홈 라이프스타일 제안 LG 전자가 CES 2025 전시장에 77형 LG 시그니처 올레드 T 28대와 상들리에 조명을 활용해 설치한 조형물. / 사진=이한듬 기자 #. 사용자가 노트북을 켜자 집안 곳곳에 설치된 인공지능(AI) 홈 허브가 해당 동작을 '업무를 준비하는 상황'으로 인식하고 이에 적합한 집안 환경 조성을 제안한다. 업무 도중 실수로 전원이 꺼져도 작업물이 손실될 걱정이 없다. 노트북에 탑재된 AI가 지속적으로 화면을 캡처하고 작업 내용을 저장해주기 때문이다. LG 전자가 7일(현지시간)부터 나흘간 미국 라스베이거스에서 열리는 세계 최대 가전·IT 전시회 'CES 2025'에서 구현한 AI 홈의 미래다. LG 전자는 이번 전시회에서 2044㎡ 규모의 전시관을 꾸리고 '공감지능과 함께하는 일상의 라이프스 굿'을 주제로 고객의 AI 경험을 다양한 공간으로 연결·확장하며 변화하는 일상을 제안했다. 관람객은 아침부터 저녁까지, 고객의 모든 시간과 일상 속 공간에 녹아든 공감지능을 자연스럽게 경험할 수 있다. 이날 본격적인 개막에 앞서 공개된 사전 부스투어에서도 집, 사무실, 취미공간, 차량 등 다양한 공간 속에 LG가 구현한 AI 라이프스타일을 체험할 수 있었다. 생성형 AI를 탑재한 'LG 씽큐 온', 온디바이스 AI 기반의 콘셉트 제품 등 다양한 AI 홈 허브는 집안 곳곳에 설치된 센서로 고객의 말과 행동, 주변 환경을 감지하고 연결된 사물인터넷(IoT) 기기를 최적의 상태로 제어한다. 예를 들어 거실에 설치된 136형 차세대 'LG 마이크로 LED'가 고객의 목소리를 식별해 노인인지, 어린이인지지를 구분하고 해당 고객에게 맞춘 웹 OS 콘텐츠를 제공하는 식이다. LG 전자가 CES 2025에서 선보인 77형 LG 시그니처 올레드 T / 사진=이한듬 기자 개인정보가 유출될 걱정보도 덜 수 있다. 고객의 동의를 받아 수집된 데이터는 LG 전자의 독자 보안 시스템 'LG 쉴드'로 안전하게 보호되기 때문이다. 초프리미엄 가전 'LG 시그니처' 신제품도 AI를 접목해 더욱 똑똑한 제품으로 거듭났다. '스마트 인스타뷰 냉장고'는 투명 올레드 디스플레이를 적용해 문을 열지 않고도 보관 중인 식품의 종류와 양을 확인할 수 있다. 내부 카메라가 식품을 자동으로 인식해 사용자는 LG 씽큐 앱으로 보관 중인 식품의 위치와 목록도 살필 수 있다. 지난달 미국 시장에 출시된 LG 시그니처 OLED T도 눈길을 끌었다. 4K 해상도와 투명 스크린, 무선 AV 송·수신 기술을 접목한 77형 TV로 현지 출시가격은 6만달러의 초고가 제품임에도 이미 판매가 이뤄질 정도로 반응이 좋다는 게 LG 전자의 설명이다. 이 제품은 국내에도 출시될 예정이다. 구체적인 일정은 아직 정해지지 않았다. LG 전자는 전시장 한복판에 77형 LG 시그니처 올레드 T 28대와 대형 상들리에 조명을 활용한 초대형 미	1.0
1	우리나라 국방부나 군사강국에서 AI와 IoT를 적용한 방산 기술이 나타트렌드에 관한 기사 있나요?	랜드에 대한 내용이 포함되어 있습니다.\n\n주요 내용은 다음과 같습니다.\n\n* 합성 데이터 활용: 한국 국방 데이터 연구단은 국방 분야 데이터 확보의 어려움을 해결하기 위해 합성 데이터를 활용하는 방안을 논의했습니다. 젠젠에이아이와 인피닉 등의 기업이 이 분야에서 활동하고 있습니다. 특히, 젠젠에이아이는 자율주행 분야에서 축적한 합성 데이터 생성 기술을 국방 분야에도 적용하고 있습니다.\n\n* AI 기반 무기 개발: 글로벌 방산 기업들은 AI 기반 무기, 특히 드론 개발에 적극적으로 투자하고 있으며, 인수합병을 통해 기술 확보에 나서고 있습니다. 미국 국방부도 AI 기술을 국방력 경쟁 우위의 핵심 요소로 보고, 직접 대출 및 민간 펀드 지원 등을 통해 AI 기술 기업에 대한 투자를 확대하고 있습니다. 특히 하드마틴, 노스롭그리먼, 팔란티어 등의 기업이 이러한 움직임을 주도하고 있습니다.\n\n* 한국 기업의 현황: 한국 기업들 중 솔빗시스템과 젠젠에이아이가 AI 기반 방산 기술 분야에서 주목받고 있습니다. 솔빗시스템의 위치 기반 인명 구조 시스템은 CES 혁신상을 수상하기도 했습니다. 하지만, 한국 기업들의 인수합병은 글로벌 기업들에 비해 소극적인 편입니다.\n\n* 미국 국방부의 전략: 미국 국방부는 AI, IoT/센싱, 로봇공학 기술에 대한 수요를 공개적으로 드러내고 있으며, CES 2025에서 관련 기술을 적극적으로 소개할 예정입니다.\n\n* 요약하자면, 제공된 기사는 한국과 미국의 국방 분야에서 AI와 IoT 기술도	"중요한 건 AI야"... 미군과 글로벌 방산업체가 꼽았다\n\nhttps://n.news.naver.com/mnews/article/417/0001050388?sid=101\n\n첨단 AI 국방기술 육성 나선 미 국방부, 직접 대출부터 민간펀드까지 금융지원 나서 록히드마틴 등 정통 기업들도 M&A에 적극 나서... CES에서 선보이는 AI 기술 주목 방산업체에 부는 디지털 트랜스포메이션(DX) 바람에 미 국방부를 선두로 정통 방산기업들도 기술 스타트업 투자와 인수합병에 나설 것으로 보인다. / 사진=쉴드 AI 글로벌 방산업체가 CES 2025에서 선보이는 인공지능(AI) 기술에 주목하고 있다. 방산업체에 부는 디지털 트랜스포메이션(DX) 바람에 미 국방부를 선두로 정통 방산기업들도 AI 기술 스타트업 투자와 인수합병에 나설 것으로 보인다. CES에 참가한 국내 스타트업의 AI 기술도 주목받으며 글로벌 최대 방산시장인 미국 진출에 대한 기대감이 오르고 있다. 8일 미국 국방부의 전략적 자본 사무국(OSC)에 따르면, 미 국방부를 선두로 미국 방산업체가 AI, IoT(사물인터넷)·센싱 기술 기업에 대한 투자와 인수합병에 나설 것으로 파악된다. OSC의 2025 회계연도 투자전략 문서는 국방 기술에서 전략적 경쟁국 대비 경쟁 우위를 제공할 수 있는 신기술 분야 12개에 대한 투자 프레임 워크를 제시했다. 미군은 AI 기술이 국방력의 경쟁 우위를 좌우하는 키체인자가 될 것으로 본다. 우크라이나 전쟁을 통해 드론과 자율주행 기술의 위력을 경험한 만큼 빠른 기술확보에 열을 올리는 모습이다. 지난해 10월 OSC는 31곳의 주요 방위기술 생산 장비를 제조하는 기업들에 대해 10억달러(1조4519억원) 규모의 직접 대출 지원을 발표했다. 소기업투자회사(SBICCT) 주요 기술 이니셔티브에 참여할 13개의 민간 펀드도 승인했다. OSC는 2022년 로이드 오스틴 미 국방장관이 민간 자본의 활용을 위해 설치한 사무국이다. 미 국방부가 본격적으로 소규모 기업 투자를 통해 신기술을 키우겠다는 방침을 발표함에 따라 주요 방산업체들도 투자에 나설 것으로 보인다. 전문가들은 첨단기술들은 급속도로 변화하는 산업인만큼 보수적인 방산기업들도 자체 개발보단 인수합병(M&A)을 택하고 있다고 말한다. AI 플랫폼 기업 팔란티어는 스타트업과 협업해 차세대 국방용 AI 사업을 확대하고 있다. 최근에는 미 국방부와 협업해 자율 드론 및 기타 군사 기술 개발용 AI인 쉴드 AI(Shield AI)를 내놓기도 했다. 전직 미 하원의원이자 팔란티어 테크놀로지스의 국방 책임자인 마이클 갤러거는 현대전에 대비해 미 국방부가 보다 다양한 신생 기술 기업과의 협력을 확대해야 한다고 강조했다. 록히드마틴·노스롭그리먼 안드릴과 같은 '정통' 방산기업들도 전쟁특수로 불어날 인형연금을 활용해 투자와 M&A에 나서고 있다. AI, 드론 등 첨단 기술을 적용한 새로운 무기 체계 개발 생산 역량을 키우기 위해서다. 안드릴은 오픈 AI와 첨단 AI 모델을 활용한 항공 드론 방어 시스템을 개선하고 있다. 미 육군은 이번 CES 전시에서 'xTech' 상금 경진대회를 열고 AI, 사물인터넷(IoT)/센싱, 로봇공학 기술에 대한 수요를 직접적으로 드러냈다. 세계 최대 IT(정보기술)가전 전시회 'CES 2025'는 '다이브 인'을 주제로 오는 7일부터 10일까지 열린다. / 사진=뉴스1 미 육군은 이번 CES 전시에서 'xTech' 상금 경진대회를 열고 AI, 사물인터넷(IoT)/센싱, 로봇공학 기술에 대한 수요를 직접 드러냈다. 방산 관련 AI 기술 기업으로 국내 기업 솔빗시스템, GenGen AI(젠젠 AI)등도 주목 받고 있다. 솔빗시스템의 전파 음영지역 내에서도 작동하는 위치 기반 인명 구조 시스템 'SOLVIT-ISAR'는 'CES 혁신상'을 수상하며 주목받았다. SOLVIT-ISAR는 첨단 모바일 통신 포렌식 기술과 지능형 추론 기술을 활용해 기존에 3시간 이상 소요되던 탐색 과정을 10분내로 단축했다. 현대전의 주요 무기인 전자기밀(EMP)탄, 정전탄 등의 공격에도 신속하게 전력 상환을 파악하고 조치할 수 있는 방어기술인 셈이다. 네이버 D2 스타트업 펍토리가 투자한 젠젠 AI는 방위 및 보안 모니터링 시스템에 특화시킬 수 있는 합성 데이터 플랫폼을 선보였다. 합성 데이터는 기존 데이터를 합성에 새로운 데이터를 창출해내는 기술이다. 실제 데이터 수집없이 알고리즘·시뮬레이션을 통해 인공 데이터를 생산할 수 있어 유출에 민감한 방위 보안 시스템에 적합하다. 시간과 비용에 있어서도 경제적이고, 적의 기습 등 변수에 대한 대응력도 높고, 도심형 항공 모빌리티용 AI 플랫폼을 선보인 미국의 SES AI, 소리 AI로 분석 시스템을 전시한 Cochl(코칠)도 있다. 드론 분야 역시 AI가 접목된 기술들이 전시됐다. 미국의 WEFLO(위플로)는 미래 항공 모빌리티를 위한 검사 솔루션을, DroneUA(드론유에이)는 드론에 로봇 공학과 데이터 분석을 통합한 응용 보안 기술을 선보였다.	1.0

