

건설산업에서의 sLLM 활용 방안

- 특정업무 맞춤형 학습 Case Study

2025. 05. 28.





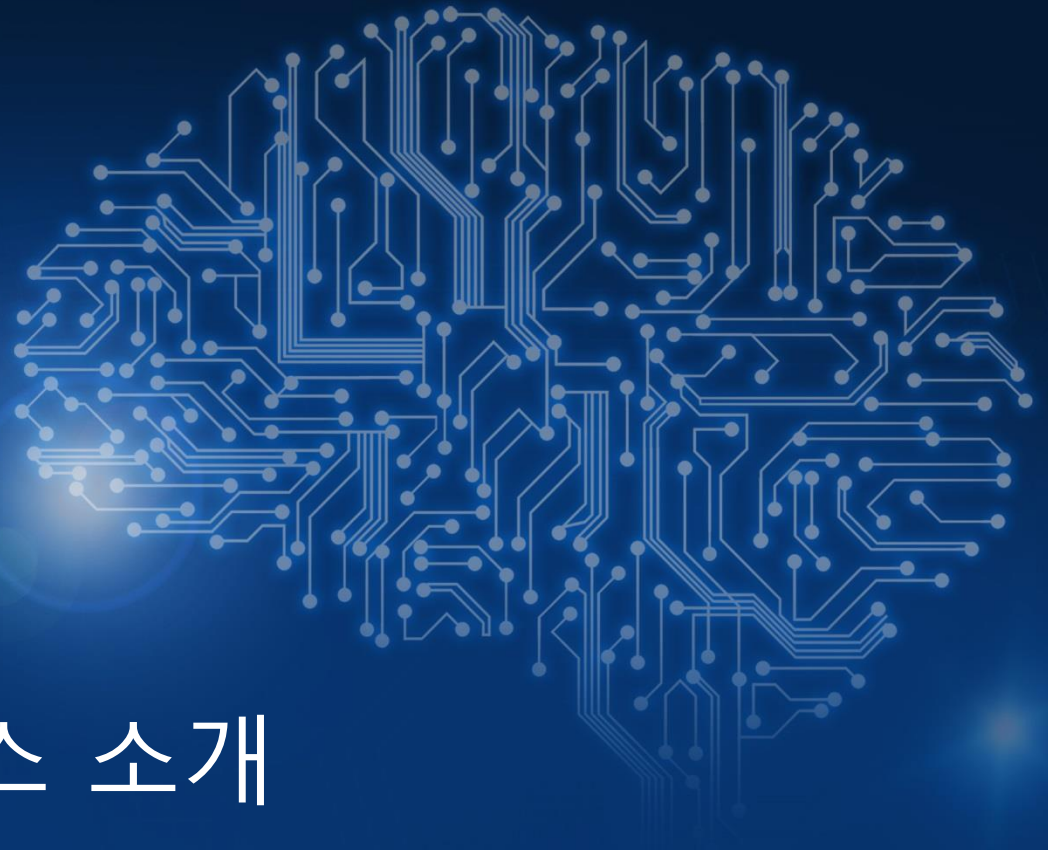
스펠릭스 소개

AI 기술 동향

건설산업의 AI 적용 사례

수행 Use Case 소개 - sLLM의 특정목적 문서분석 업무 맞춤형 학습

Q&A



스펠릭스 소개



회사 개요

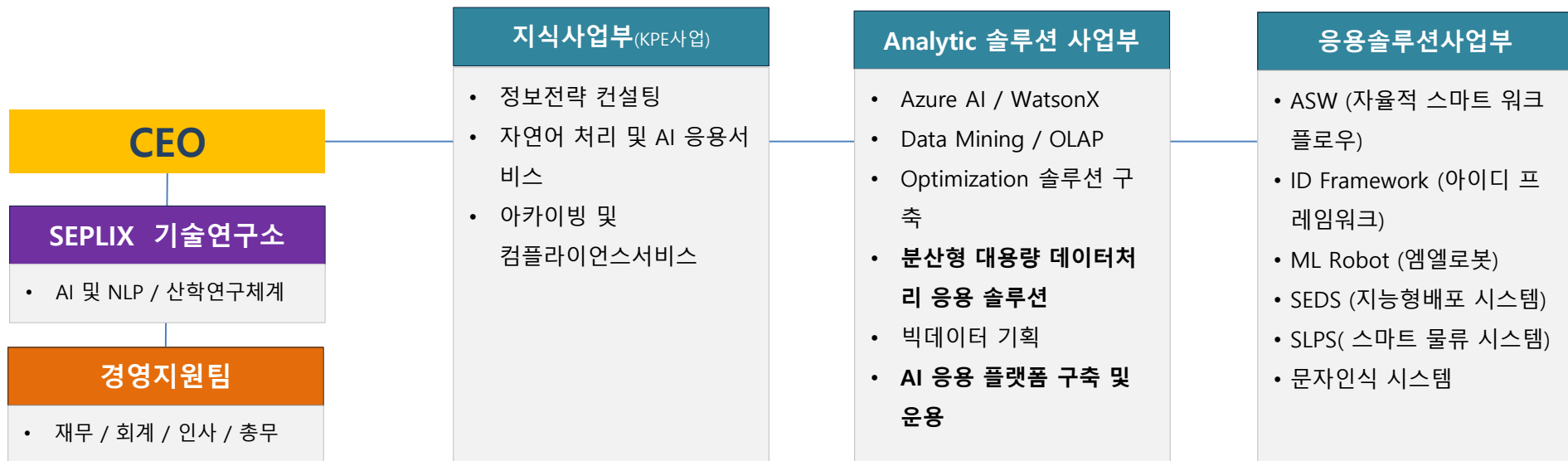
- (주)스펠릭스는 신선한 감각과 사고의 혁신으로 인간중심의 기술 구현을 목표로 합니다.
- DT, AI를 선도하며 Smart Digital Network 세상을 만들어 가는 기술중심의 기업입니다.
- 2005년 설립되어 검색엔진, 워크플로우, 대용량 데이터 분석 및 처리, 최적화, 통계 등에 대한 기술적 노하우를 보유하고 있으며 **대용량 데이터, 분산처리**와 오픈 소스에 대한 철저한 이해를 바탕으로 **국산 빅데이터 분석 시스템**을 자체 개발하였고, **통계적 알고리즘, 검색 엔진** 및 **데이터 분석기법**을 활용하여 국내 최초로 **개인화 분석 전문 패키지**를 개발하였습니다.
- 최근에는 **Machine Learning** 및 **AI 기반의 업무 개선, 자동화 프로세스**를 구축하는 사업을 수행하여 장차 다가올 본격적 **AI시대를 준비**하게 되었습니다.
- 2024년 현재 **MS Azure AI 제품**과 IBM watsonx 제품의 **LLM/생성형AI 솔루션**, ILOG CPLEX **최적화엔진**을 기반으로 여러가지 다양한 프로젝트를 수행, 솔루션 공급에 힘쓰고 있으며 **Cloud-native 방법론**을 적용하고 있습니다.

KEY

일반현황

회사명	(주)스펠릭스	대표자	유 영 보
사업의 종류	자연어 및 AI 처리 전문, 빅데이터 솔루션 및 컨설팅 서비스, 클라우드 기반 빅데이터 분석 시스템 구축/운영		
주소	07281 서울특별시 영등포구 선유로9길 10, 1601호 (문래동6가, 문래 SK V1센터)		
연락처	전화 : 02-6203-6228 FAX : 02-6203-6229		
회사설립	2005년 12월		
해당부문 종사기간	2005년 12월 ~ 2023년 11월 현재 (17년 11개월)		

“AI/ML 전문가 및 관련 플랫폼 구축 전문가 다수 보유, 신기술 학습조직 운영”

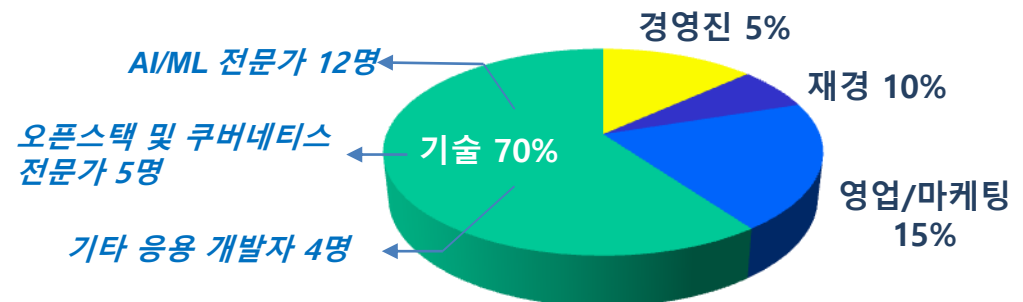


❖ 인원 현황 (계약직 포함)



경영진	영업/마케팅	컨설팅	기술
1명	3명	4명	17명

❖ 엔지니어 등급

특급	고급	중급	초급
5명	5명	4명	3명



연혁 및 최근 실적

기간	사업내용	고객사
2024.12 – 2025.03	개방형 sLLM 기반 건설업무 특화 언어모델 개발	 대우건설
2024.08 – 2025.02	생성형 AI 플랫폼 및 학습Data 구축	 한국남부발전(주)
2022.07 – 2023.03	AI 기반 SK하이닉스 자동문서분류 시스템 구축	 SK 하이닉스
2022.07 – 2022.12	서울대학교병원 통합CDW 구축 (진료정보 정형화 NLP 시스템 구축)	 SNUH
2022.07 – 2023.02	삼성화재 AI-QA시스템 구축	 SAMSUNG 삼성화재
2022.06 – 2022.12	포스코 AI를 활용한 돌발유형 자동분류 지원시스템 구축	 posco
2021.10 – 2022.06	ML기반 Bics 구축을 위한 AI플랫폼 고도화 구축	 * KB 국민은행
2022.02 – 2022.12	PLM 문제점 데이터 분석 서비스 구축	 SAMSUNG 삼성전자
2021.03 – 2021.09	삼성전자 DS Cloud Container 서비스 내재화	 SAMSUNG 삼성전자
2020.07 – 2021.06	현대미포조선 생산PLM 1단계 구축	 ▲ 현대미포조선
2020.07 – 2021.05	DB손보 스마트 컨텍센터 구축	 SAMSUNG 삼성SDS
2020.07 ~ 2021.03	COE 사업 솔루션 구축 계약	 SAMSUNG 삼성SDS
2020.06 ~ 2021.01	KB국민은행 영업점 이동배치 자동화시스템 고도화	 * KB 국민은행
2020.08 – 2020.12	LG생활건강 이상징후탐지시스템 구축 (오픈시프트/Kubernetes 기반 AI플랫폼)	 LG 생활건강
2019.07 – 2020.03	KB 국민은행 정기인사 이동배치 최적화 모델링 및 알고리즘 개발	 * KB 국민은행
2019.05 – 2019.09	롯데케미칼 MaaS	 LOTTE CHEMICAL

저작권 및 특허



저작권 등록 프로그램(13종) + 특허출원(5종)

개체명
인식(Named -
Entity
Recognition)의
정확성 제고를
위한 후처리
과정을 제공하기
위한
방법 및 이를
사용한 서버

특허증
CERTIFICATE OF PATENT

특 허	제 10-2153127 호
출원번호	제 10-2018-0174277 호
출원일	2018년 12월 31일
등록일	2020년 09월 01일

발명의 명칭 Title of the Invention
개체명 인식(Named-Entity Recognition)의 정확성 제고를 위한 후처리 과정을 제공하기 위한 방법 및 이를 사용한 서버

특허권자 Patentee
(주) 스펠릭스(110111-*****)
서울특별시 영등포구 선유로9길 10,1601호 (문래동6가, 문래SKV1center)

발명자 Inventor
등록사항란에 기재

위의 발명은 「특허법」에 따라 특허등록원부에 등록되었음을 증명합니다.
This is to certify that, in accordance with the Patent Act, a patent for the invention has been registered at the Korean Intellectual Property Office.

2020년 09월 01일

특허청장
COMMISSIONER,
KOREAN INTELLECTUAL PROPERTY OFFICE
김 용래

특허청
Korean Intellectual Property Office

복수의
범죄구성요소를
기초로 하여 관련
자료를 검색하기
위한 사용자
인터페이스를
제공하는 방법 및
이를 이용한 서버

특허증
CERTIFICATE OF PATENT

특 허	제 10-2681457 호
출원번호	제 10-2022-0166133 호
출원일	2022년 12월 01일
등록일	2024년 07월 01일

발명의 명칭 Title of the Invention
복수의 범죄구성요소를 기초로 하여 관련 자료를 검색하기 위한 사용자 인터페이스를 제공하는 방법 및 이를 이용한 서버

특허권자 Patentee
(주) 스펠릭스(110111-*****)
서울특별시 영등포구 선유로9길 10,1601호 (문래동6가, 문래SKV1center)

발명자 Inventor
등록사항란에 기재

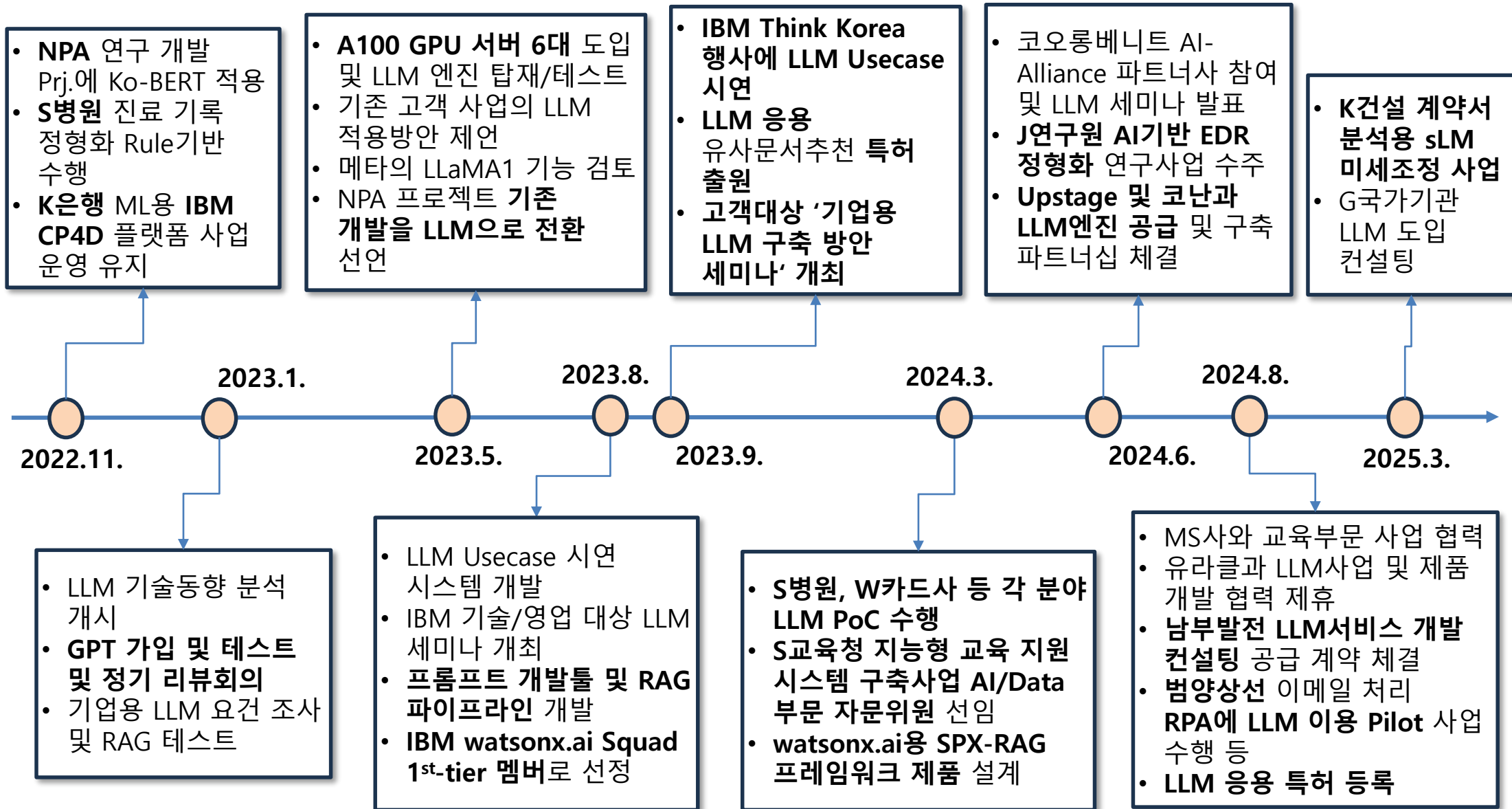
위의 발명은 「특허법」에 따라 특허원부에 등록되었음을 증명합니다.
This is to certify that, in accordance with the Patent Act, a patent for the invention has been registered at the Korean Intellectual Property Office.

