



OpenAI의 생성 AI가 불러온
범용 인공 지능 시대 개막

Microsoft 기술전략매니저 김현수

Agenda

1. AI 시대 개막
2. MS와 OpenAI의 협업
3. OpenAI 사용 사례
4. AGI 시대 개막

The background features a solid black field with two sets of abstract, wavy blue lines. One set of lines is located in the upper-left corner, and another set is in the lower-left corner. These lines are composed of many thin, parallel curves that create a sense of motion and depth.

PART 1. AI 시대 개막

THE AGE OF AI HAS BEGUN



빌게이츠가 The age of AI has begun 이라고 제목을 지은 이유가 무엇인지 분석해서 요약해줘.

✓ “빌게이츠 The age of AI has begun 제목 분석” 검색 중

✓ 답변을 생성하는 중...

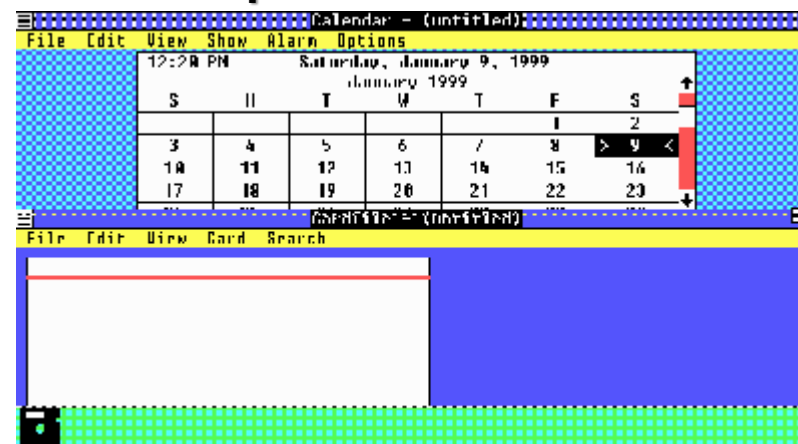
빌 게이츠는 인공지능이 모바일 폰과 인터넷과 같은 혁신적인 기술로서 혁명적이라고 생각한다고 합니다. 그는 1980년대에 그래픽 사용자 인터페이스를 처음 접했을 때와 같은 혁신적인 기술을 본 것이라고 말합니다¹. 그는 인공지능이 세상에서 가장 큰 문제를 해결하는 데 도움이 될 수 있다고 믿으며, 이를 통해 세상의 불평등을 줄일 수 있다고 생각합니다².

¹: The Age of AI has begun | Bill Gates ²: The Age of AI Has Begun

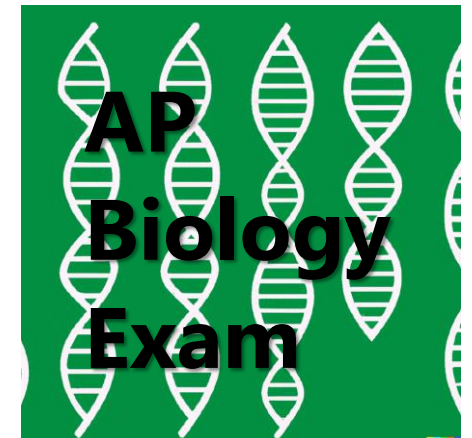
자세한 정보: [1. gatesnotes.com](https://www.gatesnotes.com) [2. pelayoarbues.github.io](https://pelayoarbues.github.io) [3. smallake.kr](https://smallake.kr) +2 더 보기

1 / 20

Graphical User Interface



OpenAI GPT (59/60)