Low Context Precision

	query	generated_answer	answer	context_precision
22	여러 분야에서 AI 기술을 둘러싼 글로벌 패권 경쟁에서 주요 국가들의 전략은 어떻게 될까?	제공된 텍스트에서는 주요 국가들의 AI 기술 경쟁 전략에 대한 구체적인 내용을 찾을 수 없습니다. 한국 정부가 AI G3 달성을 목표로 기술산업화 생태계 구축에 나서겠다는 내용만 언급되어 있습니다. 다른 국가들의 전략에 대해서는 알 수 없습니다.	<p>‘20세기 핵경쟁과 닮은꼴’... ‘AI 쏠 냉전’ 시작됐다\</p> <p>https://n.news.naver.com/mnews/article/029/0002928731?sid=105</p> <p>美, ‘우방국·적대국·기타’ 3등급 분류 일부 국가 제외한 AI 칩 수출통제 확대 AI 기술 독점 시도, 핵무기 경쟁과 비슷 국제사회 대응 따라 세계질서 바뀔수도 MS Bing 이미지크리에이터를 이용해 생성한 이미지. 인공지능(AI)이 21세기 글로벌 패권 경쟁의 중심에 서면서 ‘AI 냉전’이 격화되고 있다. AI는 단순한 기술 혁신을 넘어 국가 안보와 경제, 군사력의 핵심으로 자리 잡았으며, 이는 20세기 중반 핵무기를 둘러싼 국제적 경쟁과 유사한 양상을 보인다. 미국은 AI 기술을 ‘21 세기의 핵’으로 간주하며, 글로벌 패권을 유지하기 위해 전방위적인 전략을 전개하고 있다. AI 칩 수출 규제, 동맹국과의 협력 강화, 자체 기술 생태계 보호 등 다양한 조치가 이러한 전략의 일환으로 실행되고 있다. ◇“핵심 우방과만 AI 공조”... AI 칩 수출 막는 미국 미국 정부는 전 세계 국가를 우방국·적대국·기타 3개 등급으로 나눠 소수 우방국만 미국산 AI 반도체를 제한 없이 수입할 수 있게 하는 내용을 담은 AI 칩 수출 규제정책을 조만간 발표할 것으로 알려졌다. 이 경우 중국·러시아·북한·이란·베네수엘라·쿠바·벨라루스·이라크·시리아 등 미국의 적대국으로 분류된 20여 개국은 AI 칩을 수입할 수 없도록 이들 국가에 대한 수출 경로가 차단된다. 100여 개의 중간 등급 국가는 수입 가능한 총 연산력(컴퓨팅 파워)에 상한이 설정된다. 한국·일본·대만·영국·프랑스 등 소수의 주요 우방국은 지금처럼 미국산 반도체를 제한 없이 구매할 수 있을 것으로 보인다. 이번 조치 이전에도 미국은 중국으로의 첨단 반도체 칩 수출을 제한해 왔다. 2022년부터 주요 반도체 제조사들에게 중국에 대한 고성능 AI 칩 수출을 제한하기 시작했으며, 2024년에는 이를 더욱 강화하는 조치를 발표했다. 엔비디아, AMD와 같은 기업들은 중국으로의 수출이 제한되는 주요 대상에 포함됐고, 고성능 그래픽처리장치(GPU)와 관련 소프트웨어의 중국 수출은 사실상 전면 금지됐다. 바이든 정부가 조만간 발표할 것으로 알려진 AI 칩 수출 규제정책은 여기서 한발 더 나아가 우방국 중심으로 AI 반도체 공급망을 구축하려는 의도를 담은 것으로 보인다. MS Bing 이미지크리에이터를 이용해 생성한 이미지. ◇반발하는 빅테크들 “미국 경제에 해” 미 정부의 움직임에 대해 엔비디아·MS 등 빅테크들은 미국 정부의 AI 칩 수출 통제 확대 움직임에 공개적으로 반발하고 나섰다. 미국의 핵심 산업에 심각한 타격을 줄 수 있다는 것이다. 특히 AI 반도체 시장의 90%를 차지하고 있는 엔비디아는 강하게 목소리를 내고 있다. 네드 핀클 엔비디아 글로벌 업무 담당 부사장은 “바이든 행정부의 막판 정책은 미국 산업과 세계 사회에서 비판 받을 유산이 될 것”이라며 “우리는 바이든 대통령이 미국 경제에 해를 끼치고, 미국을 후퇴시키고, 적대자들의 손에 돌아날 뿐인 정책을 제정해 트럼프 대통령의 앞길을 막지 않길 바란다”고 말했다. 구글·MS·아마존 등을 대변하는 정보기술산업위원회(ITIC)도 목소리를 냈다. 켄 글릭 오라클 부회장은 “미국 기술 업계를 타격할 역대 가장 파괴적인 규제로 기록될 것”이라며 “업계와 협의 없이, 행정부가 바뀌기 불과 며칠 전 이런 규모의 규칙을 비밀리에 발행하는 것은 매우 심각한 일”이라고 비판했다. ◇안보도 경제도 AI로 통한다 미국 정부의 행보는 AI 기술이 단순히 민간 산업에서의 혁신을 넘어 군사적 및 감시 기술로 전환될 수 있는 잠재력을 염두에 둔 것이다. AI가 드론 제어, 자동화 무기 시스템, 사이버 전쟁 등의 군사적 활용도가 높아, 이를 제어하지 않을 경우 글로벌 안보 구조가 심각한 위협을 받을 수 있다고 보기 때문이다. AI 칩은 대규모 데이터 처리와 고속 연산을 가능케 하는 AI 모델의 핵심 자원이다. 중국의 AI 연구는 이 고성능 칩에 크게 의존하고 있으며, 이러한 칩들은 대부분 미국 기술에 기반을 두고 있다. 미국은 이러한 의존성을 전략적 카드로 활용해 중국 AI 생태계의 성장을 억제하려는 의도를 분명히 하고 있다. 월스트리트저널은 “바이든 행정부가 발표하려는 이번 정책의 목표는 중국이 동남아시아와 중동과 같은 지역의 데이터센터를 활용해 AI 칩에 액세스할 수 있는 허점을 막는 것”이라고 밝혔다. 반면 동맹국들에게는 제한 없이 칩을 수출할 수 있도록 허용하면서, 동맹국 간 AI 기술 협력을 강화하고 있다. 이는 미국 주도의 기술 생태계를 더욱 공고히 하고, 글로벌 기업들이 미국의 기준을 따르도록 유도하는 전략의 일환이다. ◇20세기의 핵 경쟁과 유사한 AI 냉전 20세기 중반, 미국과 소련은 핵무기 개발을 둘러싸고 치열한 경쟁을 벌였다. 미국은 1945년 첫 원자폭탄 개발에 성공하며 전후 국제 질서를 주도할 수 있었고, 이는 핵무기 기술의 독점이 가능했던 덕분이었다. 소련은 1949년 핵무기 실험에 성공하면서 미국과의 핵 경쟁을 본격화했고, 이로 인해 냉전이라는 장기적 대립 구도가 형성됐다. AI 기술 경쟁 또한 이와 비슷한 특징을 보인다. 21세기 들어 AI는 단순한 기술적 혁신을 넘어 국가의 전략적 자산으로 떠올랐다. 핵무기가 군사적 억지력을 상징했다면, AI는 정보와 자동화된 무기 시스템을 통해 현대 군사 전략의 중심으로 자리 잡았다. AI를 활용한 자율 무기 체계, 사이버 공격 기술, 감시 체계는 AI 기술에서의 격차가 곧 군사적 우위를 의미한다는 것을 보여준다. 이러한 상황에서 미국은 AI 기술력의 독점을 통해 글로벌 패권을 유지하려는 강한 의지를 드러내고 있다. ◇“미국 따라잡겠다”...중국의 AI ‘굴기’ 미국과 중국은 현재 AI 기술 경쟁에서 양강 체제를 형성하고 있다. 미국은 딥러닝, 자연어 처리(NLP), 생성형 AI 기술 등 AI의 핵심 기술 분야에서 여전히 우위를 점하고 있다. 구글, 오픈 AI, 메타와 같은 글로벌 빅테크 기업들은 AI 연구 및 상용화에서 선도적인 역할을 하고 있다. 특히, 오픈 AI의 챗 GPT는 생성형 AI 기술에서 미국이 세계적으로 압도적인 우위를 가지고 있음을 보여준다. 반면, 중국은 대규모 데이터와 정부 주도의 연구 개발을 바탕으로 AI 분야에서 빠르게 추격하고 있다. 텐센트, 알리바바, 화웨이와 같은 중국 대기업들은 자체 AI 칩 개발과 데이터센터 확장을 통해 기술력을 강화하고 있다. 최근 알리바바는 오픈 AI의 GPT 모델에 필적할 수 있는 새로운 AI 모델을 발표하며 기술적 진보를 과시</p>	0.0