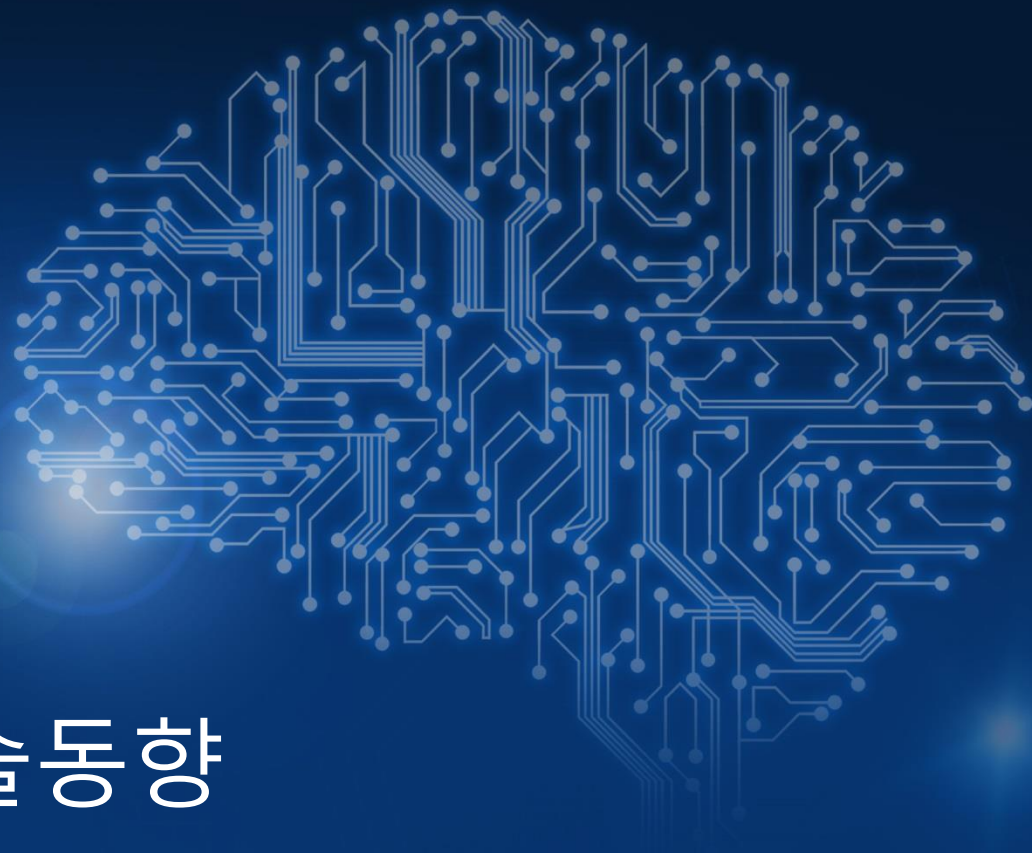
2024년 07월 01일

특허청장
COMMISSIONER,
KOREAN INTELLECTUAL PROPERTY OFFICE
김 완기

특허청
Korean Intellectual Property Office

LLM에의 여정





AI 기술동향



AGI(Artificial General Intelligence, 일반인공지능)에의 도전 - 1

OpenAI의 o1급 모델을 이용하여 인간처럼 언어를 이해하고 추론하는 과정을 거쳐 교육이나 명령없이도 문제를 해결하는 AGI로의 진입에 대한 단초를 보여주었습니다.



AGI(Artificial General Intelligence, 일반인공지능)에의 도전 - 2

Strawberry라고 불리우는 **o1 모델**(2024. 12 출시)은 기존 LLM이 못 보여주었던 복잡한 추론을 해결할 수 있는 모델로 **CoT**라고 하는 **깊은, 단계별 사고과정을 구현한** 모델로 **o3와 o4-mini**(2025. 02 출시)로 더욱 진보된 모델을 제공하고 있습니다.

Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck Varun Chandrasekaran Ronen Eldan Johannes Gehrke
Eric Horvitz Ece Kamar Peter Lee Yin Tat Lee Yuanzhi Li Scott Lundberg
Harsha Nori Hamid Palangi Marco Tulio Ribeiro Yi Zhang

Microsoft Research

Abstract

Artificial intelligence (AI) researchers have been developing and refining large language models (LLMs) that exhibit remarkable capabilities across a variety of domains and tasks, challenging our understanding of learning and cognition. The latest model developed by OpenAI, GPT-4 [Ope23], was trained using an unprecedented scale of compute and data. In this paper, we report on our investigation of an early version of GPT-4, when it was still in active development by OpenAI. We contend that (this early version of) GPT-4 is part of a new cohort of LLMs (along with ChatGPT and Google's PaLM for example) that exhibit more general intelligence than previous AI models. We discuss the rising capabilities and implications of these models. We demonstrate that, beyond its mastery of language, GPT-4 can solve novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting. Moreover, in all of these tasks, GPT-4's performance is strikingly close to human-level performance, and often vastly surpasses prior models such as ChatGPT. Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system. In our exploration of GPT-4, we put special emphasis on discovering its limitations, and we discuss the challenges ahead for advancing towards deeper and more comprehensive versions of AGI, including the possible need for pursuing a new paradigm that moves beyond next-word prediction. We conclude with reflections on societal influences of the recent technological leap and future research directions.

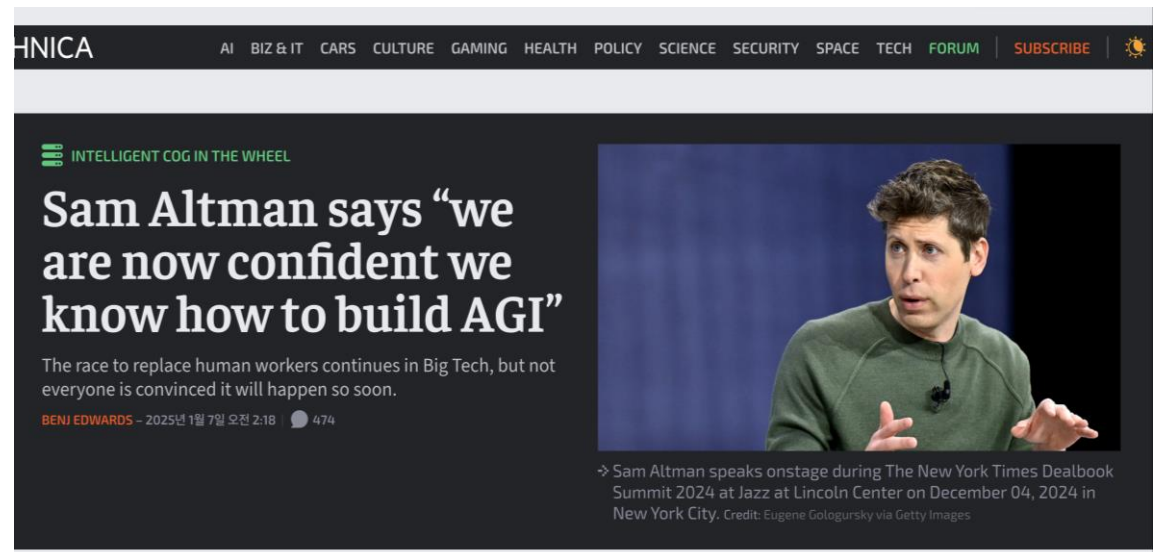
Contents

1	Introduction	4
1.1	Our approach to studying GPT-4's intelligence	6
1.2	Organization of our demonstration	8
2	Multimodal and interdisciplinary composition	13
2.1	Integrative ability	13
2.2	Vision	16
2.2.1	Image generation beyond memorization	16
2.2.2	Image generation following detailed instructions (à la Dall-E)	17
2.2.3	Possible application in sketch generation	18
2.3	Music	19

GPT4는 초기버전의 AGI

"우리는 이제 AGI를 만드는 방법을 알고 있다고 확신한다."

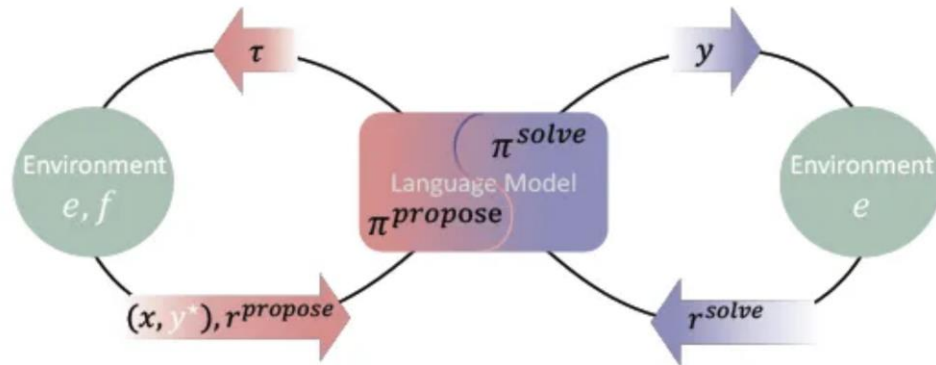
- OpenAI CEO, 샘 알트만 2025. 01. 05. 본인 페이스북



주목할 만한 기술적 성과 1 – Self Learning/Discovery

이제 AI는 인간이 주는 데이터 없이 처음부터 추론을 스스로 학습하는 단계의 연구가 활발하게 이루어지고 있으며, 이를 이용한 응용 및 서비스가 실제로 적용되는 단계로까지 발전하고 있어서 AGI로 가는 계단을 착실히 오르고 있음을 알 수 있습니다.

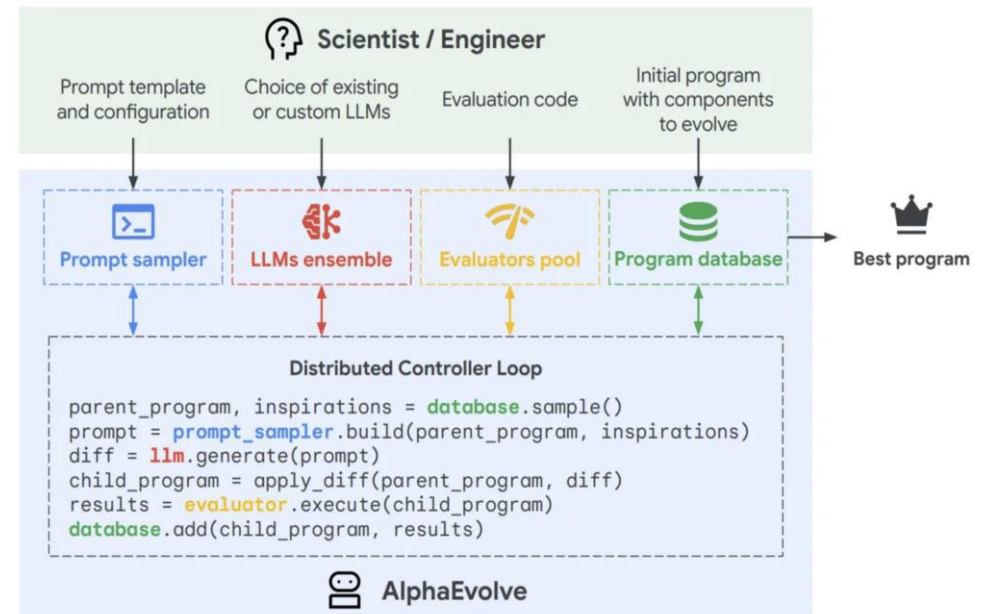
Absolute Zero Reasoner(AZR) : 외부 데이터 없이 강화된 자가학습을 통해 모델이 복잡한 추론을 스스로 학습



작업절차 : 지속적 자기개선 루프를 실행

1. AI가 작업을 제안
2. 환경은 작업의 타당성을 검증하고 학습 가능성에 대한 피드백을 제공
3. AI가 과제를 해결하려고 시도
4. 환경은 솔루션을 검증하고 정확성에 대한 피드백을 제공
5. AI는 두가지 유형의 피드백을 기반으로 스스로 업데이트하여 유용한 작업을 제안하고 이를 해결하는 능력이 향상

AlphaEvolve : 대규모 언어 모델을 올바르게 사용하면, 매우 정확한 의미에서, 증명 가능한 새로운 알고리즘 획득



- ➔ 4x4 행렬 곱셈에 대해 1969년 Strassen의 알고리즘(49번의 스칼라 곱셈)을 처음으로 능가하는 알고리즘(48번 곱셈)을 발견
- ➔ 수세기 동안의 기하학적 난제인 "키스 숫자 문제(겹치지 않는 단위 구가 중앙 구에 동시에 몇 개나 닿을 수 있는지 알아내는 문제)"를 11차원에서 593개의 구로 구성된 배열을 찾아냈으며, 이는 기존 기록인 592개를 넘어섬

주목할 만한 기술적 성과 2 - Reasoning

여기서의 추론(Reasoning)은 일반적인 LLM이 보여주고 있는 기본적 추론 능력을 뛰어 넘어서 퍼즐, 수수께끼, 수학적 증명 등과 같은 더 복잡한 추론의 능력을 말합니다.

작동원리 : Scale of time

If a train is moving at 60 mph and travels for 3 hours, how far does it go?

The train travels 180 miles.

Plain response

To determine the distance traveled, use the formula:

Distance = Speed × Time

Given that the speed is 60 mph and the time is 3 hours:

Distance = 60 mph × 3 hours = 180 miles

So, the train travels 180 miles.