AI에 대한 세상의 관심과 진입 장벽

ChatGPT Sprints to One Million Users

Time it took for selected online services to reach one million users



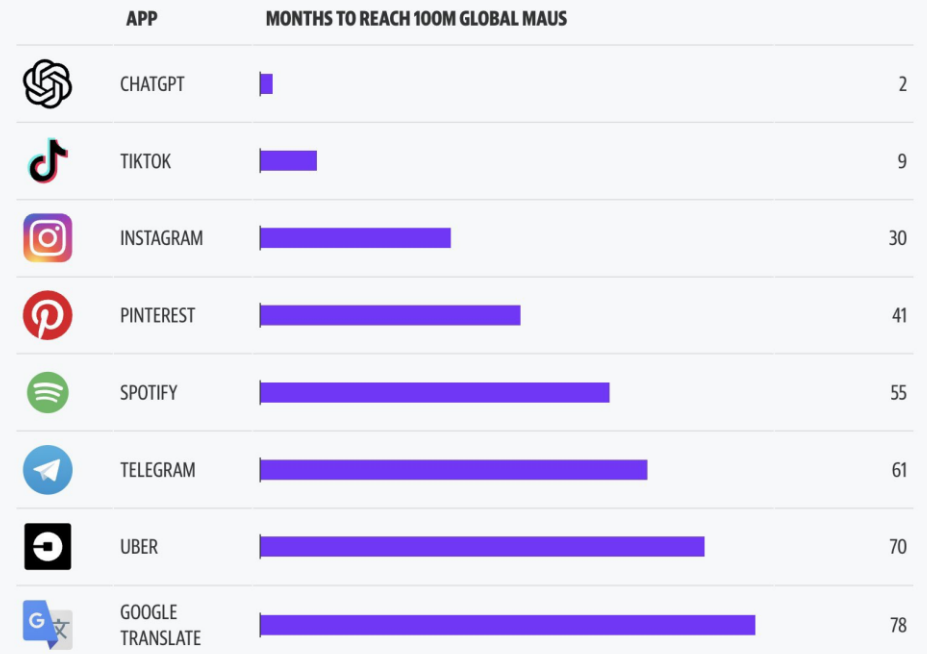
* one million backers ** one million nights booked *** one million downloads
Source: Company announcements via Business Insider/LinkedIn



statista

HOW LONG IT TOOK TOP APPS TO HIT 100M MONTHLY USERS

ChatGPT is estimated to have hit 100M users in January, 2 months after it's launch.
Here's how long it took other top apps to reach that:



SOURCE: UBS

yahoo!
finance

초거대 언어 모델 생성이 어려운 이유

D Total Compute Used to Train Language Models

This appendix contains the calculations that were used to derive the approximate compute used to train the language models in Figure 2.2. As a simplifying assumption, we ignore the attention operation, as it typically uses less than 10% of the total compute for the models we are analyzing.

Calculations can be seen in Table D.1 and are explained within the table caption.

Model	Total train compute (PF-days)	Total train compute (flops)	Params (M)	Training tokens (billions)	Flops per param per token	Mult for bwd pass	Fwd-pass flops per active param per token	Frac of params active for each token
T5-Small	2.08E+00	1.80E+20	60	1,000	3	3	1	0.5
T5-Base	7.64E+00	6.60E+20	220	1,000	3	3	1	0.5
T5-Large	2.67E+01	2.31E+21	770	1,000	3	3	1	0.5
T5-3B	1.04E+02	9.00E+21	3,000	1,000	3	3	1	0.5
T5-11B	3.82E+02	3.30E+22	11,000	1,000	3	3	1	0.5
BERT-Base	1.89E+00	1.64E+20	109	250	6	3	2	1.0
BERT-Large	6.16E+00	5.33E+20	355	250	6	3	2	1.0
RoBERTa-Base	1.74E+01	1.50E+21	125	2,000	6	3	2	1.0
RoBERTa-Large	4.93E+01	4.26E+21	355	2,000	6	3	2	1.0
GPT-3 Small	2.60E+00	2.25E+20	125	300	6	3	2	1.0
GPT-3 Medium	7.42E+00	6.41E+20	356	300	6	3	2	1.0
GPT-3 Large	1.58E+01	1.37E+21	760	300	6	3	2	1.0
GPT-3 XL	2.75E+01	2.38E+21	1,320	300	6	3	2	1.0
GPT-3 2.7B	5.52E+01	4.77E+21	2,650	300	6	3	2	1.0
GPT-3 6.7B	1.39E+02	1.20E+22	6,660	300	6	3	2	1.0
GPT-3 13B	2.68E+02	2.31E+22	12,850	300	6	3	2	1.0
GPT-3 175B	3.64E+03	3.14E+23	174,600	300	6	3	2	1.0

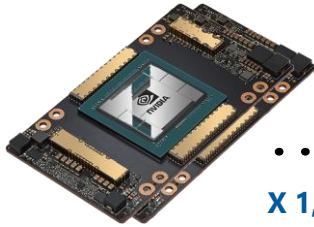
GPT-3
(175B)



X 1

NVIDIA V100

355
Years



...

X 1,024



NVIDIA A100

34
Days

\$5M



PART 2. MS와 OpenAI 협업

OpenAI와 Microsoft의 파트너십



범용 인공 지능(AGI)으로
인류에게 도움을 주자



지구상의 모든 사람과 조직이 더
많은 것을 성취할 수 있도록 지원

GPT-3

텍스트 생성 및 이해

ChatGPT

챗봇형 AI 모델
(preview)

Codex

코드 생성 및 이해

GPT-4

(preview)

DALL·E

텍스트 프롬프트에서 이미지 생성
(preview)

Microsoft AI 혁신 여정

September 2016

Microsoft achieves human parity in conversational speech recognition.

