

생성형 AI GPT를 활용한 서비스 개발



1. 생성형 AI
2. 사업 분야
3. 개발 방법
4. 보유 기술

AI 기술은 지난 수십 년을 통틀어 가장 중요한 진보. AI 에이전트가 다음 플랫폼이 될 것입니다.

`_빌 게이츠_gatesnote.com`



1.1. 새로운 기회 생성형 AI GPT

생성형 AI ChatGPT가 단순히 말장난식의 대화만 하고 있을까? 아니다, 이미 생성형 AI는 금융, 제조, 교육 등으로 무한 확장 중이며 인간의 일을 대신해 나가고 있다. 미국의 투자은행인 골드만삭스는 생성형 AI가 10년 후 세계 일자리 3억 개에 영향을 미칠 것으로 예상했다. 우리는 AI에 대체될 것인가? 아니면 AI를 이용해 더 발전할 것인가? 지금이 그 변곡점이며 기회를 잡으러 움직여야 할 때이다.



1.2. 왜 생성형 AI인가?

Chat GPT로 대변되는 생성형AI는 '인터넷', 또는 모바일 혁명을 일으킨 '아이폰'에 준하는 혁신으로 평가받고 있다. 생성형 AI는 단순한 대화를 넘어서 인간 수준의 추론을 할 수 있다. 이를 이용해 기존의 다양한 데이터와 연계해서 나만의 특별한 AI 서비스를 만들 수 있다.

생성형 AI에 대한 업계 전문가 평가

“

사용자의 질문에 놀라울 정도로 인간과 유사하게 답하는 챗봇인 Chat GPT가 인터넷만큼 중대한 발명이라고 믿는다.”

빌 게이츠, Microsoft 창업자



“

인공지능(AI)의 아이폰 시대가 시작됐습니다. 생성형 AI가 모든 산업을 재창조할 것입니다.”

젠슨 황, Nvidia 창업자



생성형 AI가 주목받는 이유

기술 측면

- 1 인간의 고유 영역에 도전하는 기술의 혁신
단순 데이터 인식을 넘어 ‘창조’의 영역에 진입한 생성형 AI 기술
- 2 가속화된 생성형 AI 기술의 진보 속도
GPT 3에서 3.5까지 2년, 3.5에서 4까지 4개월로 가속화된 업그레이드 GPT AI 모델 출시 주기

산업 측면

- 3 ‘AI의 산업화 시대’를 이끌어 갈 서비스 생태계 활성화
대형 언어 모델을 기반으로 한 수많은 AI 애플리케이션과 서비스의 출시 가속

출처: Chat GPT가 부여하는 기회: 생성형 AI 기반 혁신의 현실화 – strategy&

2.1. 생성형 AI의 사업분야

생성형 AI를 이용할 수 있는 사업분야는 대부분의 산업군을 포함한다. 생성형 AI는 단지 콘텐츠 생성과 고객응대 CS, 마케팅, 광고에만 이용되는 것이 아니다. 법률, 건강 관리, 부동산, 금융 서비스 등 다양한 산업에서 채택되고 있고 다양한 실증 사례를 만들어 가고 있다.

콘텐츠생성

글쓰기 및 편집: 블로그 글, 뉴스 기사, 책 등의 텍스트 콘텐츠 생성
이미지 및 디자인: 광고, 소셜 미디어, 웹 디자인 등의 시각적 콘텐츠 생성
비디오 및 애니메이션 생성: 유튜브, 쇼츠 비디오 등의 동영상 콘텐츠 생성

교육 및 트레이닝

교육용 콘텐츠 생성: 학습 목표에 맞춰 개별화된 교육 자료 자동으로 생성
언어 학습 자료 생성: 실시간으로 개별화된 외국어 대화 시나리오 생성
퀴즈 생성: 학습을 위한 맞춤형 연습 문제, 퀴즈 생성

법률 분야

문서 자동 생성: 계약서, 소송 문서, 기타 법률 문서를 자동으로 생성
법률 자문 및 연구: 대량의 법률 데이터 및 판례 검색 지원
준수 및 리스크 관리: 법률 규정 준수 확인 및 리스크 평가

부동산 분야

시장 분석 및 예측: 부동산 시장 트렌드 분석 및 가격 변동 예측
매물 설명 생성: 매물의 특징을 분석하여 매력적인 설명 자동 생성
가상 투어 콘텐츠 생성: 부동산 가상 투어를 위한 3D, 이미지, 비디오 생성