Response with intermediate reasoning steps

적합한 적용 분야

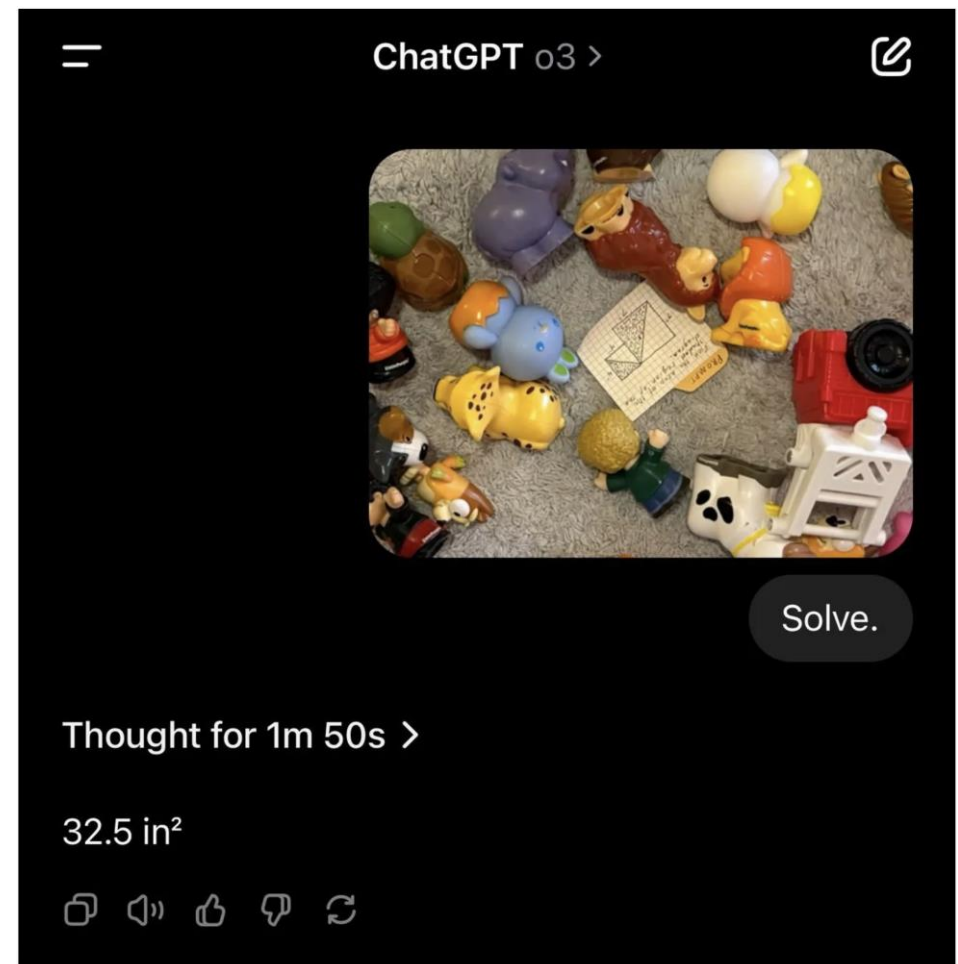
Good at

- + Deductive or inductive reasoning (e.g., riddles, math proofs)
- + Chain-of-thought reasoning (breaking down multi-step problems)
- + Complex decision-making tasks
- + Better generalization to novel problems

Bad at

- Fast and cheap responses (more inference time)
- Knowledge-based tasks (hallucination)
- Simple tasks ("overthinking")

Multi-modal Reasoning 예 : 이미지 사고



주목할 만한 기술적 성과 2 – Reasoning 응용 예

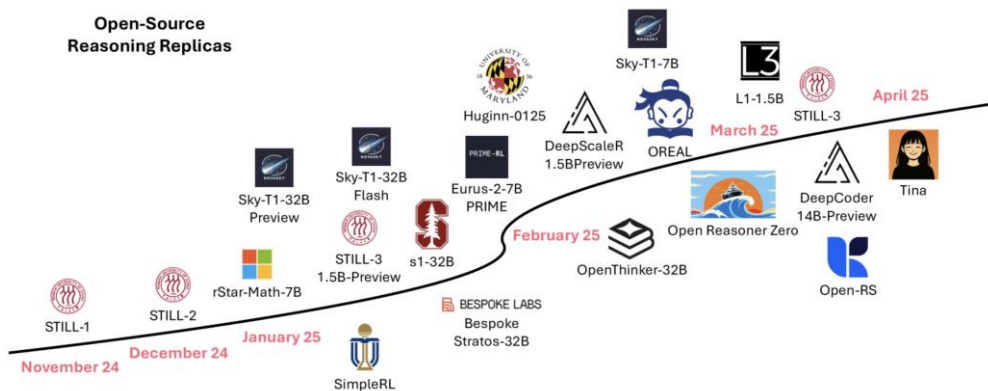
MS사에서는 OpenAI의 o3 모델을 이용하여 조사보고서와 같은 난이도가 높은 작업을 수행할 수 있는 멀티에이전트 시스템을 Copilot에 장착한 서비스를 발표하였습니다.



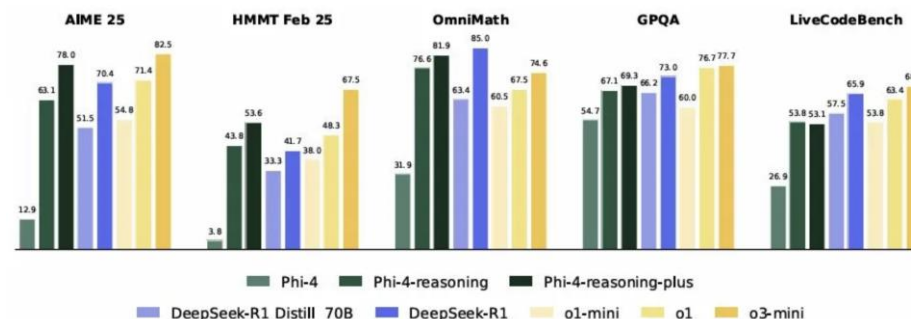
주목할 만한 기술적 성과 2 – 오픈소스 sLLM의 Reasoning 장착

여기서의 추론(Reasoning)은 일반적인 LLM이 보여주고 있는 기본적 추론 능력을 뛰어 넘어서 퍼즐, 수수께끼, 수학적 증명 등과 같은 더 복잡한 추론의 능력을 말합니다.

OpenAI o1 출시이후 오픈소스 추론모델의 등장



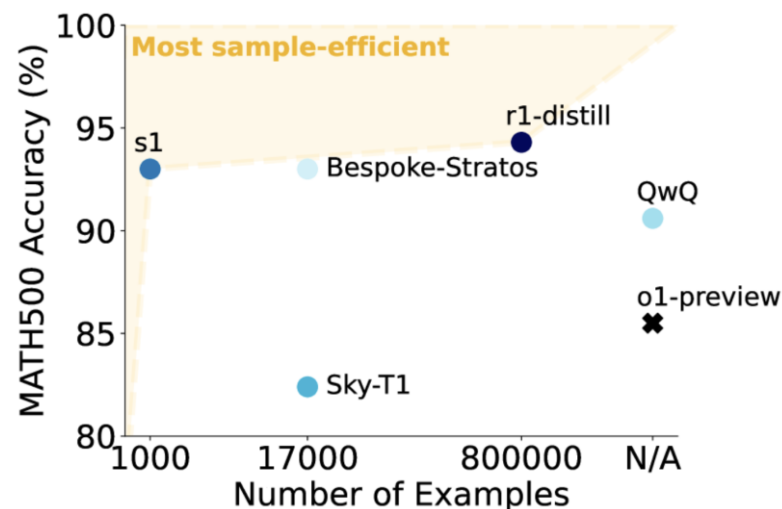
복잡한 추론작업에서 훨씬 더 큰 모델과 경쟁 가능한 모델(Phi-4)



DeepSeek-R1의 추론성능 SOTA 달성

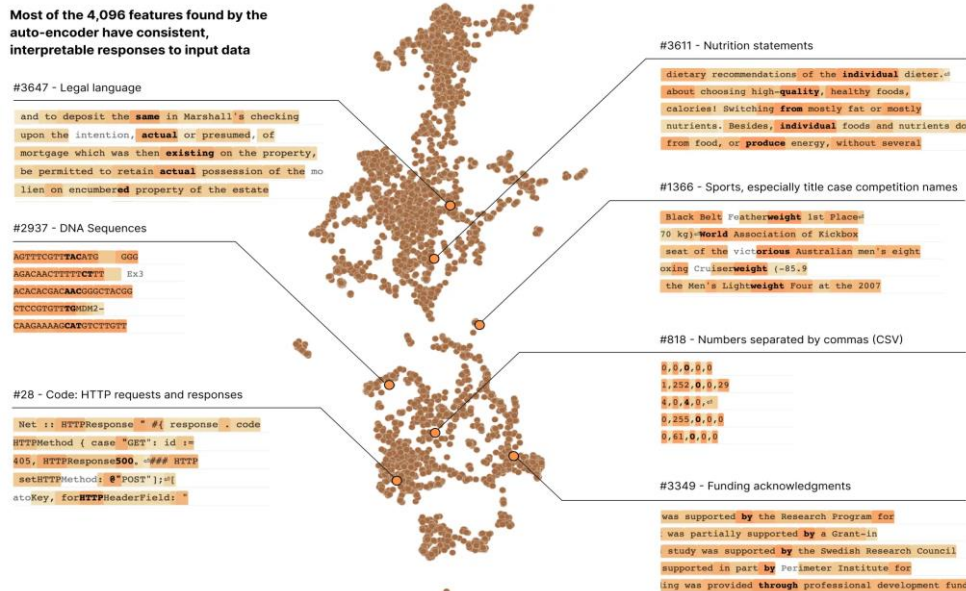
Model	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	pass@1	cons@64	pass@1	pass@1	pass@1	rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717
OpenAI-o1-mini	63.6	80.0	90.0	60.0	53.8	1820
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9	1316
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1691
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1633
DeepSeek-R1-Zero	71.0		95.9	73.3	50.0	1444
DeepSeek-R1	79.8		97.3	71.5	65.9	2029

s1 : 간단한 테스트 시간 스케일링 – 6달러 비용 소요

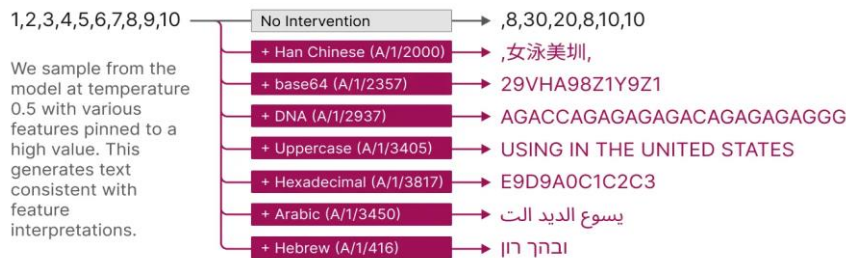


Explainable AI – AI 현미경 개발

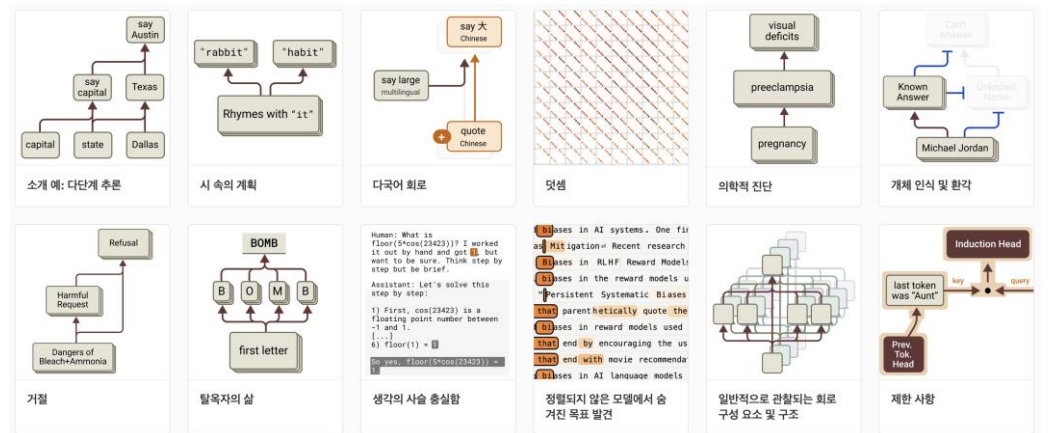
Anthropic은 Claude Sonnet이라는 대규모 언어 모델의 내부 작동 방식을 이해하는 데 큰 진전을 이루어, 수백만 개의 개념이 모델 내부에서 어떻게 표현되는지 확인함으로써 현대적이고 제품 규모 LLM의 내부를 상세히 들여다 본 최초의 사례입니다.



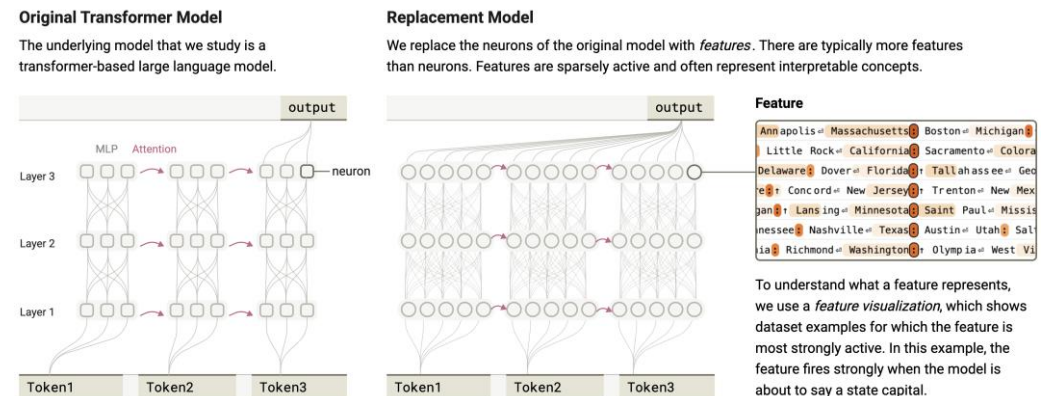
개별 뉴런이 아닌 특정한 개념에 관련이 있는 뉴런들의 집합 (Feature)으로 분해 가능



특정 Feature를 인위적으로 활성화시키면 모델의 동작을 예측 가능한 방식으로 변경 가능



회로 추적 방법론을 사용, 다양한 맥락에서 클로드 3.5 하이쿠(경량모델)가 사용하는 내부 메커니즘 조사



상세 해석 가능한 구성요소(교차계층 트랜스코더, CLT 아키텍처)를 사용하여 원본 모델의 활성화를 재현하는 대체모델 구축

LLM에의 여정

기업 업무 적용에 필요한 생성형AI의 주요 고려사항으로 1) LLM 도입 및 운용 비용, 2) 업무에 적합한 일관성 있는 결과 도출, 3) 복잡한 Workflow 지원 등을 고려해야 합니다.