March 2018

Microsoft reaches human parity in machine translation.

July 2019

Microsoft and OpenAI announce exclusive partnership to accelerate breakthroughs in AI.

May 2020

Microsoft announces its first AI supercomputers, built in collaboration with and exclusively for OpenAI.

October 2020

Microsoft announces AI breakthroughs in image captioning available in Azure, Word, Outlook, and PowerPoint.

June 2021

GitHub launches Copilot, an AI pair programmer that takes advantage of the OpenAI Codex model to support developers.

May 2022

Microsoft expands access to Azure OpenAI Service with more models and new responsible AI system.

October 2022

Microsoft announces DALL-E 2 in Azure OpenAI Service, new Designer app, and Bing Image Creator.

January 2018

Microsoft achieves human parity in reading comprehension.

June 2019

Microsoft integrates neural machine translation research breakthroughs into Azure and achieves human parity in general language understanding.

February 2020

Microsoft announces Turing-NLG language model at 17 billion parameters.

September 2020

Microsoft teams up with OpenAI to exclusively license GPT-3 models.

May 2021

Microsoft introduces its first product features powered by GPT-3 models.

November 2021

Microsoft introduces Azure OpenAI Service.

June 2022

GitHub Copilot becomes generally available.

January 2023

Microsoft announces Azure OpenAI Service is now generally available. ChatGPT coming soon.



Azure 기반의 AI 슈퍼컴퓨터

- 초거대 AI 훈련을 위해 설계된
전 세계 5번째로 큰 클라우드 기반 첫 번째 슈퍼컴퓨터
- 30만 CPU 코어, 1만개 GPU로 구성
- OpenAI 파트너십을 가지고 구축

고도로 분산된 프레임워크

- 초거대 모델 훈련을 위한 오픈소스 프레임워크
(20배 더 크고/10배 더 빠름)
- DeepSpeed: 초거대 모델 훈련이 가능한 분산 프레임워크
(2000억개 파라미터)
- Azure Machine Learning: 새로운 ONNX 분산 훈련을 통한
오케스트레이터. 훈련을 위한 Azure Machine Learning 레시피