고객 서비스 및 지원

자동화된 고객 서비스 챗봇: 고객 문의에 실시간으로 응답하는 챗봇 운영
FAQ 및 지원 문서 자동 생성: FAQ 문서를 자동으로 생성하여 고객 지원
고객 문의 피드백 처리: 고객의 문의사항을 분석하고 자동으로 처리

마케팅 및 광고

타겟 마케팅: 타겟 마케팅 캠페인을 위한 콘텐츠 생성
소셜 미디어: 소셜 미디어 전략 및 캠페인 자동화
맞춤형 광고: 소비자 피드백 및 트렌드 분석을 통한 광고 내용 생성

스마트 헬스케어

의료 기록 분석 및 생성: 환자 의료 기록의 분석 및 요약 보고서 생성
임상 연구 지원: 임상 시험 설계 및 데이터 분석을 통한 연구 개발 지원
개인화된 건강 권장사항: 환자 맞춤형 건강 관리 및 치료 계획 제안

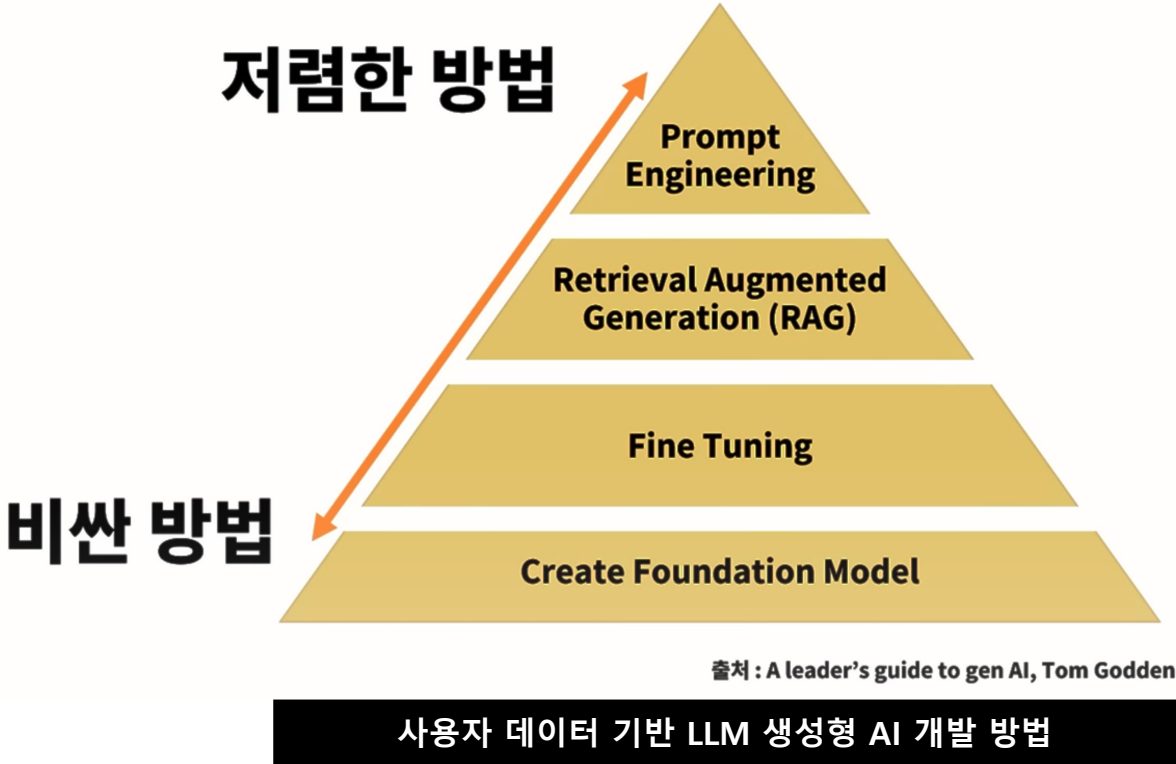
금융 서비스

금융 보고서 및 분석: 금융 보고서 생성 및 시장 데이터 분석
사기 탐지 및 리스크 관리: 거래 패턴 분석을 통한 사기 탐지
개인화된 금융: 개인 금융 상태 분석을 통한 맞춤형 금융 계획, 투자 제공



3.1. 생성형 AI 서비스 개발 방법

LLM 기반의 생성형 AI가 환각현상 없이 사용자의 데이터를 기반으로 해서 대답을 하게 하는 다양한 방법이 있다. 자체 LLM 모델을 구축하거나 기존 모델을 Fine-tuning하는 방법은 비용과 시간이 많이 들어간다. 반면 프롬프트 엔지니어링과 RAG는 비교적 저렴하여 많이 사용하고 있는 기법이다. 특히 프롬프트 엔지니어링은 지시어를 입력하는 방식에 따라서 보다 정확한 데이터를 받을 수 있어서 주목받는다.