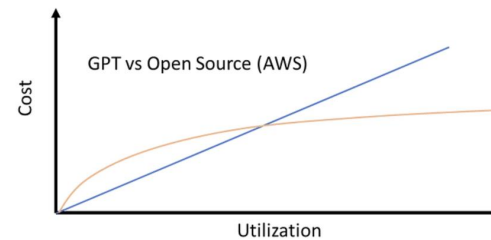
Figure 1: Costs Incurred in Different GenAI Deployment Approaches

	Consume Commercial GenAI apps	Embed GenAI APIs in custom apps	Extend GenAI models via data retrieval	Customize GenAI models via fine-tuning	Build Custom models from scratch
Use case	Coding assistants	Personalized sales content creation	Document search with RAG	Virtual assistant	Medical, insurance or financial services LLMs
Upfront costs	~\$100K to \$200K	~\$750K to \$1M	~\$750K to \$1M	~\$5M to \$6.5M	~\$8M to \$20M
Recurring costs (per user per year)	~\$280 to \$550	~\$790 to \$1.2K	~\$1.3K to \$11K	~\$8K to \$11K	~\$11K to \$21K

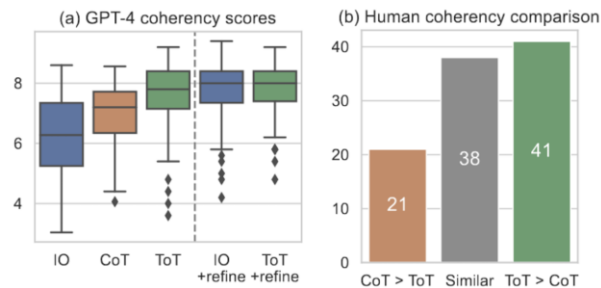
Source: Gartner (July 2024)

AI 배포에는 약
69억 원에서
약 276억 원에
이르는 비용 소요

Cloud/SaaS 서비스



small LM

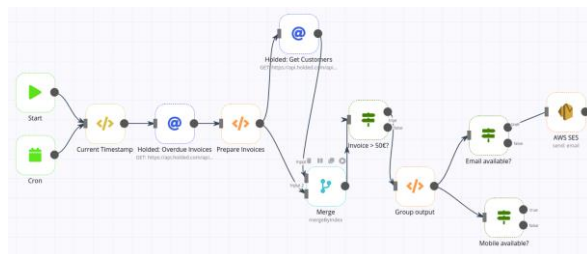
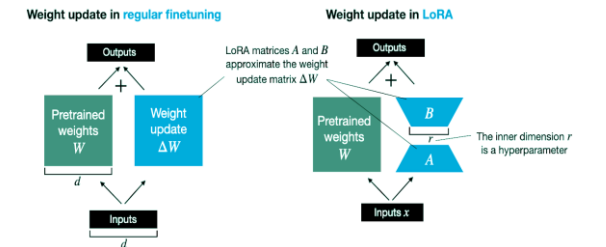


확률적 결과
생성으로
비일관성 상존

검색증강생성(RAG)

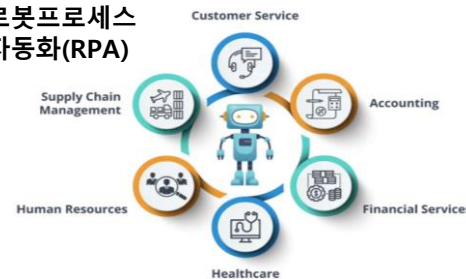


미세조정(Fine-tuning)

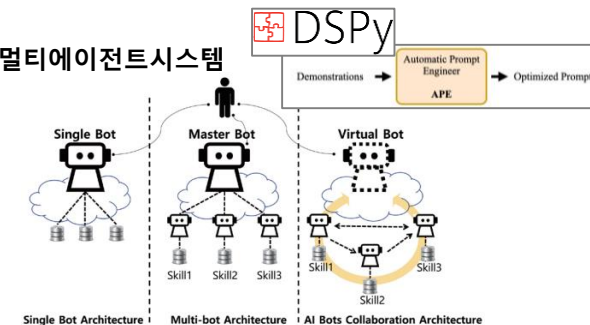


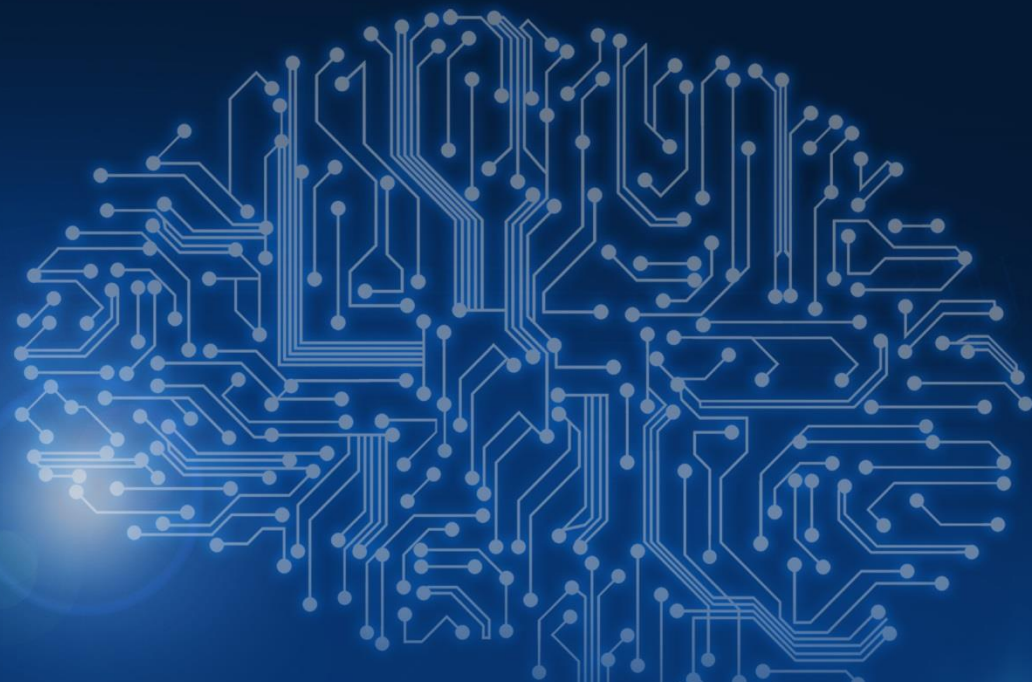
의사결정이
필요한 복잡한
업무처리

로봇프로세스 자동화(RPA)



멀티에이전트시스템





건설산업에서 AI 적용 사례

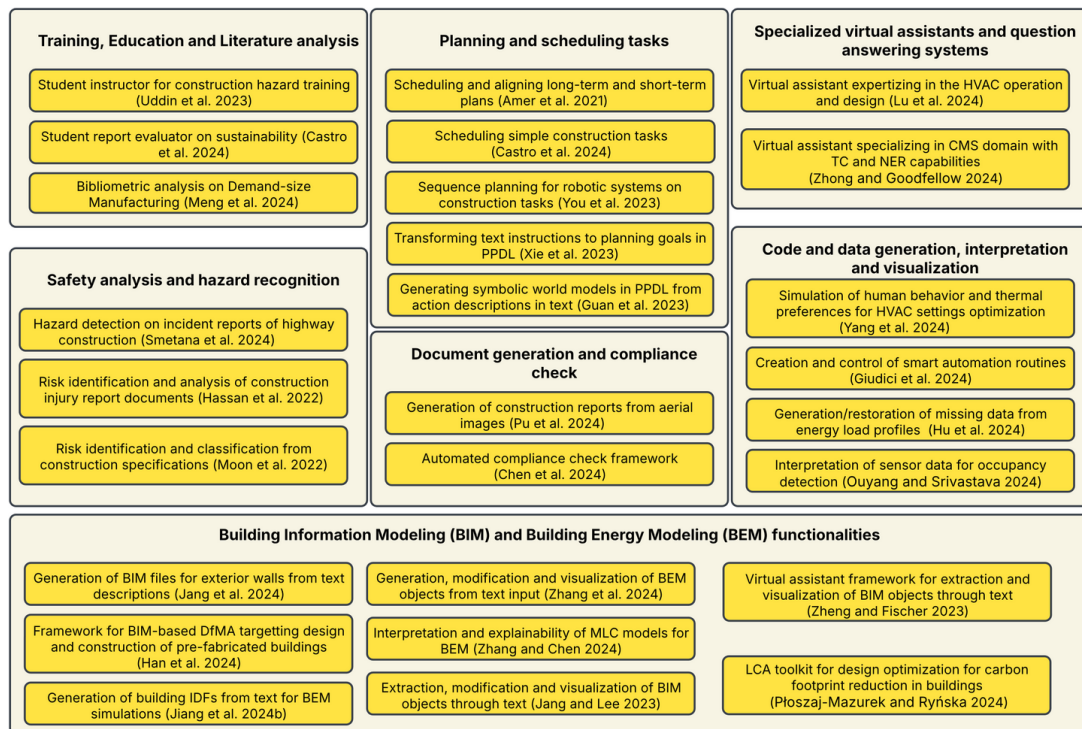


00000

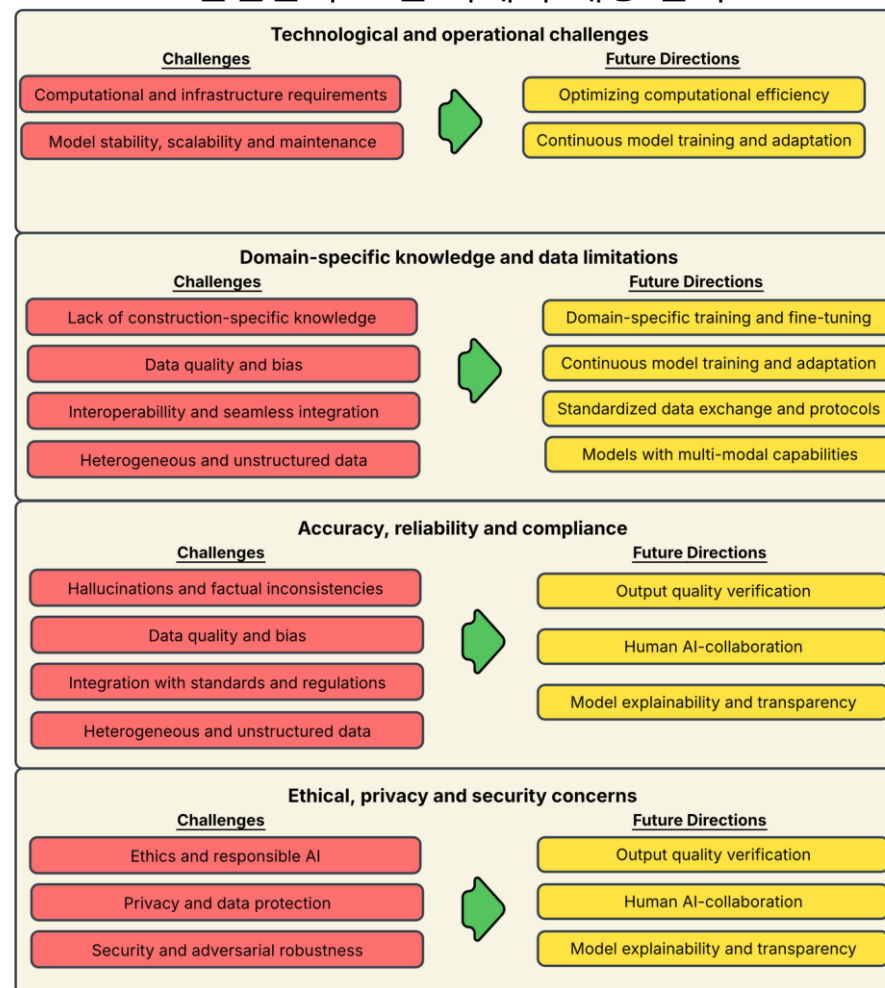
건설분야 LLM 적용 사례 및 도전 과제

최근 건축 엔지니어링 및 시공(AEC) 산업 영역에서 LLM 구현의 기회, 역량 및 잠재력을 탐구한 연구는 제한적이어서 이 분야에 상당한 공백이 존재하지만, 기업에서는 PoC/Pilot을 비롯하여 업무 시스템 구축 등 선도적인 실험과 적용을 진행하고 있습니다.

건설분야 업무별 적용사례



건설분야 도전 과제와 대응 분야

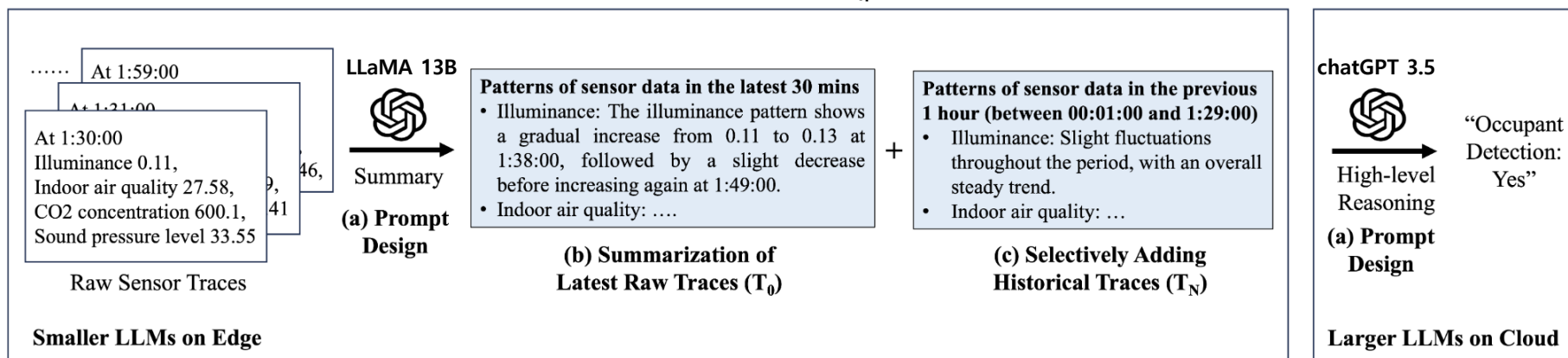


출처 : “A review of LLMs and their applications in the architecture, engineering and construction industry”, 2025
<https://link.springer.com/article/10.1007/s10462-025-11241-7>

건설분야 sLLM 적용 해외 사례 – 시공간 센서 데이터 이용 추론

대규모언어모델(LLM)과 소규모언어모델(sLLM)을 활용하여 고수준 추론 과제를 위한 장기 센서 추적에서 도출된 관측치를 분석하는 방법 및 구현 사례이며 일반화를 위한 미세조정이 필요한 프로젝트입니다.