Low Answer Relevancy

	query	generated_answer	answer	answer_relevancy
2	우리나라 기업들이 AI 기술을 활용해서 일상생활의 큰 변화를 일으킨 것에 대한 기사를 인용해	제공된 텍스트에는 한국 기업들이 AI 기술을 활용하여 일상생활에 큰 변화를 일으킨 구체적인 사례가 나와 있지 않습니다. 따라서 그 질문에 대한 답을 드릴 수 없습니다.	<p>어쩌다 만나던 AI가 이제 한집안 식구가 돼 가는 현실 [CES 2025]\nhttps://n.news.naver.com/mnews/article/082/0001306521?sid=101\n너도나도 일상 홈 AI 추진 본격화 삼성, 미래형 비전 적용 모델 공개 차량 등 연결 경험 확장도 가능해 LG, 생성형 AI로 IoT 제어 시연 자연어 대화 가능 제품도 선보여 'CES 2025' 개막을 하루 앞둔 6일(현지 시간) 미국 라스베이거스 만달레이 호텔 3층에서 전 세계 미디어와 기업 관계자들이 삼성 프레스 콘퍼런스 입장을 기다리고 있다(위). LG 전자의 'LG 월드 프리미어'에서 조주완 대표이사가 AI 혁신을 소개하는 모습. 삼성전자·LG 전자 제공</p> <p>삼성전자와 LG 전자가 7~10일(현지 시간) 미국 라스베이거스에서 열리는 세계 최대 IT·가전전시회 'CES 2025'에서 '인공지능(AI)홈을 통한 일상의 변화'를 나란히 제시했다. AI 경험을 집뿐 아니라 차량, 선박 등 다양한 공간으로 확장하고, 이를 개인 맞춤 형태로 제공하는 식으로 진화하는 모습을 보여줬다. 중국 업체들도 AI 로봇과 AI 홈 등을 내세우며 한국 업체들을 추격하는 모습이다.</p> <p>먼저 삼성전자는 참가 기업 중 가장 넓은 3368㎡(약 1019평) 규모로 전시관을 마련했다. 이번 전시를 통해 '모두를 위한 AI'를 주제로 초개인화를 위한 '홈 AI'를 선보였다. 전시관에서는 최신 AI 기술과 스마트싱스 허브를 탑재한 다양한 제품을 비롯해 상황별로 개인화된 홈 AI 경험, 차량과 선박 등으로 확장되는 홈 AI의 진화 등을 공개했다. 전시에서는 삼성전자의 연결 경험이 차량·선박·비즈니스 공간으로 확장하는 모습도 확인할 수 있다.</p> <p>삼성전자는 '갤럭시 AI'와 마이크로소프트의 '코파일럿+PC' 기능을 모두 탑재해 더욱 향상된 AI 경험을 제공하는 '갤럭시 북5 프로·북360'도 소개했다. 갤럭시 북5 프로·북360은 화면 속 이미지나 텍스트를 쉽고 빠르게 검색할 수 있는 'AI 셀렉트', 오래되거나 화질이 낮은 사진을 선명하게 보정하는 '사진 리마스터' 등 다양한 AI 기능을 갖추고 있다. LG 전자도 2044㎡(약 618평) 규모로 집, 차량, 사무실 등 다양한 공간에서 LG만의 AI 비전과 혁신을 구체화한 전시관을 꾸렸다. AI 홈 존에서는 생성형 AI를 탑재한 'LG 씽큐 온', 온디바이스 AI 기반의 콘셉트 제품 등 다양한 AI 홈 허브가 집안 곳곳에 설치된 센서로 고객의 말과 행동, 주변 환경을 감지하고 연결된 사물인터넷(IoT) 기기를 최적의 상태로 제어하는 모습을 시연했다. 또한 주변 사물, 사람, 신호 등을 빠르고 정확하게 인식하는 '비전 AI' 기반의 콘셉트 차량에선 가상 운전 시뮬레이션을 활용한 '인캐빈 센싱' 솔루션을 경험할 수 있다. 이와 함께 AI 홈을 차량으로 확대한 이동식 맞춤 공간 '드라이브 경험(AV) 플랫폼'도 선보였다.</p> <p>지난 6일 미국 라스베이거스 만달레이베이 호텔에서 열린 'CES 2025 월드 프리미어' 행사에서 영구제빙된 인공지능(AI) 기반의 '미바이온'이 선보였다. 미바이온은 수출통제 옥</p>	0.0
7	세계적으로 IoT에 AI를 접목시키는 기술 발전과 경쟁에 대해 알려줄래?	제공된 텍스트에는 IoT에 AI를 접목하는 기술 발전이나 경쟁에 대한 정보가 없습니다. 죄송하지만 그 질문에 대한 답을 드릴 수 없습니다.	<p>중 AI발전 어느 정도기에 美가 견제?...반신반의' 넘어 '놀라움'\nhttps://n.news.naver.com/mnews/article/001/0015159009?sid=104\n美바이온은 수출통제 옥</p> <p>죄기에 'AI 굴기'도 관심...관련 기업 4천500개 AI 연구 분야서 양적으론 美에 앞섰으나 한계점도 분명 지적 중국 선전의 걸어 다니는 로봇 [짐 팬의 엑스 기사물 캡처. 재판매 및 DB 금지] (서울=연합뉴스) 권숙희 기자 = 조 바이든 미국 행정부가 지난 13일(현지 시간) 중국에 대한 인공지능(AI) 반도체의 수출 통제를 한 단계 더 강화하기로 했다고 발표하면서 중국의 'AI 굴기'가 어디까지 이르렀는지에 대한 관심도 덩달아 커지고 있다. 중국은 지난 수십년간 '세계의 공장'을 자임하며 초저가 제품을 쏟아내는 데 집중했던만큼 첨단 기술과 막대한 자본이 필요한 과학기술력에 대해서는 이제껏 반신반의하는 국외 시각이 지배적이었으나 최근 들어서는 그 기류도 달라지는 분위기다.</p> <p>지난 10일 미국 라스베이거스에서 폐막한 세계 최대 가전·IT 전시회인 CES에서는 최근 몇 년간 옹성러들었던 것으로 보였던 '중국의 기술 저력'이 다시 기지개를 켜는 듯한 모습을 보였다. 미국 정부가 제재대상에 올린 화웨이와 DJI(드론업체) 등 중국 대표 테크 기업들이 빠졌음에도 AI 기술을 기반으로 업그레이드된 로봇 청소기 등 사물인터넷(IoT) 제품들과 중국의 AI 자율주행 전기차들은 관람객 눈길을 사로잡기에 충분했다. 이러한 변화의 흐름은 진작부터 감지된 듯 지난해 8월 미국 경제·혁신 정책 싱크탱크인 정보혁신재단(ITIF)은 '중국은 AI 분야에서 얼마나 혁신적인가'라는 제목의 보고서에서 "중국이 AI에 대한 끊임없는 추진력과 전략적 투자로 미국을 따라잡거나 능가하는 것은 시간문제일 뿐"이라고 지적했다. 특히 중국이 AI 연구 분야에서 양적으로는 이미 미국을 앞섰다는 것이 전문가들의 대체적인 평가다. AI 관련 기업들의 양적 성장도 두드러진다. 중국 공업정보화부의 장원밍 부부장은 지난해 11월 말 중국 저장성 우전에서 열린 '2024 세계 인터넷 정상회의'에서 "중국의 AI 산업은 급속히 발전하고 있다"면서 "중국 내 AI 관련 기업이 4천500개를 넘어섰다"고 밝혔다.</p> <p>아울러 AI 융합 응용 분야도 발전을 거듭해 중국은 421개 국가급 스마트 제조 시범 공장을 육성했다. 응용 분야에 속하는 스마트 도시(AI 기술이 적용된 교통·안전 시스템), 자율주행차, 의료 AI, 언어 AI(음성 인식과 번역) 등에서도 중국은 세계를 선도하고 있다는 평가를 받는다. 중국의 AI 발전에 대한 세간의 놀라움을 직접적으로 보여주는 사례가 최근에 있었다. AI 반도체 선두인 엔비디아(NVDA)의 수석 연구원이 지난 9일 엑스(X·옛 트위터)에 공유한 한 로봇 영상이 조회수 230만회를 넘기며 관심을 불러일으켰다. 팬 연구원은 거리에서 성큼성큼 걷는 휴머노이드 로봇의 영상을 공유하며 "선전의 이 충격적 영상이 진짜인지 알려주실 분?"이라는 글을 올렸다. 영상 속 로봇은 중국 선전에 있는 '엔진 AI'라는 스타트업 회사에서 만든 것으로 확인됐다. 영상이 사실로 밝혀지자 이는 "보행 자세가 멋지기는 하지만, 이런 로봇이 하루빨리 화제 진압에 투입될 수 있어야 한다"고 덧붙였다.</p> <p>홍콩 사우스차이나모닝포스트(SCMP) 따르면 '엔진 AI'의 영상 속 주인공인 '에스이(SE)01'은 키 170cm에 이족보행을 하며 팔굽혀펴기, 물건 집기 등 동작을 수행할 수 있다. 정부의 전폭적인 지원을 통해 급격하게 성장한 중국의 AI 분야는 물론 여전히 극복해야 할 여러 한계점을 갖고 있다. 중국 내 AI 관련 기업들은 높은 투자 대비 낮은 수익성을 보이고 있으며, 특히 휴머노이드 로봇 분야에서 상용화에 어려움을 겪고 있다. 무엇보다 핵심 기술을 미국 등 선진국에 의존하고 있다는 점이 가장 큰 문제다. 이 때문에 바이든 행정부가 임기를 일주일 남기고서 중국을 겨냥해 AI 개발에 필요한 반도체 판매와 관련한 신규 수출통제를 발표한 것으로 풀이된다. 이에 중국 상무부는 "미국의 수출 통제 남용은 글로벌 과학기술 혁신에도 심각한 영향을 미치고, 미국 기업을 포함한 전 세계 기업에 심각</p>	0.0