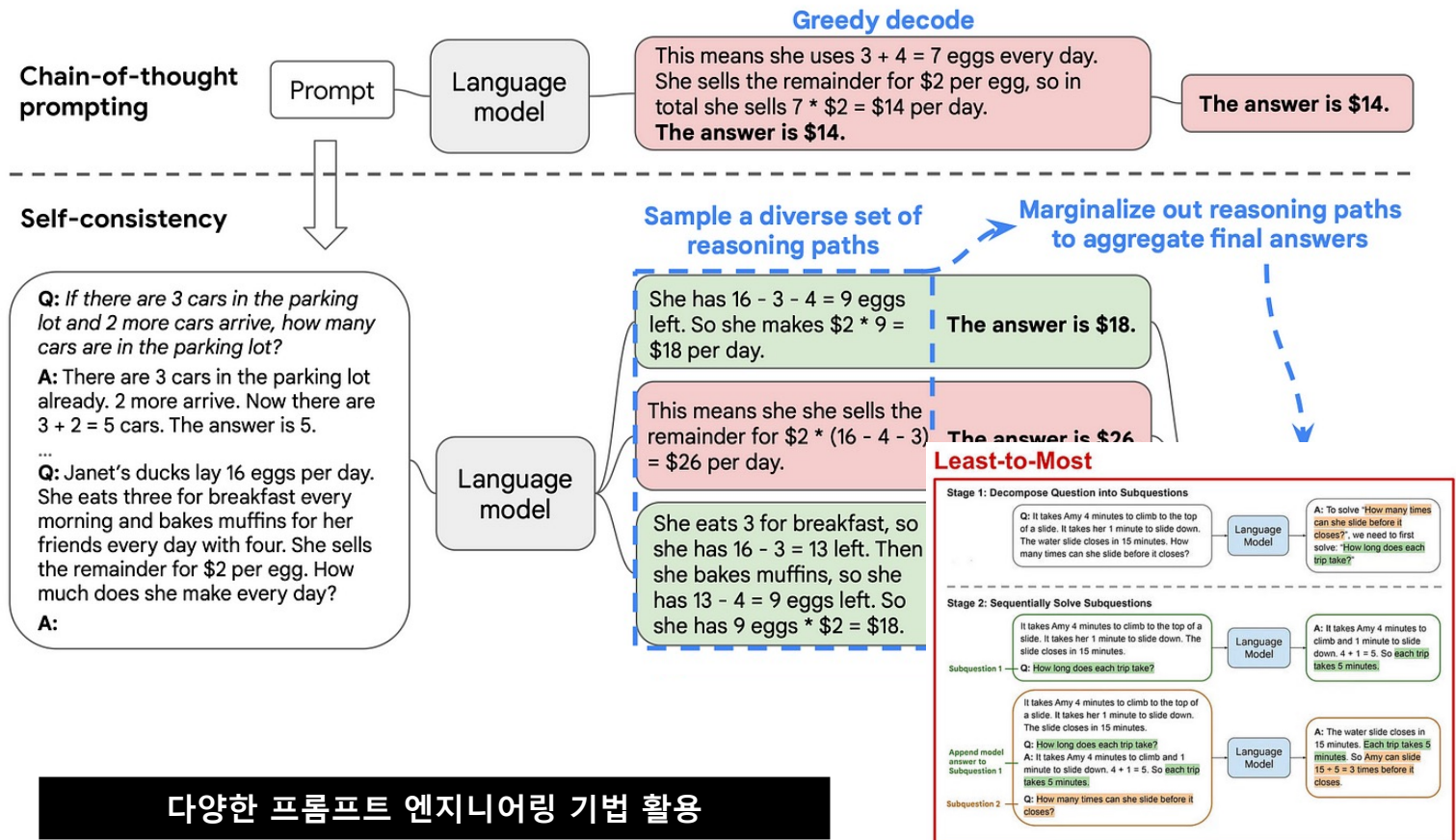


- Few-shot Learning
- Chain-of-Thought(COT)
- Self-Consistency
- Selection-Inference
- Least-to-Most
- ReAct
- Self Evaluation