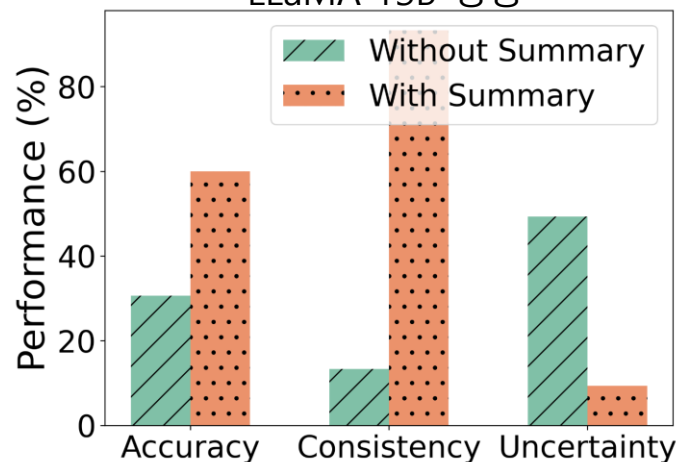
LLMSense 개요



전체 성능 요약(전체과정 각자 수행시)

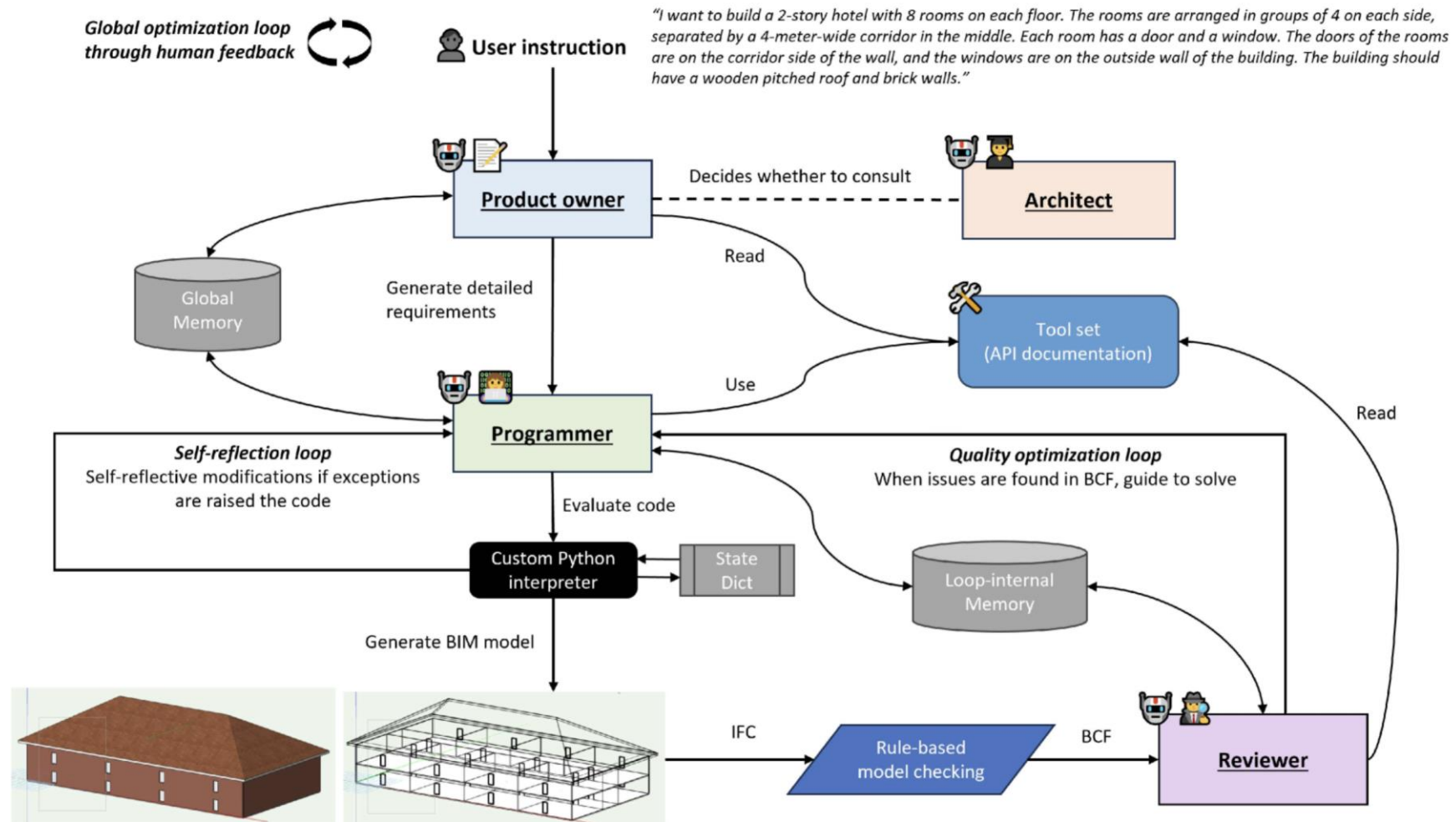
Task	Dementia Diagnosis			Occupant Detection		
	Llama-13B	Llama-70B	ChatGPT3.5	Llama-13B	Llama-70B	ChatGPT3.5
Accuracy	60%	73.3%	76%	60.83%	46.67%	72.08%
Consistency	93.3%	100%	86.7%	81.25%	87.5%	85.42%
Uncertainty	9.3%	0	0	2.08%	0	0

LLaMA 13B 성능



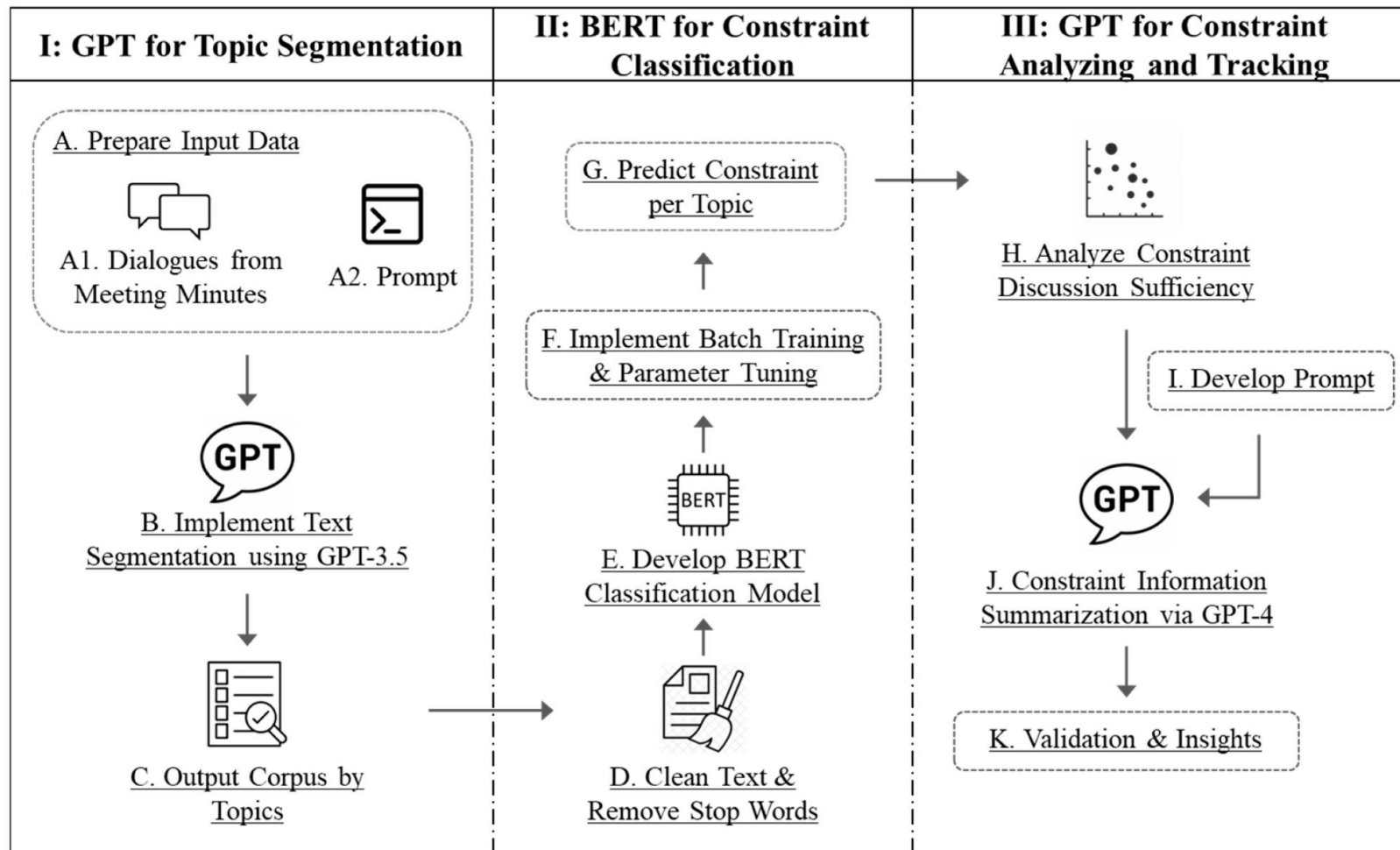
Text2BIM 프레임워크 개발 – 멀티에이전트 적용

LLM 기반 다중 에이전트 프레임워크인 Text2BIM은 설계 의도를 직관적으로 표현하기 위해 자연어 명령어로부터 3D 건물 모델을 생성할 수 있는 LLM 기반 다중 에이전트 프레임워크로서 여러 LLM 에이전트의 협업 및 추론 조율하여 BIM 저작 도구의 API를 호출하는 명령형 코드로 변환하는 시스템입니다.



건설 계획에서 제약조건 명확화 – 멀티 LLM의 활용

현장 계획 회의에서 제약 조건 논의를 식별, 분류, 분석 및 추적하는 자동화된 접근 방식이 부족하다는 심각한 격차 해소를 위해 1) 제약조건의 분류, 2) 제약조건에 대한 논의패턴 파악, 3) 제약조건의 근본원인 추출을 통해 해결책 개발을 지원하는 시스템입니다.



국내 건설분야 AI 적용 사례

AI는 건설업의 전통적인 한계를 극복하며, 설계부터 시공, 유지보수에 이르는 모든 과정에서 생산성과 효율성을 높이며, 스마트 건설 시장은 글로벌 시장조사기관에 따르면, 2023년 약 8조 원 규모로 추산되며, 연평균 10% 이상의 성장률을 기록하여, 설계와 공정, 유지보수를 포함한 모든 단계에서 새로운 가능성을 열어가고 있습니다.

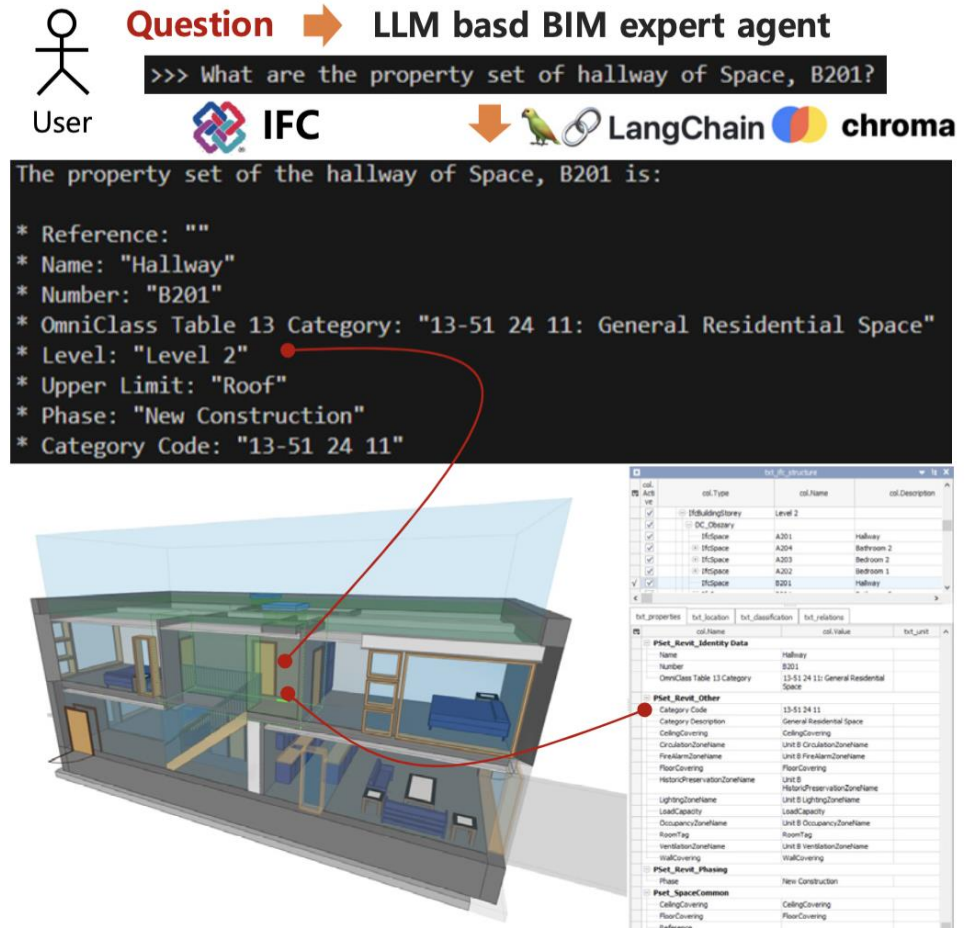
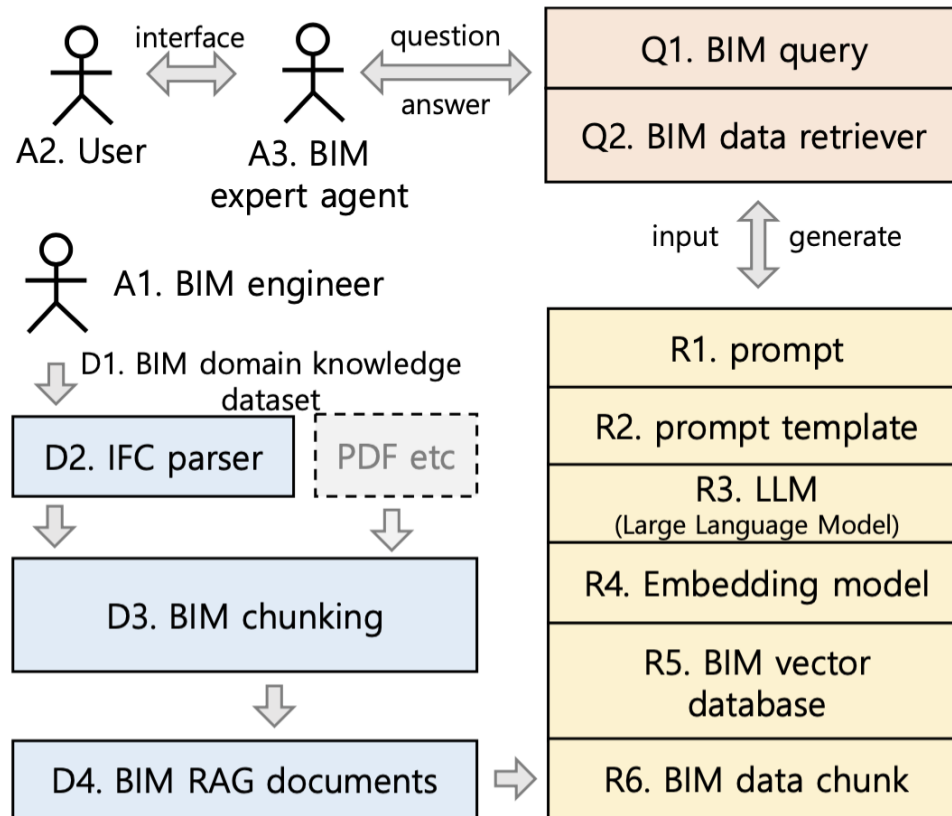
국내 건설사들의 AI 도입 사례