Microsoft의 차별화된 디지털 경험 제공

Applications

- Teams Premium
- GitHub Copilot
- New Bing / Edge

Low Code

- AI Builder
- DAX Generator

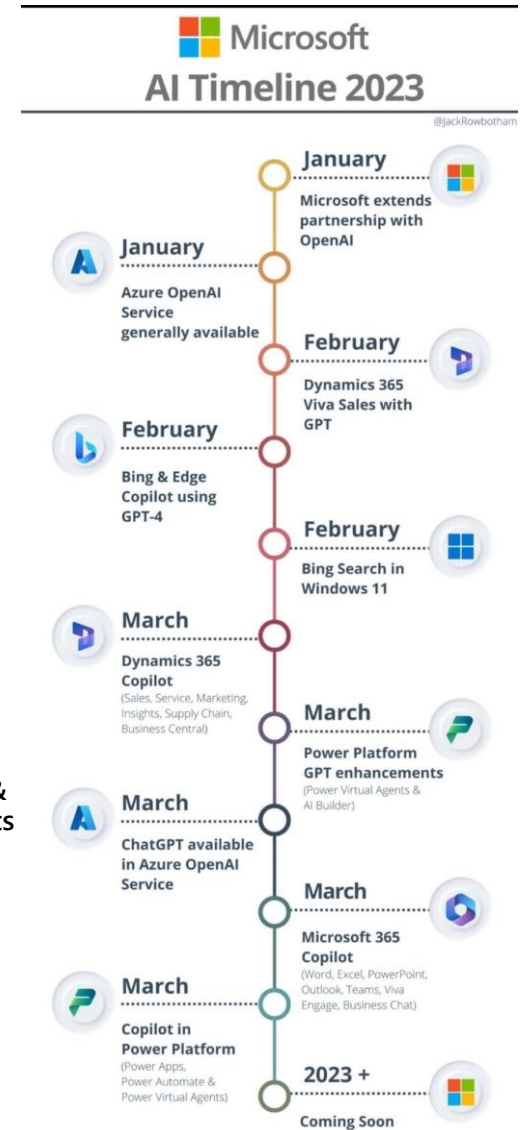
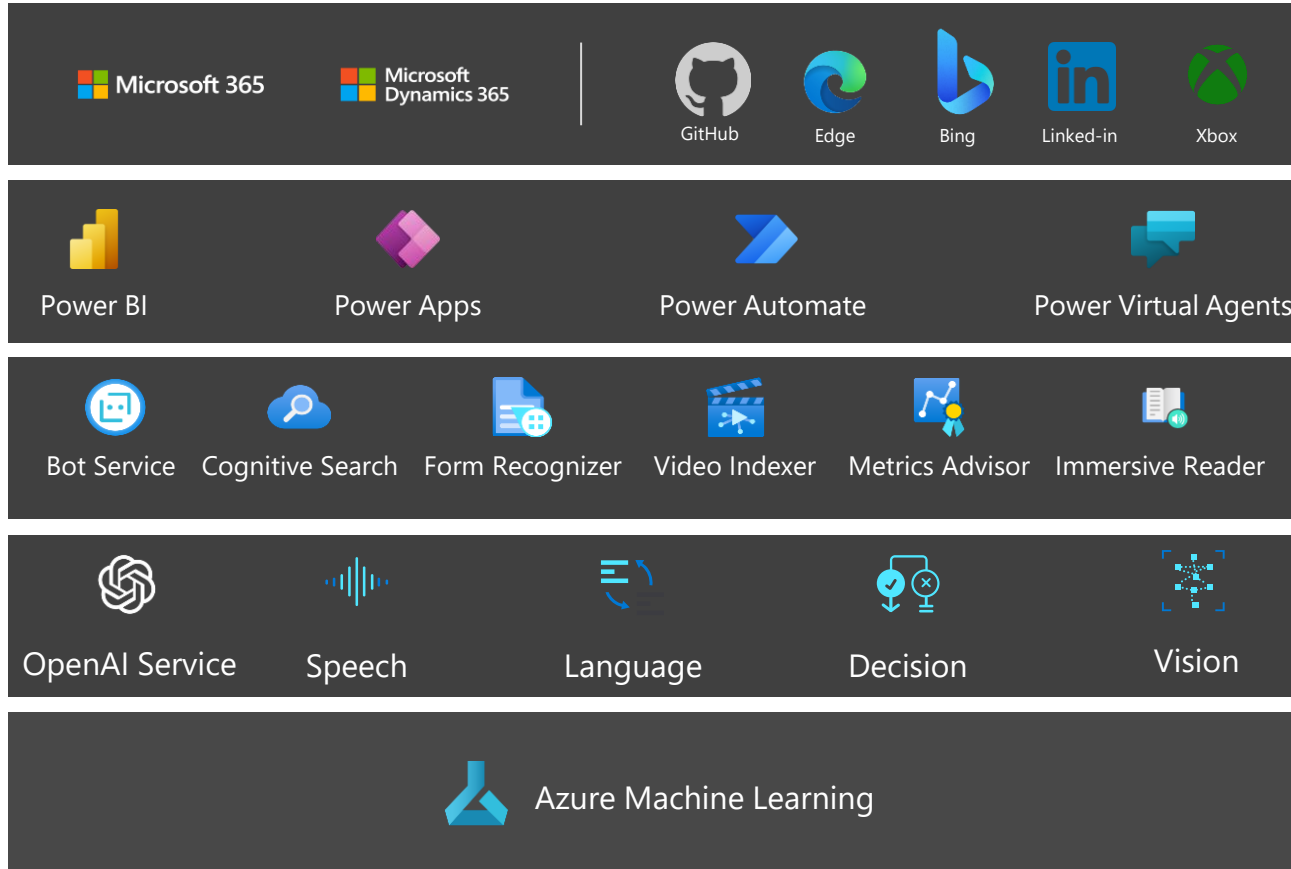
Scenario-Based Services

- Applied AI Services

Customizable AI Models

- Azure Open AI Service (GPT, DALL-E, CODEX)

ML Platform





Microsoft 365 Copilot

Embedded across Microsoft 365 apps





PART 3. OpenAI 사용 사례

주요 사용 사례



컨텐츠 생성

컨택 센터 : 고객 질의에
대한 맞춤형 답변 생성

웹 : 맞춤형 UI 생성 (개인화)



요약 및 분석

컨택 센터 : 콜 로그에 대한 요약

전문 영역 주제에 대한 요약
(회계/ 마케팅 보고서, SNS 外)



코드 생성

개발자 / 데이터 관리자 :
코드 및 SQL 쿼리 등 작성

운영자 :
각종 스크립트 작성



시맨틱 검색

검색 결과 리뷰
(제품/ 서비스 外)