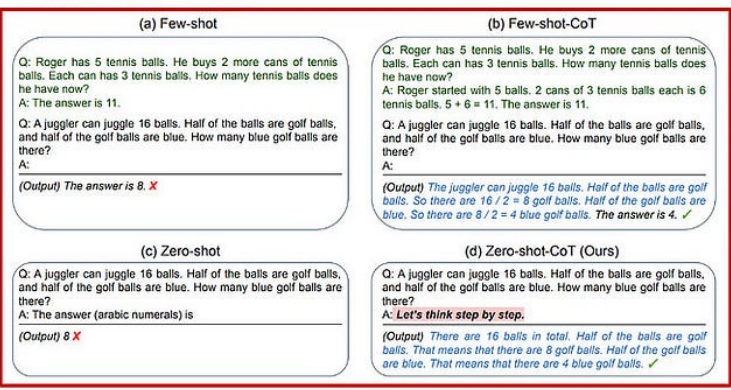
주요 프롬프트 엔지니어링 기법

3.2. 프롬프트 엔지니어링 - 생성형 AI 서비스 개발 방법

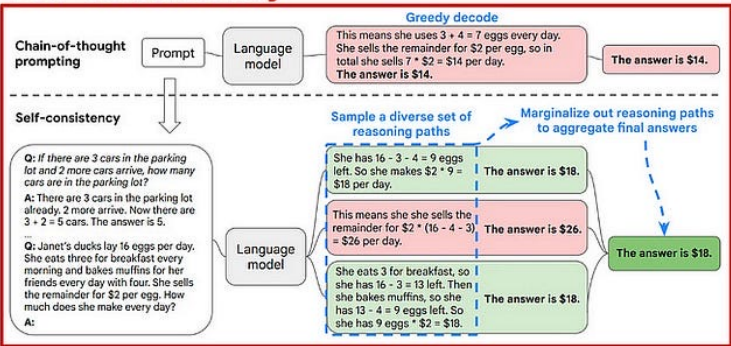
ChatGPT와 같은 LLM은 제공하는 지시어(Prompt)에 따라서 전혀 다른 결과를 낸다. LLM은 대화의 맥락을 기억하고 대답하며, 역할과 예시에 따라서 충실히 답변한다. 프롬프트 엔지니어링은 이러한 특성을 이용해서 최적의 대답을 이끌어 내는 기법들이다. 가장 많이 사용하는 기법은 1)역할(Role:페르소나)을 주고 2)말투와 같은 특성을 부여하고 3)적절한 예시(Few-Shot Prompt)를 주는 것들이다. 토큰이 허락하는 한 사전에 충분한 지식을 맥락(Context)으로 제공하는 것도 좋은 결과를 받을 수 있는 조건이다.



Zero-Shot CoT



Self-Consistency



다양한 프롬프트 엔지니어링 기법 활용

3.3. GPTs / OpenAI Assistant API 개발 - 생성형 AI 서비스 개발 방법

GPT스토어는 별도 코딩 없이 챗GPT 입력 창에서 문장을 입력해 맞춤형 챗봇인 'GPTs'를 만들고, 이를 다른 사용자와 공유하고 판매할 수 있는 서비스다. 스마트폰에 애플 앱스토어와 구글 플레이스토어가 있다면, AI에는 챗GPT의 GPT스토어가 있는 것이다. 이 새로운 스토어에 앱 이름과 기능을 선점하고 회사를 홍보하고자 많은 앱들이 등록되고 있다. 아직 진입장벽이 낮을 때 미리 앱을 등록해 놓는 것이 유리하다.

SEO Visual Describer

Draft

Create

Configure

You

Create a chatbot that generates SEO-friendly alternative text for visuals used in Zapier blog articles.

GPT Builder

Great! Let's start by giving your GPT a name. I suggest calling it **SEO Visual Describer**. Does that work for you or do you have another name in mind?

You

Sounds good!