H사	세계 최초 플랜트 및 건설 특화 LLM 개발	<ul style="list-style-type: none"> AI 연구개발 스타트업 J사와 협력하여 플랜트 및 건설 분야에 특화된 거대 언어모델(LLM)을 개발 입찰안내서(ITB) 항목을 비교 분석하고 검토할 수 있는 서비스에 적용 향후 각종 보고서 작성, 리스크 분석, 임직원 교육 등에 활용 전망
G사	다국적 근로자를 위한 AI 기반 번역 프로그램 도입	<ul style="list-style-type: none"> 작업자와 관리자 간의 실시간 소통 문제를 해결하기 위해 개발 음성 입력을 다양한 언어로 실시간 번역하는 기능을 제공 특히, 공사 일정이나 안전 지침 등 중요한 정보 전달에서 언어 장벽을 낮춰, 작업 효율성과 안전성을 동시에 향상
D사	자율주행 드론과 AI를 활용한 하자 점검 시스템 구축	<ul style="list-style-type: none"> 건물 외벽의 하자점검 위해 자율주행 드론과 AI를 결합한 프로그램 개발 지금까지 50개 이상 아파트 건설 현장에 적용 점검 과정의 안전성을 높이고, 시간과 비용을 절감하며, 보다 정확한 하자 관리를 가능
L사	AI 기반의 안전상황센터와 위험성 평가 시스템 도입	<ul style="list-style-type: none"> AI 시스템을 연계한 '안전상황센터'를 개관하여 운영 중 다양한 위험 요소를 실시간으로 모니터링하고, 데이터를 기반으로 빠르게 대응할 수 있는 체계 마련 작업 환경의 위험성을 사전에 분석하여 안정성 제고
P사	AI 기반 빅데이터 분석으로 시장 및 자재 관리 혁신	<ul style="list-style-type: none"> 전국 219개 시·군·구의 10년간 주택 수급 상황, 가격, 거래량 등을 빅데이터화하여 AI 분석으로 국내 주택시장 현황 효과적으로 파악 AI를 통해 철근과 레미콘 등 주요 건설 자재의 미래 가격을 예측 비용 관리의 정확성을 제고
K사	검색증강생성 기반 프로젝트 문서 분석 지원시스템 개발	<ul style="list-style-type: none"> 검색증강생성 기능과 GPT를 이용하여 문서분석 시스템 개발 AI를 통해 문서 내용과 주요사항, 요약 정보를 사용자 질문에 따라 제공 LLM 기반의 내용 분석 및 검토 의견을 제공하여 업무 생산성을 향상

출처 : SK C&C 참조 및 수정

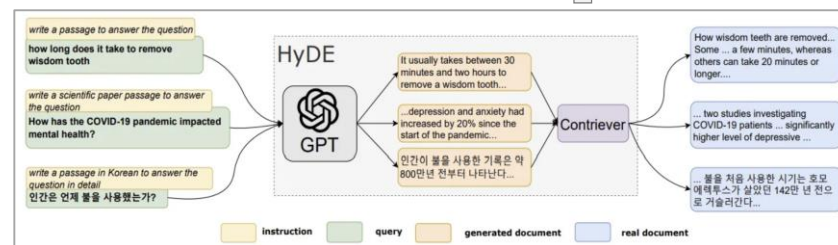
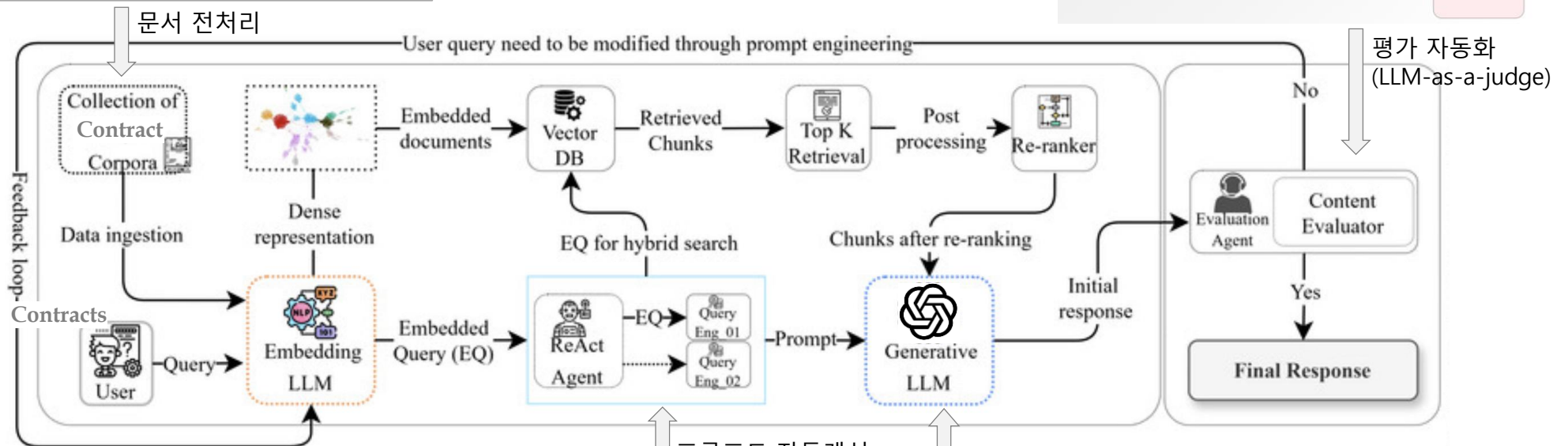
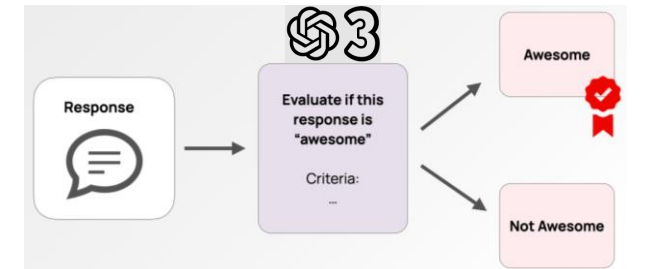
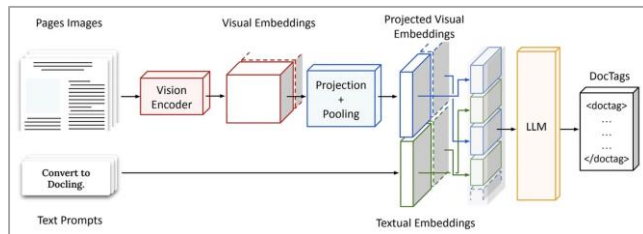
LLM과 RAG 기반 BIM 지식 전문가 에이전트 연구 - BIM 데이터 증강 처리 기술

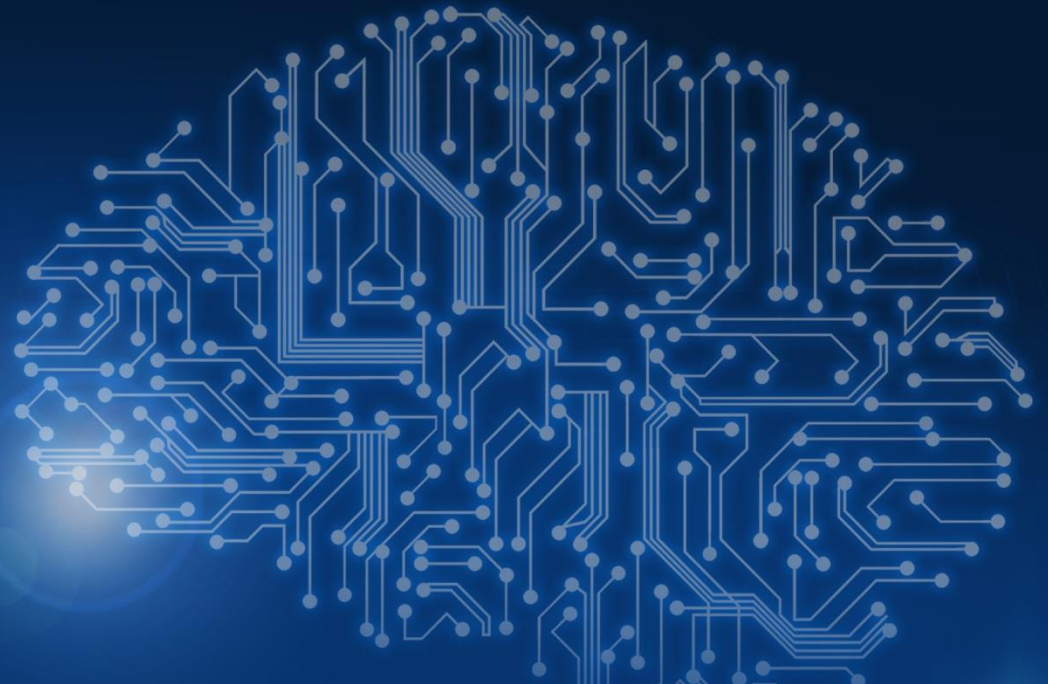
BIM 데이터를 효과적으로 검색하고, 필요한 정보를 생성하는 방법은 부족한 상황을 극복하기 위해 LLM과 RAG 기술을 적용하여 BIM 데이터의 정보를 검색하고, BIM 지식 전문가 에이전트를 개발하는 효과적인 방법에 대한 연구와 프로토타입 개발을 하였습니다.



계약서 분석을 통한 주요 내용 식별과 요약

굉장히 복잡하고 방대한 계약서의 내용을 필요한 부분의 내용을 분석하여 확인하거나 검토해야 할 사항과 위험도를 평가하여 프로젝트를 원활하게 수행할 수 있도록 LLM을 이용하여 지원하는 시스템을 개발하였습니다.





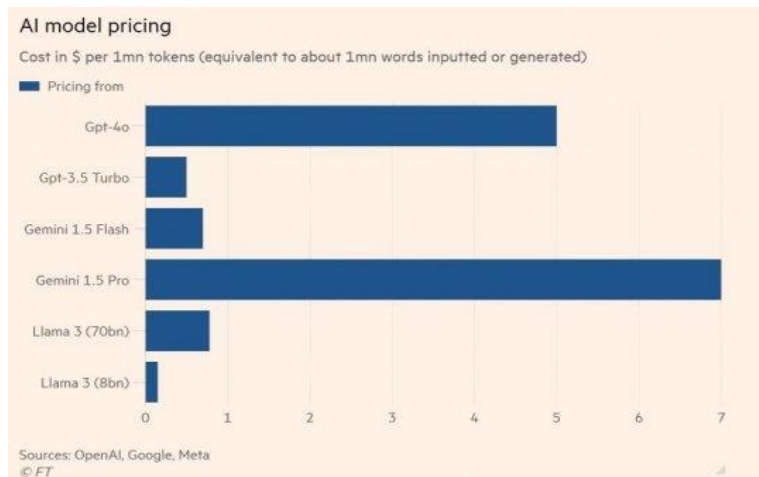
Case Study :
건설업무 맞춤형 sLLM 개발

What & Why On-premise for sLLM?

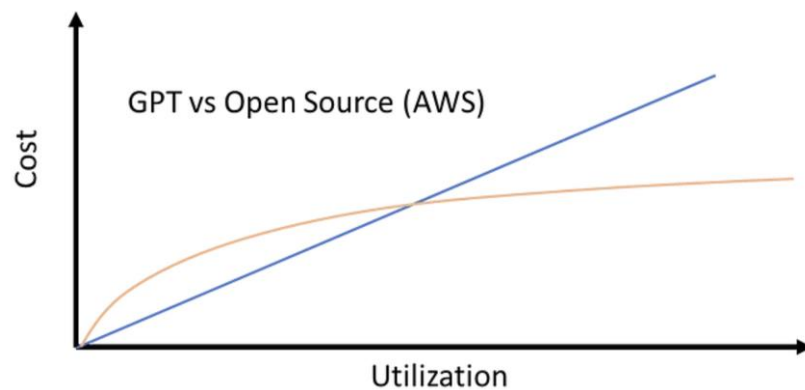
자체 구축(On-premise)형 sLLM은 1) 정보보안의 문제 해결, 2) 업무 맞춤형 모델의 구축, 3) 기업의 사업환경과 전략에 따른 제어력 향상 등에 따른 총소유비용(Total Cost of Ownership)이 가장 좋은 방법 중의 하나입니다.

독자구축형 sLLM의 장점

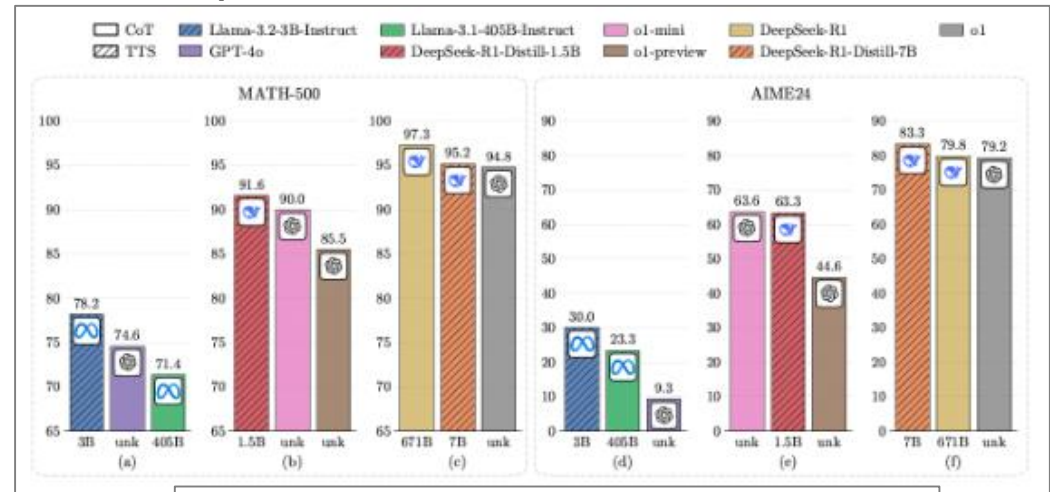
단위토큰 비용



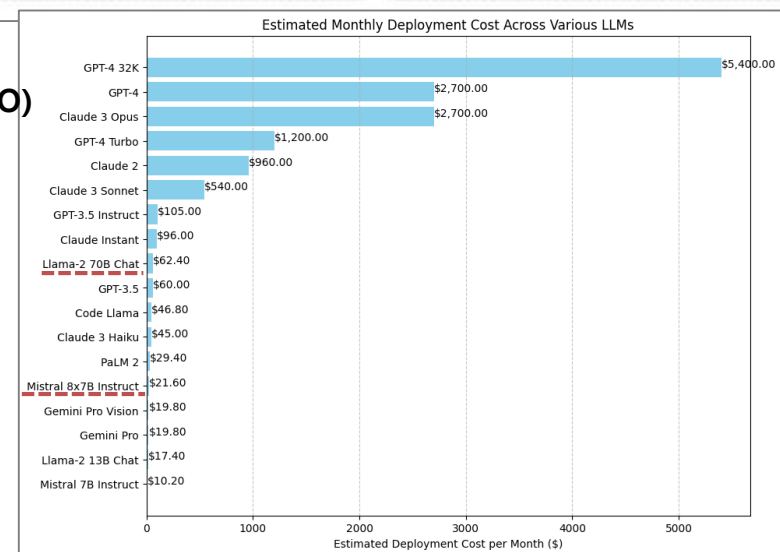
SaaS 대 Local LLM 운영비용



성능평가(DeepSeek 경우)



총소유비용(TCO)



What & Why 업무 특화 sLLM? (맞춤형, 미세조정, Fine-tuning)

기존 sLLM이 학습한 내용(사전학습, Pre-trained)에 특정 분야의 특정 업무에 필요한 지식을 학습시켜서 해당 업무에 적합한 질문과 답변을 확보할 수 있도록 수행하는 학습의 일종입니다.