Low Context Recall

	query	generated_answer	answer	context_recall
16	우리나라 정부에서 AI를 활용한 정보화 교육 프로그램에 대한 기사를 찾을 수 있을까?	네, 위 기사는 한국 정부가 2025년부터 개별 맞춤형 교육을 위해 AI 교과서 도입을 추진하고 있다는 내용을 담고 있습니다. AI 교과서는 학생들의 학습 수준에 맞춰 개별화된 학습을 제공하고, 교사는 학생들의 활동 데이터를 통해 학습 수준을 파악할 수 있도록 설계되었습니다. 하지만, 저연령 학생의 경우 디지털 시민 교육을 먼저 진행한 후 AI 교과서를 활용할 예정이며, 국회의 논의 결과에 따라 올해 학교에서의 사용 여부가 결정될 것이라고 합니다. 따라서 이 기사는 AI를 활용한 정보화 교육 프로그램에 대한 내용을 포함하고 있습니다.	논산시 시민 정보화 교육, 스마트폰·AI 등[논산소식]\nhttps://n.news.naver.com/mnews/article/003/0013026222?sid=102\n[논산=뉴시스]논산시청 경. 2025. 01. 20 photo @ newsis.com *재판매 및 DB 금지 [논산=뉴시스]광송훈 기자 = 충남 논산시가 시민들의 디지털 정보격차 해소와 IT 기기 활용 능력을 향상시키기 위해 '2025 년 시민 정보화 교육'을 실시한다. 시민 정보화 교육은 스마트폰 기초 및 활용, 컴퓨터 기초 및 활용, AI 를 활용한 블로그 제작, 쉽게 배워보는 AI , SNS 기초 및 활용 등 현 사회적 트렌드에 맞는 인공지능(AI) 교육과정을 신규 발굴·개설해 총 44개 과정으로 운영할 예정이다. 스마트폰 기초 및 활용 과정은 23일부터 29일까지 신청 접수를 받아 2월 3일 시작, 12월까지 매월 교육과정별 3~4주로 진행된다. 교육장소는 교육생이 지역별 접근이 편리한 장소를 선택하여 교육을 받을 수 있도록 농업인회관 1층에 위치한 정보화 상설교육장과 강경도서관 2층 교육장, 연무도서관 2층 교육장 등 세 곳에서 이루어진다. 시 관계자는 "이번 교육과정은 디지털 기술에 익숙하지 않은 시민들에게 친숙하게 다가갈 수 있도록 교육수준을 고려하여 편성했다"면서 "챗 GPT 등 생성형 AI 를 활용한 신규 과정 개설로 최신 기술을 활용할 수 있는 좋은 계기가 될 것"이라고 말했다. [논산=뉴시스]논산시는 수입식품의 유통·판매 안정성을 강화하고 시민들의 안전을 보호하기 위해 20일부터 24일까지 외국 식료품 판매업소에 대한 점검을 실시한다. (사진=논산시 제공). 2025. 01. 20 photo @ newsis.com *재판매 및 DB 금지 <논산시, 외국 식료품 전문판매업소 위생점검 실시> 충남 논산시는 수입식품의 유통·판매 안정성을 강화하고 시민들의 안전을 보호하기 위해 20일부터 24일까지 외국 식료품 판매업소에 대한 점검을 실시한다고 밝혔다. 2023년 기준으로 논산의 외국인 주민 비율은 8.2%에 달하며, 이에 따라 관내 외국 식료품 판매업소도 매년 증가하고 있다. 이에 따라 시는 외국 식료품 위생관리 수준을 높이기 위해 매월 1회 점검을 실시할 계획이다. 이번 점검에서는 소비기한 경과 제품 진열·보관·판매·사용 여부, 무신고·무표시 제품의 유통·판매 여부, 냉장·냉동 제품의 적정 보관 여부 등을 중점적으로 확인할 예정이다. 또 영업자들이 자율적으로 위생관리를 할 수 있도록 외국 식료품 판매업소에 준수 사항을 적극적으로 홍보할 예정이다. 시 관계자는 "소비자들의 안전을 최우선으로 생각하며, 주기적인 점검을 통해 위생적이고 안전한 소비 환경을 조성하겠다"며 "취소문·다함께·내고·말했다"개념도 그에 따라 발전합니다. 그러면 AI 에 대한 직관적인 이해를 포기해야 할까요? 아닙니다. 어떤 사물이나 현상을 바라보는 과정에서, 유일하게 정확하고 진실한 관점을 찾느라 시간을 허비해선 안 됩니다. AI 는 발전하고 있는 분야이기 때문에, 불변하는 단 하나의 정의란 존재하지 않는다고 보는 게 더 합리적 태도입니다. AI 를 어떻게 활용하고, 거기에 어떤 이익을 얻을지를 고민하는 것이 중요합니다. AI 에 대한 자체적인 관점과 프레임을 설정하고, 그에 맞춰 AI 를 바라보는 겁니다. 그렇게 하면 추상적인 개념이나 정의, 철학적인 수렁에 빠지지 않을 수 있습니다. 중요한 건 우리가 AI 로 무엇을 할 것이냐, 무엇을 할 수 있는냐는 것이니까요. 알수록 아리송... AI 를 바라보는 유용한 프레임 : '예측기계' 어제가 애그러월, 조슈아 갠즈, 아비 골드파브가 공저한 책 '예측 기계'의 원서 표지(국내 절판). 이 책에서 저자들은 '인공지능은 예측기계다'라고 주장합니다. 그런 점에서, AI 를 바라보는 유용한 프레임을 하나 꼽는다면, 'AI 는 예측하는 기계다'라는 프레임이라고 생각합니다. 캐나다 토론토대학 로트만 경영대학원 교수인 조슈아 갠즈, 아비 골드파브 세 사람은 2018년 출간한 책 예측기계(Prediction Machines)에서 이렇게 썼습니다. " AI 라는 새로운 물결이 가져다준 것은 지능이 아니라 지능의 중요한 요소인 예측이다." AI 로 무엇을 할지, 무엇을 할지 고민이라면 그때 바로 '예측'을 떠올려보세요. '지금 내 업무에서 무엇을 예측할 수 있지?', '우리 회사가 무슨 예측을 하면 좋을까?' 실제 사례를 보죠. 비즈니스의 세계는 수많은 예측으로 굴러가고 있습니다. 유니클로는 계절적 변화와 지역별 트렌드를 예측해서 옷의 재고를 조절합니다. 예측이 정확할수록 재고비용을 절감할 수 있습니다. 아마존은 소비자의 이전 구매 데이터를 바탕으로 다음 구매 상품을 예측합니다. 추천 상품이라는 이름의 예측을 통해 아마존은 매출을 극대화할 수 있죠. 