GPT Builder

ChatGPT와 대화하며
앱 개발. 앱 아이콘도
자동으로 그려줌

Here's the initial profile picture I've created for "SEO Visual Describer":

Do you like this design, or would you like to make any changes?

Provide alt text for this image of a workfl...

Create descriptive alt text for this illustr...

Suggest SEO-friendly alt text for this gra...

Generate alt text for this screenshot of a ...

Message GPT Builder...

ChatGPT4 GPTs 개발 채팅창

Preview

실시간으로 앱 상태를
확인하면서 개발

SEO Visual Describer

I generate SEO-friendly alt text for visuals in Zapier's blog articles.



새로운 앱스토어 GPTs



코딩 없이 텍스트만
으로 누구나 앱개발
가능

GPTs

Discover and create custom versions of ChatGPT that combine instructions, extra knowledge, and any combination of skills.

Search public GPTs

Top Picks

DALL·E Writing Productivity Research & Analysis Programming Education Lifestyle

Featured

Curated top picks from this week



AllTrails

Find trails that fit your nature for your next hike, ride, or run.

By alltrails.com



Consensus

Your AI Research Assistant. Search 200M academic papers from Consensus, get science-based...

By Consensus



Code Tutor

Let's code together! I'm Khanmigo Lite, by Khan Academy. I won't write the code for you, but I'll help you...

By khanacademyorg



Books

Your AI guide in the world of literature and reading.

By Josh Brent N. Villacido



Creative Writing Coach

I'm excited to read your work and give you feedback to improve your skills.



Laundry Buddy

Ask me anything about stains, settings, sorting and everything laundry.



Game Time

I can quickly explain board games or card games to players of any skill level. Let the games begin!



Tech Advisor

From setting up a printer to troubleshooting a device, I'm here to help you step-by-step.