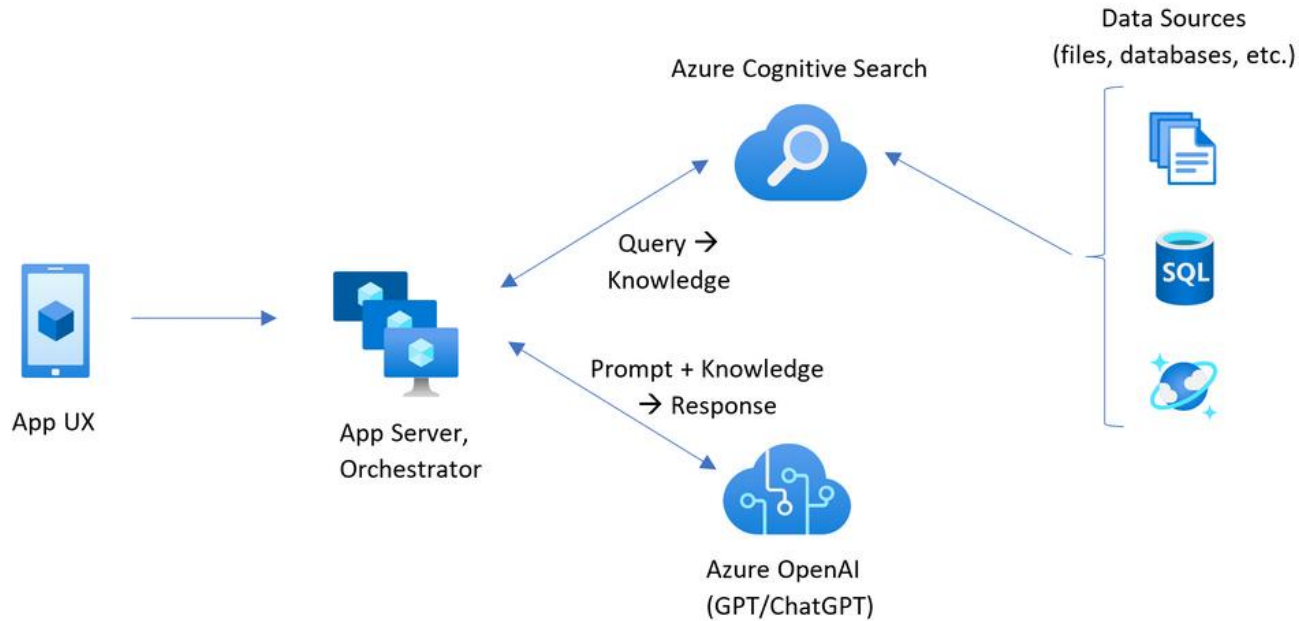
정보 검색 및 지식화

컨택 센터 활용 시 : 분류, 감정분석, 주요 정보 추출, 요약, 답변 생성

고객 관리 목적 : 개별 고객에 맞춤형 정보 제공 (고객 문의, 트렌드, 검색결과, 조회결과 생성)

업무 자동화 : 정형 비정형 문서 검색 / 매크로, 스크립트, 코드 생성 / 컨텐츠 생성

ChatGPT를 활용한 QnA 서비스



URL: <https://github.com/Azure-Samples/azure-search-openai-demo>

GPT + Enterprise data | Sample Chat Ask a question Azure OpenAI + Cognitive Search

Clear chat Developer settings

Does my plan cover annual eye exams?

Both Northwind Health Plus and Standard plans offer coverage for vision exams¹. Northwind Health Plus offers coverage for vision exams, glasses, and contact lenses, while Northwind Standard only offers coverage for vision exams and glasses.

Citations: 1. Benefit_Options-2.pdf

Hearing too?