은행은 대출 신청자의 상환 가능성을 예측해볼 수 있습니다. 신용 점수와 과거 거래 기록을 분석해 신용 위험을 분석하고, 적합한 마진이율을 설정할 수 있죠. 보험사 역시 가입자의 질병 발생률 등을 정확히 예측한다면 보험료 마진을 높일 수 있습니다. 농부는 날씨와 토양 상태 등을 예측하면 수확량을 예측할 수 있고, 역시 물류 부담을 낮출 수 있습니다. 넷플릭스와 같은 엔터테인먼트는 시청자의 각종 데이터를 기반으로 정확한 추천(예측)을 하고, 소비를 유인하죠. 언론사도 예측의 힘에 기댈 수 있습니다. 독자 데이터를 파악해서 독자가 좋아할 만한 읽을 만한 뉴스를 성공적으로 예측한다면, 트래픽 향상을 거둘 수 있죠. 즉, 예측은 모든 산업에서 의사결정 토대가 됩니다. 어떻게 예측하느냐에 따라 대응이 달라지죠. 결과도 달라집니다. 소비자를 만족시킬 수 있게 되고, 비용을 절감하며 매출을 증대시킬 수 있죠. 예측력이 바로 경쟁력입니다. AI 의 힘 : 예측 비용을 낮춘다 그런데, 잠깐 예측은 AI 만이 할 수 있는 게 아닙니다. 예측 없이 하루를 살아가는 사람이 있나요? 직장인 누구나 어느 정도 예측을 하며 일을 하고, 기업도 이미 예측에 기반해 조직을 운영하죠. 과거의 데이터를 보지도 않고 의사결정을 하는 사람은 없습니다. 예측은 AI 만이 갖고 있는 특징이 아니라, 인간의 보편적 능력입니다. 그러면 왜 AI 가 예측기계일까요? 왜 우리는 AI 를 예측기계라고 부를 수 있는 걸까요? 인간과 AI 는 모두 예측을 합니다. 그러나 AI 의 예측 능력은 그 규모와 복잡성에서 인간을 압도합니다. 인간은 한정된 데이터, 경험, 기억, 직	0.500000
33	AI가 다양한 산업에서 어떻게 예측을 활용하고 있는 지 설명한 사례가 있어? AI가 예측을 잘할 수 있는 이유는 무엇이야?	가입자의 질병 발생률을 예측하여 보험료 마진을 높입니다.\n* **농부: ** 날씨와 토양 상태를 예측하여 수확량을 예측하고 물류 부담을 줄입니다.\n* **넷플릭스: ** 시청자 데이터를 기반으로 정확한 콘텐츠 추천(예측)을 통해 소비를 유도합니다.\n* ** 언론사: ** 독자 데이터를 분석하여 독자가 선호할 만한 뉴스를 예측하고 트래픽을 향상시킵니다.\n\n\n* **AI가 예측을 잘하는 이유: **\n\nAI는 인간보다 훨씬 더 크고 복잡한 규모의 데이터를 처리할 수 있습니다. 인간은 제한된 데이터, 경험, 기억, 직관에 의존하여 예측하지만, AI는 다음과 같은 강점을 가지고 있습니다.\n\n* **방대한 데이터 처리: ** 수십 개의 변수를 포함하는 대규모 데이터를 처리하고, 인간이 감지하지 못하는 복잡한 상관관계와 패턴을 파악합니다.\n\n* **비선형적 관계 파악: ** 딥러닝과 머신러닝 등의 알고리즘을 통해 비선형적인 관계와 숨겨진 패턴을 발견합니다. 인간은 선형적 사고에 익숙하지만, 현실은	0.571429	

입력하신 검색어가 길어 '최근 ai~게 되고,' 까지만 검색된 결과입니다. [도움말 보기](#)

옵션 ▾

• **관련도순** • 최신순 | 모바일 메인 언론사 ☐

'최근 ai 업계에서 비디오 생성 모델이...'에 대한 검색결과가 없습니다.

- 단어의 철자가 정확한지 확인해 보세요.
- 한글을 영어로 혹은 영어를 한글로 입력했는지 확인해 보세요.
- 검색어의 단어 수를 줄이거나, 보다 일반적인 검색어로 다시 검색해 보세요.
- 두 단어 이상의 검색어인 경우, 띄어쓰기를 확인해 보세요. [네이버 맞춤법 검사기](#)
- 검색 옵션을 변경해서 다시 검색해 보세요.



AI로 사진 속 물체 지우기

AI 지우개 CLOVA X



매일 달라지는 타임 특가

설맞이 첫 세일! 10% 쿠폰 어택 받기 >



네이버에서 넷플릭스를

지금 멤버십 시작하기



아무리 추워도 여행은 못 참지

겨울에 가기 좋은 국내 여행지

Query : 최근 ai 업계에서 비디오 생성 모델이 뭐가 있을까? 그리고 그 모델들의 성능 비교는 어떻게 되고, 앞으로의 전망은 어때?


```
# 대화형 QA 시스템 실행
interactive_qa()
```

AI 질문 답변 시스템 (종료하려면 'quit' 입력)

질문을 입력하세요: 최근 AI 업계에서 비디오 생성 모델이 뭐가 있을까? 그리고 그 모델들의 성능 비교는 어떻게 되고, 앞으로의 전망은 어때?

답변 생성 중...

답변:

최근 AI 업계에서 주목받는 비디오 생성 모델로는 오픈AI의 Sora, 구글의 Imagen Video (비오2), 메타의 Make-A-Video (무비젠)이 있습니다.

****성능 비교:**** 세 모델 모두 장단점이 있으며, 단순한 성능 비교는 어렵습니다. 기사에 따르면, 메타는 자사의 Make-A-Video가 오픈AI의 Sora보다 더 나은 결과물을 만들어냈다고 주장하지만, 정식 출시되지 않아 실제 성능(

****앞으로의 전망:**** 동영상 콘텐츠 수요 증가와 AI를 통한 영상 제작 시간 및 비용 절감 효과로 인해 AI 비디오 생성 도구 시장은 계속 성장할 것으로 예상됩니다. 각 기업은 기술력을 더욱 발전시켜 해상도, 물리적 정확도,

참고 출처:

1. <https://n.news.naver.com/mnews/article/009/0005429554?sid=105>
2. <https://n.news.naver.com/mnews/article/009/0005429554?sid=105>
3. <https://n.news.naver.com/mnews/article/009/0005429554?sid=105>
4. <https://n.news.naver.com/mnews/article/052/0002142109?sid=101>

계속하려면 Enter를 누르세요...