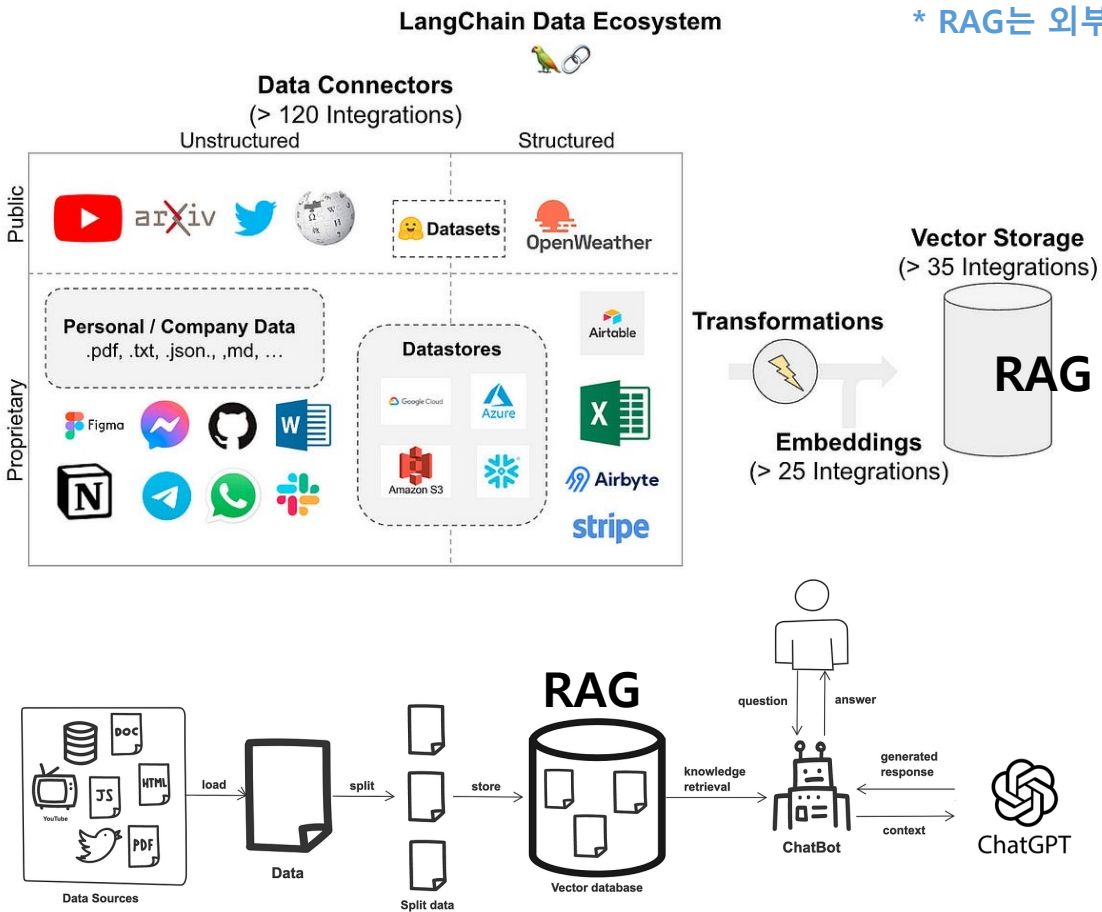
GPTs Store

See more

3.4. Langchain과 RAG를 이용한 GPT 챗봇 서비스 - 생성형 AI 서비스 개발 방법

Langchain은 대규모 언어 모델(LLM)을 사용하여 애플리케이션을 개발하기 위한 오픈 소스 프레임워크이다. 이를 통해 ChaGPT를 비롯해 구글의 Bard, 메타의 LLaMa까지 LLM을 선택해서 이용할 수 있다. 또한 다양한 형태의 정형, 비정형 데이터를 Vector Storage에 저장해 RAG(검색 기반 답변 생성: Retrieval Augmented Generation)를 구성하여, 보다 정확하고 신뢰성 있게 AI 챗봇 답변을 만들 수 있다.

* RAG는 외부 소스에서 가져온 정보로 생성 AI 모델의 정확성과 신뢰성을 향상시키는 기술이다.



pdf, 엑셀, SQL 등 다양한
정형, 비정형 데이터와
연동해서 답변 생성



벡터DB + RAG(검색 기
반 답변 생성)를 이용해
LLM의 성능 향상. 로컬
에 있는 보안 데이터도
연동해서 안전하게 검색
가능.



3.5. sLLM을 이용한 Private 생성형 AI - 생성형 AI 서비스 개발 방법

sLLM은 GPT와 같은 대형 LLM과 비교했을 때 매개변수의 크기가 비교적 작은 언어모델을 의미한다. sLLM은 2023년 메타가 라마(LLaMA)를 오픈소스로 공개하면서부터 주목받기 시작했다. 저렴하면서도 유연해 기업 맞춤형으로 적합하다. sLLM을 이용하면 보안이 유지된 기업 AI 서비스를 개발할 수 있다. 다만 미세조정(fine-tuning) 등의 인공지능 기술이 있어야 하고 비싼 GPU를 구매해야하는 단점이 있다.

대규모언어모델과 소형 대규모언어모델 비교

구분	대규모언어모델(LLM)	소형 대규모언어모델(sLLM)
크기(파라미터)	약 수천억 개	약 수십억 개
예시	오픈AI GPT-4, 구글 팜2, 네이버 하이퍼클로바X, LG 엑사원	메타 라마, 스탠퍼드대 알파카, 스캐터랩 핑퐁
특징	정확하고 복잡한 작업, 방대한 컴퓨팅 자원 필요	적은 컴퓨팅 활용, 특정 영역 언어에 특화, 신속한 파인튜닝



* 매개변수 줄이고 파인튜닝으로 최적화
저렴하면서도 유연해 기업 맞춤형으로 적합



LLaMa2 70B 모델은 ChatGPT 3.5 성능 도달

Llama 2 was trained on 40% more data than Llama 1, and has double the context length.

Llama 2

MODEL SIZE (PARAMETERS)	PRETRAINED	FINE-TUNED FOR CHAT USE CASES
7B	Model architecture:	Data collection for helpfulness and safety:
13B	Pretraining Tokens: 2 Trillion	Supervised fine-tuning: Over 100,000
70B	Context Length: 4096	Human Preferences: Over 1,000,000