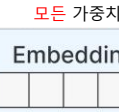
sLLM의 활용 방안 및 미세조정

1) 프롬프트 개발



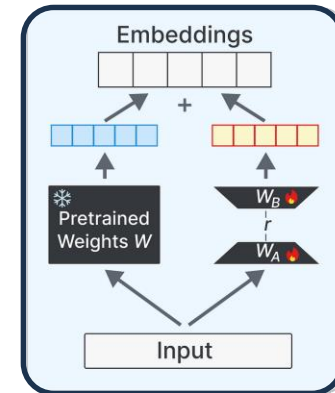
3) 미세조정(Fine-tuning)

Regular Fine-Tuning



Low-Rank Adaptation

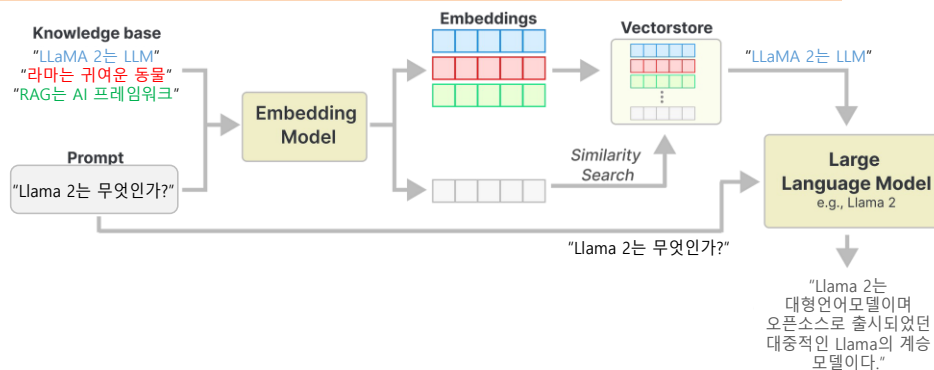
가중치의 작은 표현 수정



사고 기반 방식

- 비즈니스 사용 사례에 중요한 특정 작업에 집중하여 비용을 절감하거나 지연시간을 최소화하고 싶을 때
- 모든 데이터를 비공개로 유지하여야 할 때
- 특정 작업에 대한 답변의 품질을 개선하고 싶을 때

2) RAG(Retrieval Augmented Generation, 검색증강생성)



1), 2), 3) 방법의 비교

1 Prompt Engineering

Natural Language

- Quick Iteration
- Requires no Training
- (Sometimes) No coding

Few-shot

"Here are a few examples..."

Chain-of-Thought

"Solve this step-by-step..."

ReAct

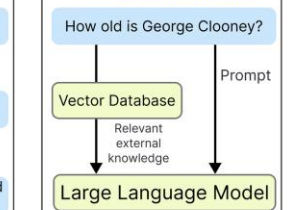
"Create thoughts, actions, and observations..."

2 Retrieval Augmented Generation (RAG)

External Knowledge Base

- Query Database
- Requires no Training
- Allows for Fact Checking

Using vector database:

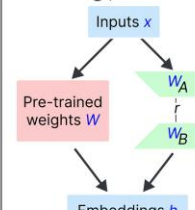


3 Fine-Tuning

PEFT

- Best Performance
- Requires Training
- Quality Dataset Necessary

e.g., LoRA:



Complexity
Quality

프로젝트 개요

GPT4-o를 이용한 문서 분석 업무의 적용을 성공적으로 개시함에 따라, 향후 다양한 업무 환경에의 확대 적용 가능성을 고려하여 특정업무에 적합한 경량 LLM(sLLM)의 적용 가능성과 타당성을 검토하여 생성형AI 적용 전략 선택의 다양성을 확보하고자 합니다.

● 건설분야 업무 관련 맞춤형 sLLM(sLM*) 구축 및 업무 적용 타당성 검토

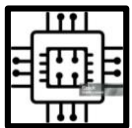
프로젝트 명	개방형 sLLM 기반 건설업무 특화 언어모델 개발 및 성능 평가
프로젝트 매니저	OOO 상무
수행기간	2024. 12. 19 ~ 2025. 03. 24.
발주처	건설 고객사
주요 수행내용	<ul style="list-style-type: none">● sLLM을 이용하여 건설분야 특정업무 맞춤형(Task-specific) 언어모델의 구축과 적용 타당성을 검증<ul style="list-style-type: none">- 건설 관련 업무 분야에 대한 학습 정도를 파악하여 적합한 sLLM을 선정- 안전관리, 보안 등 적합한 특화업무에 대한 맞춤형 학습(Fine-tuning)- 선정 특화업무 분야에 적용한 결과에 대한 평가- 건설분야 특정업무 적용 타당성 분석 및 보고서 제출

* sLM : small Language Model의 약어로 small Large Language Model(sLLM)로도 불림

프로젝트 수행 - 주요 내용

업무 적용성 평가를 위한 대상 응용 업무를 정의하고, 개방형 sLLM의 검토 및 비교 분석을 통해 적용 대상 모델의 선정, 분야특화 학습을 위해 학습데이터 구축 및 평가지표의 설정, 특정 업무 적용 학습(Fine-tuning), 학습 결과의 평가 및 재학습, 업무 적용성 평가 및 보고서 작성이 주요 업무 내용입니다.

주요 수행 내용



sLLM의 업무 적용성 평가를 위한 **대상 응용서비스** 정의 및 선정

적용할 대상 업무에 대한 응용서비스를 텍스트 분석 및 처리 분야, 이미지 인식 등의 비전 분야, 다양한 미디어를 처리하는 멀티모달(Multi-modal) 분야 등으로 구분하고 세부 요구사항을 정의한 후 적용 서비스 후보를 선정



개방형 **sLLM**의 기본 성능 분석 및 적용 대상 **모델** 선정

10B 전후의 파라미터를 보유한 개방형 sLLM의 기본 성능을 비교하여 적용하려는 업무 분야의 학습 정도를 분석하고 업무에 적합한 개방형 sLLM을 선정



학습을 위한 **데이터**의 구축 및 **평가 지표** 정의

특정업무 적용 학습, 즉 미세조정을 위한 학습데이터의 구축과 함께 학습 평가를 위한 기본 정량지표를 선정하고, 실제 업무에 적용한 결과에 대한 정성평가(기존 수행 업무의 결과와 비교 등) 지표를 정의



특정업무 **맞춤형 학습(Fine-tuning)** 및 업무 적용 **평가**

준비된 학습데이터를 통한 업무특화 학습을 수행하고, 그 결과를 평가지표에 근거하여 평가를 수행하며, 결과에 따른 재학습을 수행하고, 일정 수준의 학습된 모델을 특정업무에 적용하여 그 결과를 정성평가(소규모 현업 평가)하고 필요한 평가 프로세스에 대한 자동화*를 모색

**자동화 : 하이퍼파라미터 자동화 방안 가능성 조사 필요*



업무 적용 타당성 분석 및 보고서 작성

특정업무에 적용한 결과를 분석하여 그 타당성을 평가하고 장단점 및 품질개선* 방안과 적용의 한계 등을 분석한 보고서를 작성하여 제출

**품질개선 방안 : 다수의 sLLM의 결과를 병합하거나, 모델을 순차 또는 병렬 처리를 하여 품질을 고도화하는 방안 연구*

sLLM의 업무 적용성 평가를 위한 대상 응용 서비스 정의 및 선정

미세조정 적용할 대상 업무에 대한 응용서비스를 연구원 담당자들과 온라인 및 오프라인 브레인스토밍을 거쳐 총 22개 업무 시나리오와 기술 난이도, 데이터 준비도 등을 기준으로 최상의 난이도로 판단한 문서 내용 분석 업무를 선정하였으며, 미세조정을 적용하지 않은 기본 이미지 인식 성능을 테스트하는 것으로 추가 선정하였습니다.

● 미세조정 대상 업무 및 응용 서비스 일부 목록과 대상 선정

SN	업무 시나리오	최종 의견	미세조정	모델 사이즈	우선순위	기술난이도	데이터 준비도 /준비 난이도
3	문서의 일부 문단을 던지고, 그 문단에 대해서 어떤 내용인지 분석하고 요약 하는 방식?? (요약을 200자 내외로 한글로 요약?)	LLM은 어차피 요약을 잘하는데, 요약을 할게 아니라, 주요 문장을 추출한다면?? 요약이 아니라, 주요 정보 정확히 추출 하는 문제로 진행, 메타데이터 추출	O	Medium		최상	대상 문서 기반 합성
16	현장 이미지(오픈데이터)를 보고, 위험성평가 위험 원인 분류 방식에 따라, 위험 상황을 분류하고, 가능하다면 필요한 액션을 제시. (2단계)	텍스트 모델 > 이미지 비전 > 멀티모달	O				공개된 건설현장 사진 데이터

개방형 sLLM의 기본 성능 검토 및 적용 대상 모델 선정 3/3

10B 전후의 파라미터를 보유한 개방형 sLLM의 후보를 선정한 후, 기본 성능, 특징, 학습 및 추론 환경, 다국어 지원 여부, 대상 적용 분야 학습 정도 분석 등을 수행하였으며, 프로젝트 개시 후 등장한 DeepSeek 모델도 테스트하였으나 기대 품질이 예상보다 낮아서 제외하였습니다.

개방형 sLLM의 검토 대상 후보 및 선정 모델 (★)

	모델	제조사	파라미터 개수
요약	Gemma2	구글	2B
이미지 인식	LLaMA 3.2	메타	11B Vision
계약 분석	Phi-4	마이크로소프트	14B

학습데이터의 구축 – 문서 분석 업무를 위한 미세조정용

현재 운영 중인 문서 분석 서비스에서 사용하고 있는 대상 문서와 서비스의 결과 데이터를 취합하여 Instruct Set을 구축하되, 고객이 제공하는 정제된 Seed 문서를 이용하여 LLM으로 정답데이터를 증강생성합니다. (이미지 인식 데이터는 공공데이터 사이트에서 제공하는 건설현장 사진 데이터를 이용)

미세조정 학습 데이터의 구축

문서 원문 예시

3.14. All vent lines(including silence, if required) up to atmosphere shall be provided by the Supplier.

3.15. The terminal points of circulating water piping will be located one(1) meter above ground on the pipe, nearby condenser.

3.16. All flanges, gaskets, bolts, nuts and washers for large connections including counter flange at the terminal point are supplier's scope (Counter flange and welding preparation shall be considered Purchaser's piping thickness and the welding preparation shall be provided by supplier).

3.17. The utility connection from Purchaser's to supplier's provision at terminal point shall be only one line for each system and followed by Purchaser at the detail design stage. The number and location of terminal point about the utility connection from supplier's to Purchaser's for each system shall be informed by Purchaser at the detail design stage.

3.18. All chemical dosing and sampling points shall be provided by Supplier. The scope of supply shall be equipped by Supplier. All sampling rack and chemical dosing rack shall be placed at same place for easy sampling.

3.19. All steel parts to be embedded in concrete for supporting or fixing equipment, instruments and pipes of the Supplier's scope, which includes anchor bolts and nuts, flanges, base plates, sole plates, sub-sole plates, shims, sleeves, drains, vent pipe and electrical conduits, etc. shall be provided by the Supplier.

3.112. All internal piping, valves including root valves for instruments, and expansion joints, including hanger and support within the terminal points, shall be provided by the Supplier.

3.113. For fuel gas line, the terminal point connection and preparation shall be in accordance with ASME B16.25 to match with Purchaser's pipe thickness. If it's not possible, transition pieces shall be supplied by Supplier.

3. Fuel gas (FG) supply system

Terminal point of main piping for fuel gas shall be provided from the isolation valve at upstream of fuel gas treatment such as scrubber, etc. The terminal points of vents for fuel gas lines shall be up to the roof level of the building.

There is no terminal points of heating water or water lines for fuel gas heater.

3. Closed cooling water line.

문서 분석 결과 데이터 (정답 데이터)

항목	주요계약내용	계약조항
Advance Payment	<ul style="list-style-type: none"> 금 액: Total Lump Sum Price의 10% 지급 시기: 3% - Commencement Date로부터 45일 내 	GCC 3.1.1.1
Advance Payment Bond	<ul style="list-style-type: none"> 반 환: 4% - Engineering & Procurement의 Master Schedule 승인 시 반환 기간: 매 기성에서 선수금의 5% 공제 	GCC 3.1.1.1
Progress Payment	<ul style="list-style-type: none"> 금 액: 선수금의 10% 유료 기간: Date of the last date of the Acceptance Certificate 또는 선수금 	GCC 3.1.1.2
Retention	<ul style="list-style-type: none"> 지급 방식: Monthly Progressive Payment 지급 조건: 매월 10일 전에 전월 공사 분량 인보이스 제출 지 급 일: 인보이스 수령 후 30일 내 	GCC 3.1.1.2
Retention Bond	<ul style="list-style-type: none"> 금 액: 5% (Retention Bond로 대체 가능) 환급 시기: Acceptance Certificate 발행/ 인보이스 제출 후 60일 내 	GCC 3.1.1.2

정답데이터 작성/생성

```

input
1. 계약 체결일자는 (signing date of the CONTRACT)?

Data (JSON):
{
  "id": 291495,
  "type": "text",
  "source": {
    "doc_nm": "01 EPC_Contract (IPP Plant)",
    "rev": "1",
    "page_no": 1286,
    "pj_no": "11-047",
    "table_name": "contract_elems"
  },
  "input": "### SIGNATORIES\n\nIN WITNESS whereof this Direct Agreement has been executed and is intended to be and is hereby delivered on the date stated at the beginning of this Direct Agreement The Owner",
  "metadata": {
}
}

output
{
  "summaryAnswer": "계약의 체결일자는 문서에 명시적으로 언급되어 있지 않습니다.",
  "descriptiveAnswer": "계약 문서에는 특정 체결일자에 대한 직접적인 언급이 없습니다. 예를 들어, 일부 문서는 '계약의 유효한 날짜로부터' 관련된 일정이나 승인 과정에 대해 설명하고 있으나, 구체적인 체결일자에 대한 언급은 명시되지 않았습니다.",
  "keyKeywords": [
    "Contract Award Date",
    "Effective Date",
    "Document Execution",
    "Legal Compliance"
  ],
  "keyEntities": [
    "Effective Date of December 2008: 일부 절차는 이 날짜를 기준으로 설명됩니다.",
    "Executed Date of Direct Agreement: 문서상 구체적인 날짜에 관한 언급이 없음."
  ]
}

```

고객사 제공 99개 정답셋(Seed)의 샘플

최고수준 LLM을 이용한 데이터 합성(증강)



```

input
"project_id": "20-009E2",
"query": "5. 계약 해지 사유는 (termination reason of the CONTRACT)?",
"chunkedData": [
  {
    "input": "### 5 TERMINATION\n\n5.1 Termination by the Company\n\nThe Company may terminate this Contract at any time upon giving thirty...(중략)"
  }
]

output
{
  "summaryAnswer": "계약의 해지 사유는 두 당사자 모두가 명확히 명시되어 있으며, 가장 일반적인 사유로는 계약 불이행과 행위자의 파산이 포함됩니다.",
  "descriptiveAnswer": "회사 측에서는 계약자가 계약 불이행을 하고 10일 이내에 이를 수정하지 않을 경우, 계약자가 파산하거나 법적 금지로 인해 의무를 수행할 수 없을 경우 계약을 해지할 수 있습니다. 계약자는 회사가 90일 이내에 대금을 지급하지 않거나 잘못된 정보 제공 및 회사의 핵심 운영이 금지될 경우 계약을 해지할 수 있습니다.",
  "keyKeywords": [
    "Termination by Company",
    "Termination by Contractor",
    "Written Notice",
    "Failure to Perform",
    "Insolvency"
  ],
  "keyEntities": [
    "회사 해지 사유: 계약 불이행, 파산",
  ]
}

```

GPT4-o 생성 정답셋 20,000여 건의 샘플

미세조정 학습

모델의 업무적용 평가 결과에 대한 지표 분석, 결과 평가 관련 학습 데이터의 완전성 및 정확성 등 평가, 학습된 sLLM 모델의 응답속도, 시스템 자원 사용량 등 다양한 관점에서 종합적인 연구결과를 평가하고 보고서를 작성합니다.