최근 AI 업계에서 주목받는 비디오 생성 모델로는 오픈AI의 Sora, 구글의 Imagen Video (비오2), 메타의 Make-A-Video (무비젠)이 있습니다.

성능 비교 : 세 모델 모두 장단점이 있으며, 단순한 성능 비교는 어렵습니다. 기사에 따르면, 메타는 자사의 Make-A-Video가 오픈AI의 Sora보다 더 나은 결과물을 만들어냈다고 주장하지만, 정식 출시되지 않아 실제 성능에 대한 객관적인 평가는 부족합니다.

Sora는 일반 사용자를 대상으로, Imagen Video와 Make-A-Video는 전문가를 대상으로 하는 것으로 보이며, 각 모델의 강점이 다릅니다.

Sora는 텍스트 외 이미지 입력이 가능하고, Imagen Video는 높은 해상도를 자랑하며, Make-A-Video는 기존 영상 편집에 강점이 있다고 언급되어 있습니다. 물리 엔진 구현 능력 또한 모델마다 차이가 있을 것으로 예상됩니다.

앞으로의 전망 : 동영상 콘텐츠 수요 증가와 AI를 통한 영상 제작 시간 및 비용 절감 효과로 인해 AI 비디오 생성 도구 시장은 계속 성장할 것으로 예상됩니다.

각 기업은 기술력을 더욱 발전시켜 해상도, 물리적 정확도, 편집 기능 등을 개선하려 할 것이며, 사용자 경험을 향상시키는 데 집중할 것으로 보입니다.

하지만, 모델 훈련에 사용된 데이터의 투명성 부족과 저작권 문제 등 해결해야 할 과제도 남아 있습니다.

```
query = input("질문을 입력하세요: ")

if query.lower() == 'quit':
    print("프로그램을 종료합니다.")
    break

print("\n답변 생성 중...")
answer, sources = generate_answer(query)

print("\n답변:")
print(answer)
print("\n참고 출처:")
for i, source in enumerate(sources, 1):
    print(f"{i}. {source}")

input("\n계속하려면 Enter를 누르세요...")

# 대화형 QA 시스템 실행
interactive_qa()
```

... AI 질문 답변 시스템 (종료하려면 'quit' 입력)

질문을 입력하세요:

최근 AI 업계에서 비디오 생성 모델이 뭐가 있을까? 그리고 그 모델들의 성능 비교는 어떨까?

전체

이미지

뉴스

도서

동영상

지도

쇼핑

더보기

모두의연구소

<https://modulabs.co.kr> > blog > openaisora

OpenAI Sora, 비디오 생성 AI의 시대를 열다 | 블로그

2024. 2. 19. — OpenAI가 공개한 비디오 생성 AI모델 Sora에 대해 알려드립니다. 텍스트 프롬프트를 통해 최대 1분 길이의 비디오를 생성할 수 있는 AI 모델 Sora를 ...

누락된 검색어: 뭐가 있을까? 그리고 전망 어때?

hahahahr.com

<https://hahahahr.com> > business-ai

영상 생성 AI 비교 : runway VS Lumalab 성능을 비교해보자.

2024. 6. 18. — LumaLabs AI의 Dream Machine은 텍스트와 이미지에서 고품질의 현실적인 비디오를 생성하는 최신 AI 모델이다. 주요 기능으로는. 1) 사용자가 입력한 ...

AI 히어로즈

<https://aiheroes.ai> > community

구글 영상 생성 AI '비오2(Veo2)' 소라(Sora)와 비교하기

2024. 12. 27. — 오픈 AI에서 공개한 소라(Sora) AI 영상 생성 모델이 올해 많은 관심을 받았는데, 구글 딥 마인드의 비오2 모델은 소라에 비해 해상도는 4배, 생성 ...

누락된 검색어: 최근 뭐가 있을까? 그리고 전망 어때?

네이버 프리미엄콘텐츠

<https://contents.premium.naver.com> > aipost > contents

"더 사실적인 영상 생성"...루마 AI, 새로운 영상 생성 모델 '레이 2 ...

5일 전 — 동영상 생성 인공지능(AI) 스타트업인 루마 AI(Luma AI)가 새로운 비디오 생성 모델인 레이 2(Ray 2)를 공개했다. 드림 머신 홈페이지와 모바일 앱을 ...

최근 AI 업계에서는 다양한 비디오 생성 모델이 개발되고 있습니다. 주요 모델과 그들의 성능 및 전망을 살펴보겠습니다.

주요 비디오 생성 모델

1. KlingAI:

- 특징: 중국에서 개발된 이 모델은 텍스트를 기반으로 비디오를 생성하는 'T2V'와 기존 이미지를 활용한 'I2V' 기능을 제공합니다. 현재 1.5버전까지 출시되었으며, 무료 버전은 5초 길이의 비디오만 생성할 수 있습니다. 생성 속도가 다소 느리지만, 유료 버전으로 업그레이드하면 더 긴 비디오와 빠른 생성 속도를 경험할 수 있습니다.

2. RunwayML의 Gen-3 Alpha Turbo:

- 특징: 이 모델은 이미지에서 비디오로 변환하는 기능을 제공하며, 기존 모델보다 7배 더 빠른 속도로 작업을 수행할 수 있습니다. 가격이 절반으로 줄어들어 많은 사용자들에게 접근성이 높아졌습니다.

3. Haiper AI:

- 특징: 비교적 신생 플랫폼으로, 텍스트 기반 비디오 생성에 최적화되어 있습니다. 사용자가 입력한 간단한 텍스트를 기반으로 고품질의 비디오 콘텐츠를 생성할 수 있으며, 빠른 렌더링 속도와 사용자 친화적인 인터페이스를 제공합니다.

4. 구글의 Veo 2:

- 특징: 구글이 최근 발표한 동영상 생성 AI로, 텍스트 프롬프트를 기반으로 고품질의 동영상을 자동으로 생성합니다. 시각적 스토리텔링 기대됩니다.

5. 메타의 Movie Gen:

- 특징: 메타가 발표한 새로운 AI 모델로, 텍스트 프롬프트를 통해 현실감 있는 비디오와 오디오 클립을 생성할 수 있습니다. 이 모델은 비디오 편집, 사운드 매칭, 개인화된 콘텐츠 생성 등 다양한 기능을 제공합니다.

지식의 힘

최신 AI 영상 생성 플랫폼 대결 : Haiper AI, Lumalabs.ai, RunwayML Gen-3 ...

2024년 10월 28일 — 최근 AI 영상 생성 플랫폼의 신흥 강자로 떠오르고 있는 Haiper AI, Lumalabs....

서비스 고도화를 위한 작업

RAG 적용 초기 성과

- 네이버 뉴스 API를 활용한 RAG 모델 개발로 웹 기반 서비스 데이터 베이스 구축에 대한 가능성을 확인함
- RAGAS 평가 지표 기반 Faithfulness(0.91), Answer Relevancy(0.42), Context Recall(0.89), Context Precision(0.96)의 성능이 확인됨