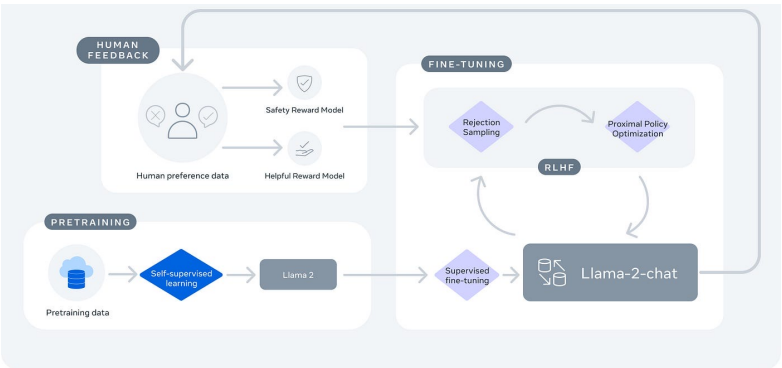
LLaMa2 한국어 파인튜닝 모델

id4thomas/llama2-koalpaca

Training llama2 weights with koalpaca dataset



1 Contributor, 0 Issues, 4 Stars, 1 Fork



4.1. 당사 보유 기술

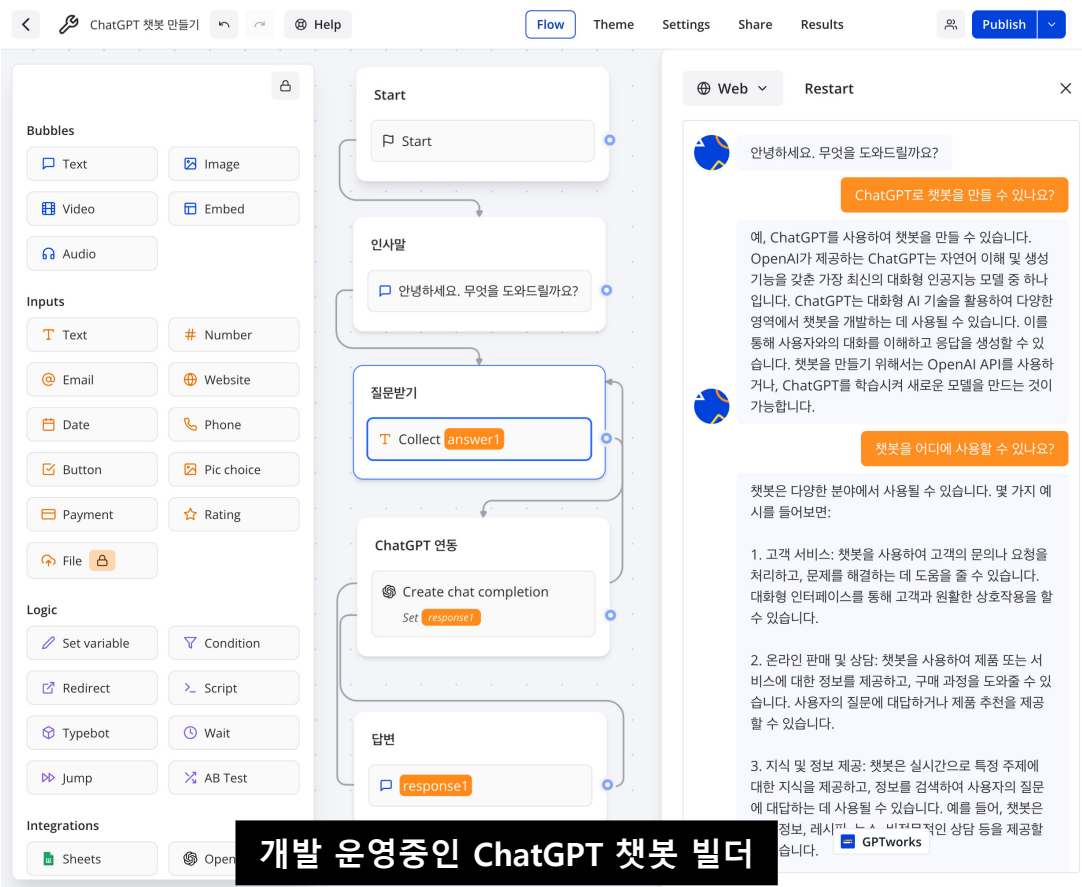
당사는 Botpress처럼 ChatGPT를 이용해 지능형 챗봇을 직접 만들 수 있는 빌더를 개발해 운영하고 있다. 이에 챗봇을 구성하는 LLM 연동과 프론트엔드부터 데이터와 연동하는 백엔드까지 모두 구성할 수 있다. 고객의 사이트의 필요에 따라 프롬프트 엔지니어링과 RAG(검색 기반 답변 생성)를 효과적으로 사용하여 정확성과 신뢰성을 갖춘 챗봇을 만들고 있다. 모바일 앱 개발에도 많은 경험을 갖추고 있다.