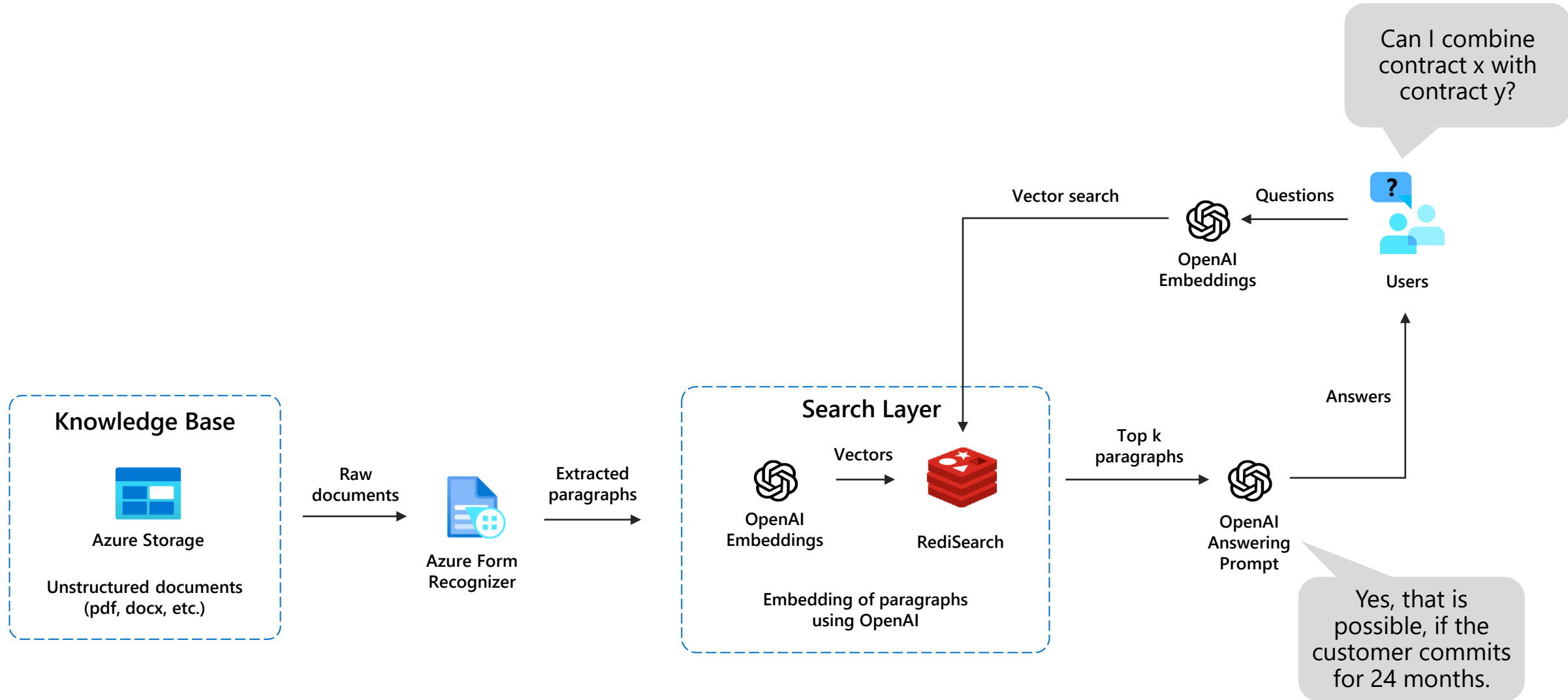
Both Northwind Health Plus and Standard plans offer comprehensive hearing care coverage, including hearing tests and evaluations, hearing aids, and other associated services. You can receive hearing care services from any in-network provider and enjoy comprehensive coverage for all hearing care services^{1 2}.

Citations: 1. Northwind_Health_Plus_Benefits_Details-29.pdf 2. Northwind_Standard_Benefits_Details-29.pdf

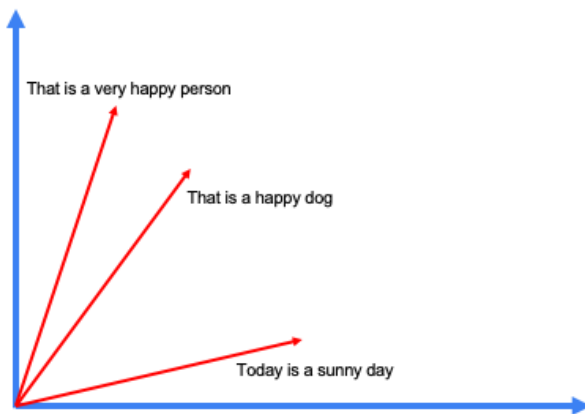
Follow-up questions: Does Northwind Health Plus cover hearing aids? What is Northwind Standard's coverage for hearing tests? What are "other associated services"?

Type a new question (e.g. does my plan cover annual eye exams?)

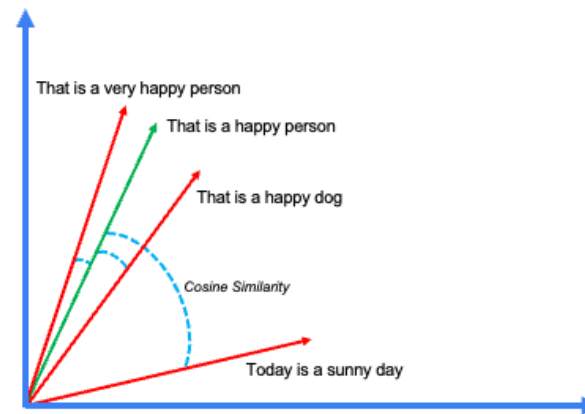
지식관리서비스(KMS)를 위한 시맨틱 Q&A 에이전트



임베딩: 데이터는 벡터로 저장하고 검색 가능



1 데이터 벡터화



2 검색: "That is a happy person"