서비스 고도화의 필요성 및 방안

<워크플로우>



사용자 쿼리 검색



쿼리에서 키워드 추출



검색 API 호출



텍스트 분할



임베딩 모델 생성



벡터 스토어 생성



벡터 스토어 검색



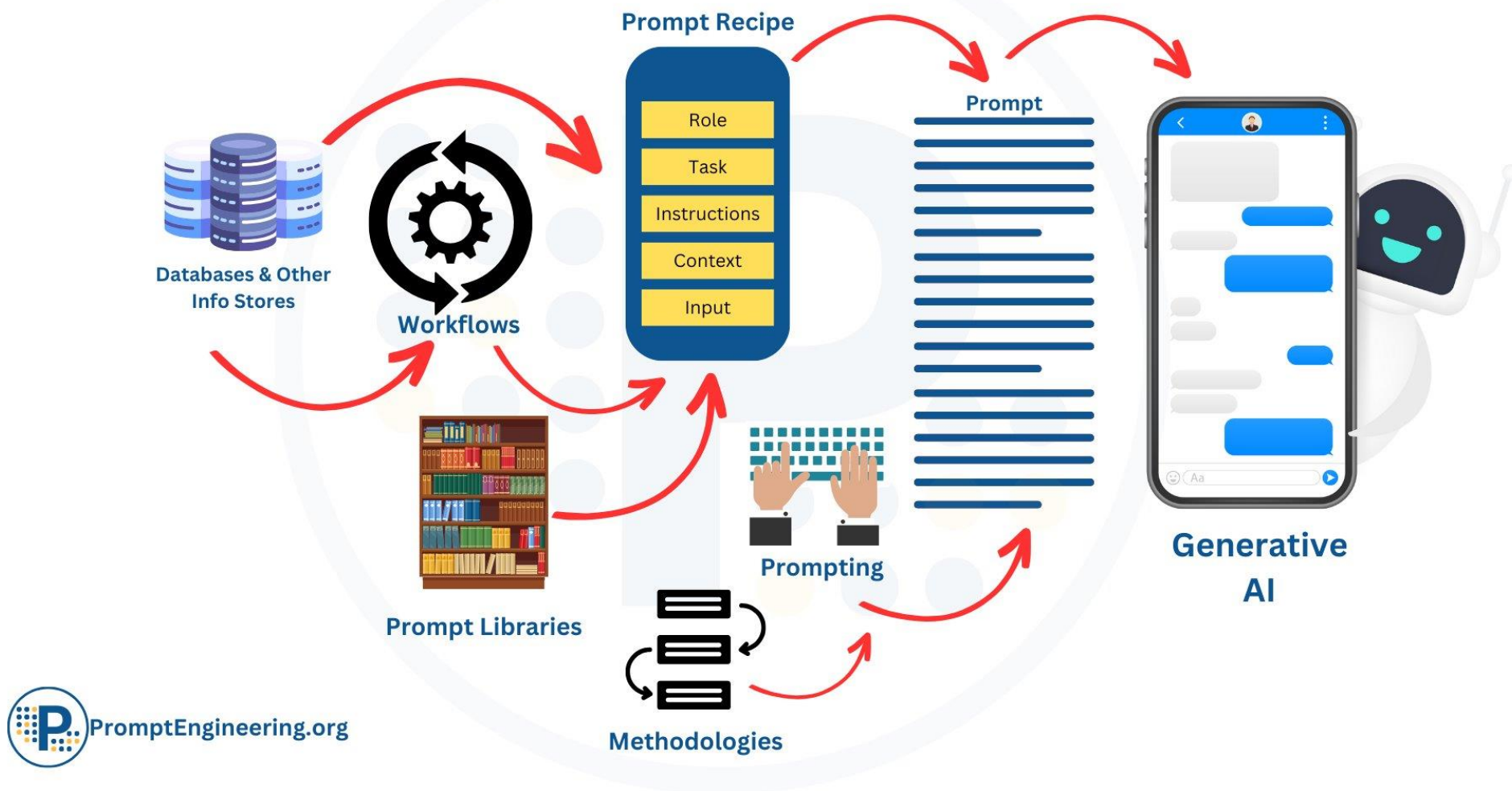
응답 생성

사용자 쿼리에 맞는 키워드를 추출하여
양질의 데이터 확보를 위한 개선 필요

내부 데이터 베이스 외 **외부 데이터 구축**을
통해 **실시간 사용자 맞춤형 데이터 베이스** 구축

뉴스레터 양식에 맞는 요약문을 제공할 수 있도록
LLM의 결과물에 대한 세부 조정 필요

What is Prompt Engineering? Everything that goes before the prompt



서비스 고도화 - Prompt Engineering

Phase 1: Prompt to Keyword Extraction

모델 한계

- 문장 형식의 쿼리를 사용하는 경우, 답변이 없거나 매우 희소한 결과값을 얻을 확률이 높음
- 사용자의 쿼리를 단순 활용하는 경우, 검색 키워드가 한정적으로 충분한 데이터를 얻지 못함
- 사용자의 쿼리가 모호하거나 불명확한 경우 원하는 정보값을 얻기 어려움

개선 사항

- 문장 형식의 쿼리에서 문맥을 유지하는 핵심 키워드를 추출하여 네이버 뉴스 검색에 적합한 포맷으로 변환함
- 핵심 키워드 추출 및 중복되거나 모호한 표현을 배제함으로써 더욱 명확한 쿼리를 생성하도록 지원함
- 동의어 및 유의어 범위까지 확장하여 유사한 의미의 키워드를 포괄하는 키워드 리스트셋을 생성함

Phase 2: Chunk Data to Summarization

- 상용 모델을 사용하더라도 구체적 입력값이 주어지지 않았을 때 일반적으로 출력되는 요약문의 형식은 한정되어 있음. (개조식 서술, 근거-결론식 구조)
- 각각 다른 쿼리를 입력할 때마다 상이한 형식, 어투의 요약문이 출력될 수 있다.
- 기존에 LangChain에 내장된 기본 Q&A 프롬프트 서식을 사용해 출력.

- 서비스에서 제공하고자 하는 의도에 맞는 일관된 어투, 서로 다른 쿼리에 대해서도 동일한 형식의 요약문 출력
- 뉴스 본문을 읽지 않아도 될 정도의 세부정보 유지, 불필요한 정보를 제거한 요약문.
- 키워드 추출 과정에서 얻은 증강된 쿼리 정보를 활용

Prompt Engineering: Prompt to Keyword Extraction

Base LLM Model : Llama-3.2-3B-Instruct (Quantized Model)

Step 1: Prompt for Key Phrases Extraction

"" You are an intelligent assistant specializing in generating search-friendly keyword phrases. Your task is to extract concise, contextually accurate keyword phrases from a given user query. These keyword phrases should be optimized for news search and adhere to the following guidelines:

1. Ensure each phrase is specific and clear.
2. Maintain the user's intended context.
3. Include relevant dates, time periods, or locations when applicable.
4. Use Named Entities (e.g., people, places, organizations, events) mentioned in the query.
5. Incorporate synonyms or related terms to broaden the search scope.
6. Keep phrases concise and avoid redundant words.
7. Construct keyword phrases suitable for Boolean operators (e.g., AND, OR).
8. Include regional or contextual relevance when specified.
9. Reflect actions or dynamic processes, not just static terms.
10. Ensure the phrases are structured for search engine optimization.

****Output Format**:**

- Provide 3-5 keyword phrases as a numbered list.
- Each phrase should be a complete and logical search query.

****Example User Query**:**

"2025년에 헬스케어 분야에서 AI 기술이 어떻게 발전하고 있나요?"

****Example Output**:**

1. 2025년 AI 기술 헬스케어 발전
2. AI 헬스케어 트렌드 2025
3. 2025년 헬스케어 AI 최신 동향
4. AI 기반 헬스케어 혁신 2025

Now, extract keyword phrases from the following query:

"{USER_QUERY}"

LLM 역할 명시

검색 엔진 기반 포맷에 맞춰 사용자 쿼리 기반 핵심 구문을 파악하여 추출하는 역할 부여

가이드라인 제공

뉴스 검색에 적합한 키워드의 특징(i.e. 고유명사 포함, 키워드 관련 유의어 생성 등) 및 결과물에 필수적으로 포함되어야 하는 내용 명시

One-shot Learning

사용자의 쿼리 기반 핵심 구문 및 관련 동의어를 추출한 Key Phrases를 숫자 리스트로 제공할 수 있도록 예시 제공

<실행 결과>

• User Query:

"최근 AI 업계에서 비디오 생성 모델이 뭐가 있을까? 그리고 그 모델들의 성능 비교는 어떻게 되고, 앞으로의 전망은 어떤?"

• Output:

1. AI 비디오 생성 모델 전망
2. 최근 비디오 생성 모델 성능
3. AI 비디오 생성 모델 최신 동향
4. 비디오 생성 모델 성능 향상 전망

Step 2: Prompt for Keywords Extraction

"" You are an intelligent assistant specializing in keyword extraction.

Your task is to extract ****Matched Keywords**** and ****Related Keywords**** from a given document.

The output should be presented as two separate numbered lists for clarity.

Make sure that the list should be written in query language.

****Guidelines**:**

1. ****Matched Keywords****: Extract up to 5 keywords explicitly mentioned in the document.
2. ****Related Keywords****: Extract up to 5 keywords contextually related to the document but not explicitly mentioned.

****Output Format**:**

- Matched Keywords:

1. [Keyword 1] #n2. [Keyword 2] #n3. [Keyword 3]...

- Related Keywords:

1. [Keyword 1] #n2. [Keyword 2] #n3. [Keyword 3]...

****Example Input**:**

Document: "트랜스포머 기법을 쓰지 않은 LLM모델의 트렌드"

****Example Output**:**

- Matched Keywords:

1. 트랜스포머 #n2. 기법 #n3. LLM #n4. 모델

- Related Keywords:

1. 자연어 #n2. 언어모델 #n3. AI #n4. 챗봇

Now, extract keywords for the following document:

Document: {USER_QUERY}

LLM 역할 명시

사용자의 쿼리를 기반으로 매칭 키워드와 연관 키워드를 추출하는 LLM으로써의 역할 부여

가이드라인 제공

각 매칭 키워드와 연관 키워드 추출 방식에 대한 세부적인 가이드라인 제시

추출 포맷 및 예시 제공

리스트 형태의 결과물을 반환할 수 있도록 Output 형식 제공 및 One-shot Learning을 위한 예시 제공

<실행 결과>

• User Query:

"최근 AI 업계에서 비디오 생성 모델이 뭐가 있을까? 그리고 그 모델들의 성능 비교는 어떻게 되고, 앞으로의 전망은 어떤?"