● Phi-4의 특정 문서 주요내용 추출 및 요약, 구조화 태스크 미세조정 실험






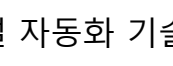
A	B	C	D	E	F	G	H	I	J	K	L
파일명	사용 데이터	동작 여부	학습 시간 (초)	Json Schema	Only Json	summaryAnswer BERTScore	descriptiveAnswer BERTScore	keyKeywords BERTScore	keyEntities BERTScore	BERTScore 평균	LLM Judge
gpt-4o				700 / 700	700 / 700	0.7325	0.7384	0.8536	0.7069	0.7579	92.6
gpt-4o-mini				698 / 700	700 / 700	0.8245	0.7658	0.8206	0.7242	0.7838	91.4
phi-4				696 / 700	696 / 700	0.8051	0.7486	0.785	0.7108	0.7624	80.4
tuning-phi-4-q-e1		H100		698 / 700	697 / 700	0.7996	0.7421	0.812	0.7025	0.7641	93.4
rain_2025-03-13-13-21-54		O	1885.85	700 / 700	699 / 700	0.8155	0.7595	0.7897	0.7182	0.7707	92.6
tuning-phi-4-q-e5		H100		697 / 700	696 / 700	0.7981	0.7414	0.812	0.7021	0.7634	93.2
tuning-phi-4-q-e5-2		H100		697 / 700	696 / 700	0.7985	0.7409	0.8121	0.7009	0.7631	93
tuning-phi-4-q-e10		H100		687 / 700	694 / 700	0.7627	0.7131	0.7997	0.6823	0.7395	88.8
tuning-phi-4-q-e10-6		H100		675 / 700	694 / 700	0.7567	0.7081	0.7901	0.6758	0.7327	89
tuning-phi-4-q-e10-5-16000		H100		699 / 700	700 / 700	0.8039	0.7467	0.7995	0.7028	0.7632	83.0
tuning-phi-4-q-e10-6-16000		H100		697 / 700	694 / 700	0.8015	0.7448	0.7933	0.7015	0.7603	82.97
tuning-test-phi4-e20-16000		H100									
rain_2025-03-24-16-46-16_phi-4_re1	클러스터링 중복 제거 증강데이터	A100		699 / 700	694 / 700	0.818737911	0.770902861	0.832415021	0.72635594	0.7871	81.18
				99 / 99	85 / 99	0.8328	0.7518	0.7857	0.7157	0.7715	66.82

sLLM의 이미지 인식 실험

멀티모달 Phi-4(5.6B)와 LLaMA 3.2 Vision(11B) 모델에 대해 건설현장 이미지에서 위험요소 인식을 테스트 하였고, 한글의 경우에는 결과 품질이 기대에 미치지 못하였으나 영문의 경우에는 좋은 결과를 확인하여, 미세조정을 거치면 더 좋은 결과를 기대할 수 있을 것으로 판단합니다.

● 건설현장 이미지 인식 실험 (일부)

Llama-3.2-11B-Vision-Instruct 이미지 설명 테스트(영어)

이미지	id	Human description (정답)	단순 description	안전 description
	1	빔 모서리에 의한 슬링(스트랩) 손상 및 추락사고 위험	The image depicts a yellow metal bar secured with a green and orange cable tied to a metal grate.	The image depicts a yellow beam wrapped in electrical tape on a metal grate, potentially posing a tripping hazard and requiring immediate attention from site safety management.
	2	파레트를 임시발판으로 사용함에 따른 추락사고 위험	This image depicts a man standing on a metal grate platform, with a wooden pallet and a fan visible beneath him.	The image shows a wooden pallet on top of a metal grate, which poses a risk of falling debris and entrapment for workers below.
	3	노후-파손된 목재 블록 사용으로 인한 중량물 전도 및 낙하 위험	This image shows a blue flatbed trailer loaded with wooden planks, some of which are stacked on top of one another and others that are resting against the trailer's side.	This image shows a large piece of lumber protruding from a pile of wood on a flatbed trailer, posing a potential risk of falling off and causing injury to workers or damage to equipment.
	4	무(無)버팀 굴착 작업으로 인한 토사 붕괴 및 추락사고 위험	This image depicts a large hole in the ground, likely a sinkhole or excavation site, with a person standing at the top and another at the bottom, surrounded by a metal ladder and other tools.	From a construction site safety management perspective, this image depicts a significant hazard: an open and unguarded excavation pit with a steep drop-off, posing a substantial risk of injury or fatality to workers and others in the vicinity.
	5	난간이 없는 높은 구조물 위에서 안전대(랜야드) 미제결 상태로 작업함에 따른 추락사고 위험	The image depicts three workers in hard hats and safety gear, engaged in constructing a metal structure on a metal platform.	The image shows three construction workers in high-visibility vests and hard hats working on a metal structure, highlighting the importance of proper safety gear and protocols in preventing accidents and ensuring a safe work environment.
	6	계단 하부에 방치된 아르곤 가스 실린더(용기)로 인한 전도-충돌 및 가스 누출 사고 위험	This image depicts a metal scaffolding structure, likely being used for construction purposes, with a prominent yellow oval highlighting a specific section of the scaffolding and a sign that reads "CAUTION: Cable pulling activity in progress. Pinch points for".	This image depicts a construction site with a scaffolding system that poses a risk of falling objects, highlighting the need for proper safety measures to prevent accidents.
	7	슬링을 날카로운 철골 구조물에 임시로 걸어둠에 따른 마모-파단 및 중량물 추락사고 위험	This image shows a close-up of a large industrial structure, with a red rope or cable suspended from the top of the structure.	This image depicts a red cable that poses a risk of entanglement for workers on a construction site, highlighting the importance of proper cable management and safety protocols.

수행 결과 - 문서 분석 1

Phi-4(14B) 모델을 이용하여 문서 내용을 분석한 후 **주요 항목 검출**과 **내용을 요약**하여 **구조화**하는 태스크에 대한 미세조정 후 테스트를 한 결과 **GPT4급에 준하는 성능**을 확인하는 고무적인 결과를 얻었습니다.

● 주요항목 검출, 구조화 및 주요내용 요약 학습 결과(Phi-4)

700건 테스트 (대상 모델)	BERT Score	LLM Judge (by GPT-4o)	미세조정 방법
gpt-4o	<u>0.7579</u>	<u>92.6</u>	-
gpt-4o-mini	<u>0.7838</u>	<u>91.4</u>	-
phi-4(original)	0.7624	80.4	미세조정 전
tuing-phi-4-q-e5	0.7634	93.2	Epoch 5, 6000건

수행 결과 - 문서 분석 2

Gemma 2(2B) 모델을 이용하여 문서 내용을 분석한 후 주요 내용을 요약하는 단일 태스크에 대한 미세조정 후 테스트를 한 결과 **GPT4급에 준하는 성능**에는 미치지 못하였으나 **파라미터의 절대적 소규모(천배~십만배 이상 소규모)**에 비하면 **고무적인 결과**를 얻었습니다.

● 주요내용 요약 학습 결과(Gemma 2)








700건 테스트	BERT Score	LLM Judge (by GPT-4o)	미세조정 방법
gpt-4o	0.6952	94.2	-
Gemma-2-2b-it	0.6001	8.99	미세조정 전
train_gemma-2-2b-it_6	0.7634	72.91 (평균)	Epoch 20, 16000건 추론 파라미터 조정
train_gemma-2-2b-it_7	0.7820	70.22 (평균)	Epoch 30, 16000건 추론 파라미터 조정

수행 결과 – 이미지 식별

LLaMA 3 Vision(11B) 모델과 **Phi-4 Multimodal(5.6B)** 모델을 이용하여 공개된 건설현장 사진 이미지 데이터를 이용하여 이미지 설명 및 위험요소 식별 테스트를 **미세조정 없이** 진행하여 **영문의 경우 중간 이상의 성적**을 얻었으며, 미세조정을 수행한다면 더 나은 결과를 기대할 수 있습니다.

● 이미지 설명 및 안전문제 식별 결과

Llama-3.2-11B-Vision-Instruct 이미지 설명 테스트(영어)

이미지	id	Human description (영문)	단순 description	안전 description
	1	헬 모서리에 의한 슬링(스트랩) 손상 및 추락사고 위험	The image depicts a yellow metal bar secured with a green and orange cable tied to a metal grate.	The image depicts a yellow beam wrapped in electrical tape on a metal grate, potentially posing a tripping hazard and requiring immediate attention from site safety management.
	2	파레트를 임시발판으로 사용함에 따른 추락사고 위험	This image depicts a man standing on a metal grate platform, with a wooden pallet and a fan visible beneath him.	The image shows a wooden pallet on top of a metal grate, which poses a risk of falling debris and entrapment for workers below.
	3	노후 파손된 목재 블록 사용으로 인한 중량물 진동 및 낙하 위험	This image shows a blue flatbed trailer loaded with wooden planks, some of which are stacked on top of one another and others that are resting against the trailer's side.	This image shows a large piece of lumber protruding from a pile of wood on a flatbed trailer, posing a potential risk of falling off and causing injury to workers or damage to equipment.
	4	무(無)바닥 굴착 작업으로 인한 토사 붕괴 및 추락사고 위험	This image depicts a large hole in the ground, likely a sinkhole or excavation site, with a person standing at the top and another at the bottom, surrounded by a metal ladder and other tools.	From a construction site safety management perspective, this image depicts a significant hazard: an open and unguarded excavation pit with a steep drop-off, posing a substantial risk of injury or fatality to workers and others in the vicinity.
	5	난간이 없는 높은 구조물 위에서 안전대(벨트) 미착용 상태로 작업에 따른 추락사고 위험	The image depicts three workers in hard hats and safety gear, engaged in constructing a metal structure on a metal platform.	The image shows three construction workers in high-visibility vests and hard hats working on a metal structure, highlighting the importance of proper safety gear and protocols in preventing accidents and ensuring a safe work environment.
	6	계단 외부에 방치된 아르곤 가스 실린더(용기)로 인한 전도-충돌 및 가스 누출 사고 위험	This image depicts a metal scaffolding structure, likely being used for construction purposes, with a prominent yellow oval highlighting a specific section of the scaffolding and a sign that reads "CAUTION: Cable pulling activity in progress. Pinch points for".	This image depicts a construction site with a scaffolding system that poses a risk of falling objects, highlighting the need for proper safety measures to prevent accidents.
	7	승강을 날카로운 철골 구조물에 임시로 걸어놓음에 따른 마모 파단 및 중량물 추락사고 위험	This image shows a close-up of a large industrial structure, with a red rope or cable suspended from the top of the structure.	This image depicts a red cable that poses a risk of entanglement for workers on a construction site, highlighting the importance of proper cable management and safety protocols.

이미지 설명

3점 척도

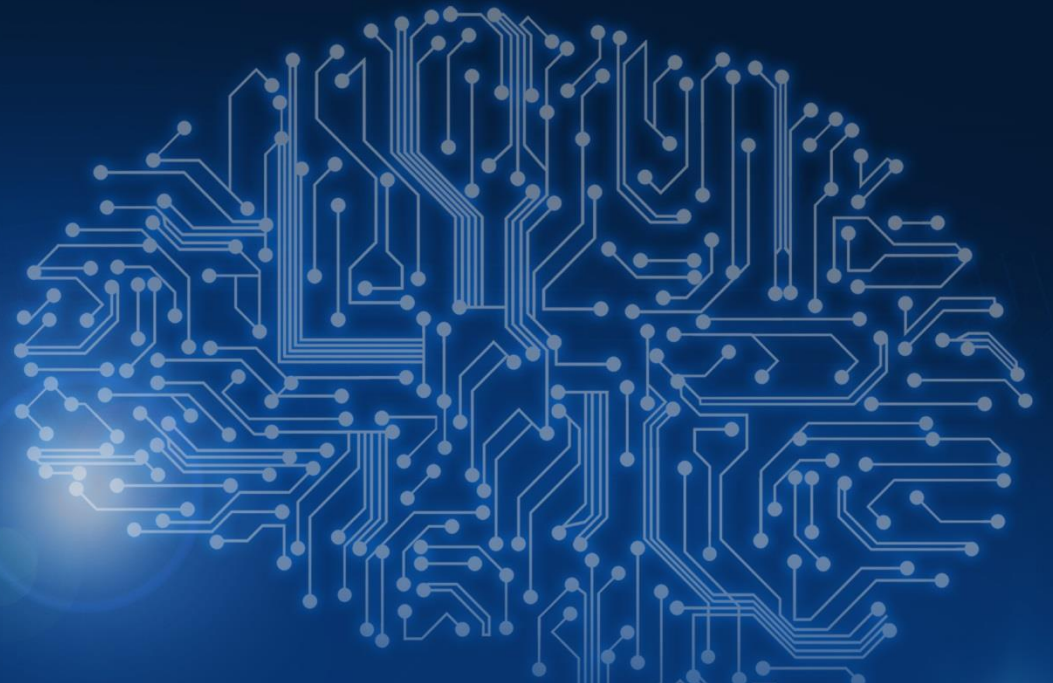
모델	한글 프롬프트 평균	영문 프롬프트 평균
Llama-3.2-11B-Vision-Instruct	1.5	2.6
Phi-4-multimodal-instruct	1.8	2.1

위험요소 식별

3점 척도

모델	한글 프롬프트 평균	영문 프롬프트 평균
Llama-3.2-11B-Vision-Instruct	1.4	2.3
Phi-4-multimodal-instruct	1.1	1.3

Q&A



End of Document