GPT 인공지능 챗봇 앱 개발

 **OpenAI**
ChatGPT API

ChatGPT 기반 자동화 서비스



개발 운영중인 ChatGPT 챗봇 빌더

보유기술 리스트

- LLM: GPT, BERT, LLaMa
- Llama.cpp, Ollama
- GPTs Development
- OpenAI GPT API
- Prompt Engineering
- sLLM Fine-tuning
- Langchain Framwork
- RAG(Retrieval-Augmented Generation)
- Vector Store
- Chroma DB
- FAISS
- Pinecone
- Whisper AI (TTS/STT)
- Chatbot Dev (Web/Mobile App)
- Web Front-end / Back-end
- Mobile App Dev



4.2. 제공 서비스

당사는 Botpress처럼 ChatGPT를 이용해 지능형 챗봇을 직접 만들 수 있는 빌더를 개발해 운영하고 있다. 이에 챗봇을 구성하는 LLM 연동과 프론트엔드부터 데이터와 연동하는 백엔드까지 모두 구성할 수 있다. 고객의 사이트의 필요에 따라 프롬프트 엔지니어링과 RAG(검색 기반 답변 생성)를 효과적으로 사용하여 정확성과 신뢰성을 갖춘 챗봇을 만들고 있다. 모바일 앱 개발에도 많은 경험을 갖추고 있다.

ChatGPT 이용 서비스 개발

- 프롬프트 엔지니어링
- GPTs 개발
- ChatGPT API 연동
- Assistant API 연동
- GPT 3.5 Fine-tuning
- GPTs 연동 외부 API 개발
- ChatGPT 기반 챗봇 웹 개발
- ChatGPT 기반 RAG 적용
- ChatGPT + Langchain 연동



ChatGPT API

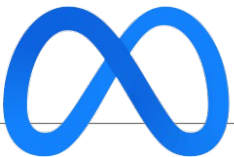
챗봇 기능 앱 개발

- 웹 기반 LLM 챗봇 개발
- LLM 챗봇 기능 모바일 앱 개발
- Whisper API 이용 음성인식
- LLM 연동 API 개발
- Langchain + RAG 연동
- Flutter를 이용 멀티플랫폼 지원
- 네이티브 iOS / Android 개발
- Web Front-end / Back-end



Private LLM 개발

- LLaMa2, Alpaca, Vicuna
- sLLM Fine-tuning
- Hugging Face 연동
- Open-Source Embedding
- Langchain Framework
- RAG
- Vector Storage
- 자체 sLLM 서버 구축





상 호 : 올리브스토리 주식회사(OliveStory Corp.)
대 표 이 사 : 곽 용 진 (sales@olivestory.co.kr)
주 소 지 : 서울특별시 영등포구 국제금융로 10 Three IFC 43층
사 업 분 야 : IT 컨설팅, System Integration
설 립 일 자 : 2012년 3월 13일
벤 처 인 증 : 2013년 8월 13일
파 트 너 쉽 : 여의도 성모병원 인공지능 영상분석 자문
천주교서울대교구 모바일 개발사, 서울특별시 북촌 IoT 시범사업단

	주요 사업 연혁
2012년 01월	천주교서울대교구 주요 모바일 어플리케이션 개발(매일미사, 가톨릭성경, 성무일도, 굿뉴스, 평화방송앱 등)
2014년 02월	비콘을 활용한 천주교 서울대교구 명동성당 무인 안내서비스 개발
2014년 08월	MBC 신사옥 콘텐츠 전시장 비콘 안내 솔루션 개발
2015년 05월	CBS레인보우라디오 모바일 및 방송관리 플랫폼 개발. LG CNS IoT 시스템 개발
2015년 09월	서울시 북촌 IoT 시범사업 (Beacon과 증강현실-AR을 이용한 안내서비스)
2016년 01월	중국 광저우 전시관 (Canton Fair) 비콘안내 시스템 개발
2016년 11월	LoRa IoT 네트워크를 이용한 화성시 국가기준점 관리 시스템 개발 (국토부 과제)
2017년 01월	코오롱 스포츠 IoT 센서 웨어러블 모듈 개발
2018년 10월	인공지능 자연어 음성처리 및 음성 데이터 디지털 아카이빙 솔루션 개발 - KT연구과제

Thank You.

OLIVESTORY[®]



sales@olivestory.co.kr