• Output:

['비디오', '생성', '모델', '성능', '전망', 'AI', '컴퓨터', '화면', '영상', '인공지능']

Prompt Engineering: Prompt for Summarization 🧐

Base LLM Model : Gemini-1.5 Pro API

```
...
SystemMessagePromptTemplate(
    prompt=PromptTemplate(
        input_variables=['context'],
        input_types={},
        partial_variables={},
        template="Use the following pieces of context to answer the user's question. \nIf you don't know the answer, just say that you don't know, don't try to make up an answer.\n-----\n{context}"),
    additional_kwargs={},
    HumanMessagePromptTemplate(
        prompt=PromptTemplate(
            input_variables=['question'],
            input_types={},
            partial_variables={},
            template='{question}'),
        additional_kwargs={})
    ...
```

위의 기존 LangChain 내부의 프롬프트 서식은 너무 단순하다!

- 퓨샷 프롬프트 - 예시 제공 추가
- 쿼리와 Retrieve된 컨텍스트 뿐 아니라 제목, 연관키워드 등을 추가할 수 있다.
- 쿼리가 시스템 메시지와 혼동되지 않게 수정

```
Template = '''\
<SYSTEM>
당신은 출판사의 담당 편집자입니다.
아래 질문(query)을 바탕으로 주어진 문서(document)를 다섯 문장 이내로 요약(summary)해야 합니다. 키워드 문장(keywords)과 제목(title)을 참고해서 생성하세요.

<EXAMPLE 1>
query:
미국 정치권에 발생한 중요한 이슈에 따라서 첨단 산업에 어떤 변화가 있을까?
keywords:
미국 정치 이슈와 첨단 산업 트렌드
title:
"트럼프 2.0시대 개막 4대 테마는?" 에너지, AI·로봇, 우주, 안보
document:
미국 제47대 대통령 선거뿐만 아니라 상·하원 모두 공화당이 승리함에 따라...
summary:
미국 정치 이슈와 첨단 산업의 트렌드에 대해 트럼프 2.0 시대의 미국 우선주의와...</s>

<EXAMPLE 2>:
query:
보험 산업에서 AI 기술을 활용하여 어떤 새로운 서비스와 사업이 등장하고 있나요?
keywords:
보험산업과 AI
title:
시니어·AI·해외시장...신사업 시동거는 보험사
document:
삼성생명 시니어리빙 TF →시니어 Biz 팀 격상 교보생명 '보장분석 AI 서포터'...
summary:
보험 산업에서 AI 기술을 활용한 새로운 서비스와 사업이 등장하고 있습니다...</s>

<ASSISTANT>:
query:
AI 에이전트가 기업의 생산성과 효율성을 향상시키는 데 어떻게 사용되고 있나요?
keywords:
에이전트 AI
title:
"AI와 AI가 협업하는 시대"...새해 핵심 트렌드 'AI 오케스트레이션'
document:
아마존웹서비스·마이크로소프트·카카오 등 주력 서비스로 제시 2025년 인공지능( AI ) 산업...
summary:
...
'''
```

예시 문제

목표 문제

Prompt Engineering: Prompt for Summarization 🧐

Base LLM Model : Gemini-1.5 Pro API

요약 전 기사원문

... 이를 극복하기 위해 AI 기술을 유기적으로 결합해 복잡한 문제를 해결하고 업무 효율을 높이려는 기업들의 수요가 급증하고 있다. 특히 AI 에이전트의 등장과 지속된 불황으로 인한 기업들의 생산성 향상 요구가 맞물리며 AI 오케스트레이션에 대한 관심이 높아지는 추세다. 얼라이드 마켓 리서치(AMR)는 글로벌 AI 오케스트레이션 시장이 2031년까지 연평균 21.5% 성장하여 약 352억 달러에 이를 것이라는 시장 조사 결과를 발표했다. AWS의 베드록 에이전트(이미지=아마존웹서비스) **AWS는 지난 12월 '아마존 베드록 에이전트(Amazon Bedrock Agents)'를 발표했다.** 이는 하나의 AI 에이전트가 여러 보조 에이전트를 조율해 복잡한 작업을 수행하도록 설계된 멀티 에이전트 프레임워크다. 복잡한 프로젝트를 여러 단계로 나눠 각 단계에 가장 적합한 전문화된 AI 에이전트를 할당할 수 있어 AI의 업무 정확성을 높이고 낭비되는 컴퓨팅 자원을 최소화할 수 있도록 지원한다. **마이크로소프트도 이그나이트 2024에서 애저 AI 파운드리를 발표하며 코파일럿 등 다양한 AI 제품을 통합 활용할 수 있는 환경을 제시했다.** 유아이패스는 2025년 AI와 자동화 분야의 핵심 트렌드로 'AI 오케스트레이션'을 지목하며 에이전트 생태계의 핵심이 될 것으로 전망했다. 이와 함께 RPA와 AI 오케스트레이션을 결합해 금융, 헬스케어 등 여러 산업의 복잡한 업무를 효율적으로 처리하기 위한 방안을 제시한다. 마이크로소프트의 애저 AI 파운드리(이미지=마이크로소프트) 국내 주요 기업도 AI 오케스트레이션 기술을 적극 도입하고 있다. **카카오는 AI 오케스트레이션 전략을 통해 다양한 AI 모델을 조합해 최적의 성능과 비용 효율성을 추구하고 있다.** 현재 개발 중인 AI 메이트 '카나나'의 경우 개인 AI 메이트 '나나'와 그룹 AI 메이트 '카나'로 구성돼 있으며 AI 오케스트레이션을 통해 두 메이트를 효율적으로 활용할 수 있는 환경을 구축했다. 더불어 카가오의 클라우드 AI 플랫폼과 연결해 다양한 외부 데이터 소스와 AI 모델과 연계도 가능하다. 카카오의 카나나나 나나와 카나 2개의 AI 모델을 AI 오케스트레이션으로 최적화한다(이미지=카카오) **LG CNS는 AI 오케스트레이션 기술을 스마트 팩토리와 물류 자동화 시스템에 적용하고 있다.** 스마트 팩토리에서는...

프롬프트 적용 후 생성

AI 오케스트레이션은 여러 AI 모델과 시스템을 통합하여 복잡한 작업을 효율적으로 수행하는 기술입니다. 이 기술은 기업의 생산성과 효율성을 향상시키는 데 사용되고 있습니다. 예를 들어, **Amazon Web Services는 여러 보조 에이전트를 조율하여 복잡한 작업을 수행하도록 설계된 Amazon Bedrock Agents를 제공합니다.** 마이크로소프트는 다양한 AI 제품을 통합 활용할 수 있는 환경을 제공하는 **Azure AI Foundry를 발표했습니다.** 국내 기업인 **카카오는 AI 오케스트레이션을 통해 다양한 AI 모델을 조합하여 최적의 성능과 비용 효율성을 추구하고 있습니다.** **LG CNS는 AI 오케스트레이션 기술을 스마트 팩토리와 물류 자동화 시스템에 적용하여 제조 공정을 최적화하고 물류 운영 효율을 높이고 있습니다.**

Limitation & Future Work

모델 한계

- **다양한 모델에 대한 성능 평가 부족**
Gemini-1.5-Flash 외 다른 모델(Gemini-1.5-pro, Llama-3.2 등)의 성능 평가가 수행되지 않아, 모델 간 비교와 선택 기준이 부족함
- **정성적 평가 기준 부족**
성능 평가에서 RAGAS, BERT-Score, BLEU, ROUGE 등 정량적 지표가 사용되었으나, 사용자의 경험에 대한 정성적 평가 과정이 결여됨
- **키워드 추출 및 요약 생성문의 랜덤성**
키워드 추출 과정 및 요약문 생성을 위한 Prompt Engineering 과정에서 답변 생성에 대한 랜덤성이 확인되었으며, 결과적으로 안정적인 키워드 확보 및 일관적인 응답 생성이 어려움

개선 방안 및 향후 과제

- **모델 성능 평가 확장**
Gemini-1.5-pro, Llama-3.2 등 다양한 모델에 대해 성능을 비교 평가하여 서비스에 최적화된 모델 선정
- **정성적 평가 프로세스 도입**
정량적 평가 지표 외, 사용자 경험 중심의 객관적인 정성적 평가 방법을 도입하여 모델의 신뢰성 및 서비스 가능성에 대한 이해가 필요함
- **답변 생성의 일관성 강화**
Prompt Engineering을 고도화하여 키워드 추출 과정에서의 랜덤성을 줄이고, 추가적인 LLM 모델의 Fine-tuning 수행을 통해 요약 작업에 대한 성능을 강화