Cosine Similarity

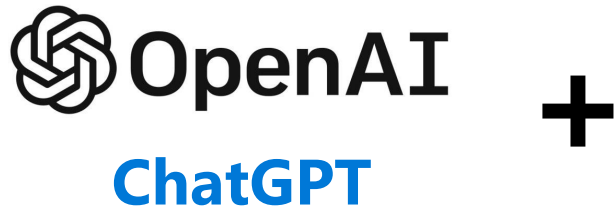
That is a happy dog	0.695
That is a very happy person	0.943
Today is a sunny day	0.257

3 유사도 결과












벡터 데이터베이스 종류



최신 정보로의 접근



Plugins

 Expedia 여행 계획을 현실로 만드세요. 그곳에 가고, 머물고, 보고 할 일을 찾으세요	 FiscalNote 법적, 정치적, 규제 데이터 및 정보에 대해 시장을 선도하는 실시간 데이터 세트를 엄선하여 이에 대한 액세스를 제공하고 가능하게 합니다.	 Instacart 좋아하는 지역 식료품점에서 주문하세요.
 KAYAK 항공편, 숙박 및 렌터카를 검색합니다. 예산 내에서 갈 수 있는 모든 장소에 대한 추천을 받으세요.	 Klarna Shopping 수천 개의 온라인 상점에서 가격을 검색하고 비교하십시오.	 Milo Family AI 매일 20분씩 조증을 마법으로 바꿀 수 있는 초능력을 부모에게 제공합니다. 질문: 헤이! 마일로, 오늘 마법은 뭐야?
 OpenTable 예약에 대한 직접 링크와 함께 레스토랑 추천을 제공합니다.	 Shop 세계 최고의 브랜드에서 수백만 개의 제품을 검색하십시오.	 Speak AI 기반 언어 튜터인 Speak와 함께 다른 언어로 무엇든 말하는 방법을 배우십시오.
 Wolfram Wolfram Alpha 및 Wolfram 언어를 통해 계산, 수학, 선별된 지식 및 실시간 데이터에 액세스합니다.	 Zapier Google Sheets, Trello, Gmail, HubSpot, Salesforce 등과 같은 5,000 + 앱과 상호 작용하십시오.	

The background features a solid black field with abstract, flowing blue lines on the left side. These lines are composed of many thin, parallel curves that create a sense of movement and depth, resembling a stylized wave or a digital signal.

PART 4. AGI 시대 개막

범용 인공 지능(AGI) 시대

AGI에 대해서 중요한 기능들을 몇 가지 추려줘.

✓ “AGI 기능” 검색 중

✓ 답변을 생성하는 중...

AGI의 기능 중 일부는 다음과 같습니다:

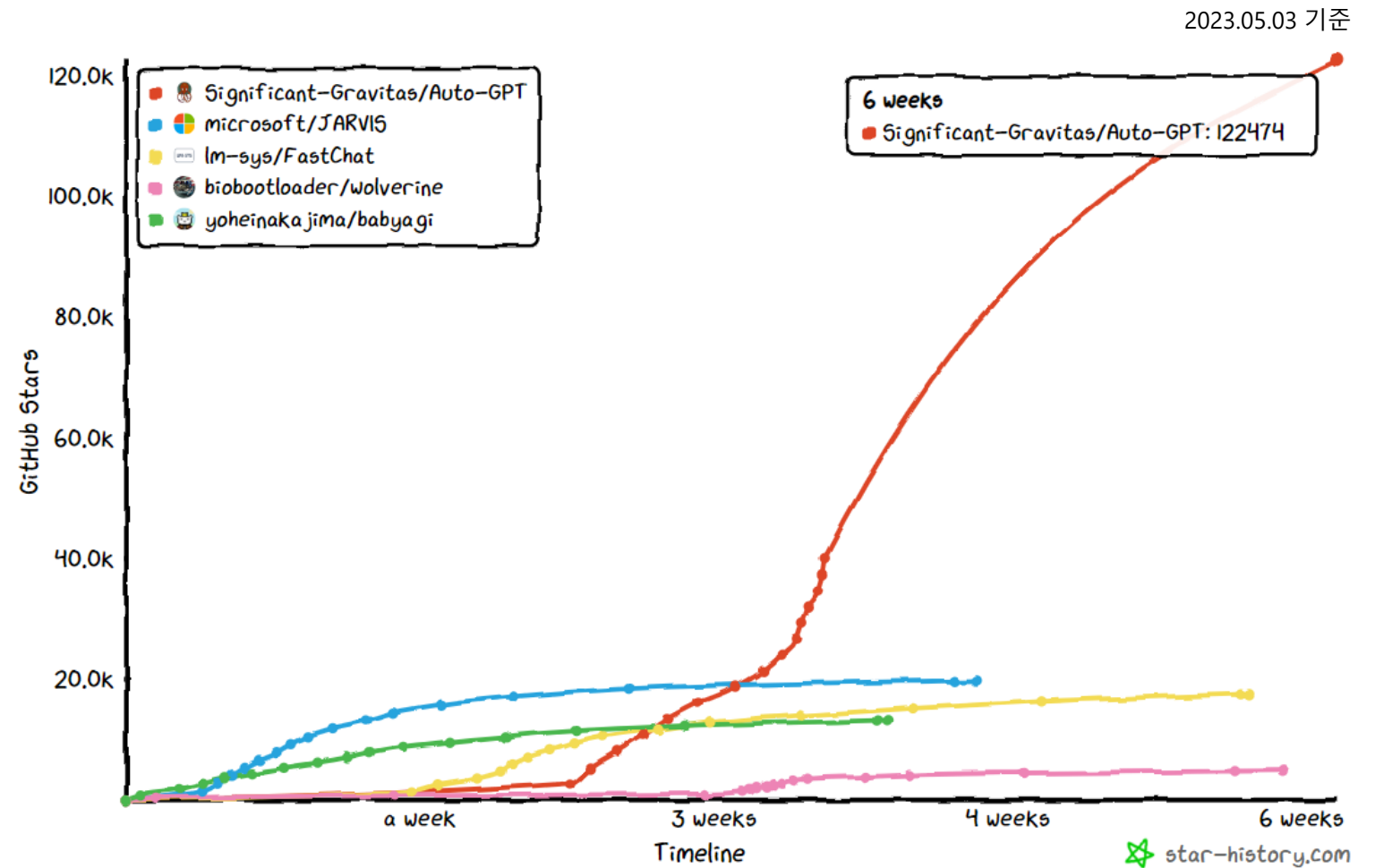
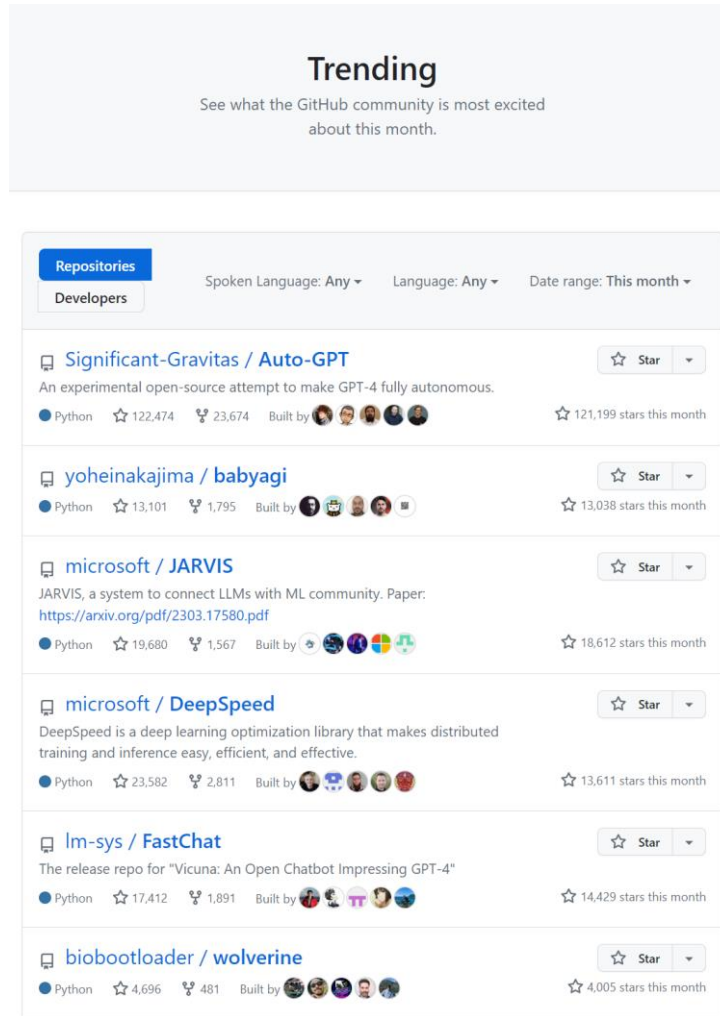
- 인간이나 동물이 수행할 수 있는 모든 지적 작업을 수행할 수 있습니다.
- 인간의 능력을 뛰어넘는 자율 시스템으로 정의될 수 있습니다.
- 인간의 능력을 뛰어넘는 대부분의 경제적 가치가 있는 작업을 수행할 수 있습니다. ¹
- 인간처럼 생각하고 추론할 수 있습니다. ²

: [Wikipedia](#)

자세한 정보: [1. news.sbs.co.kr](#) [2. technologyreview.kr](#) [3. yna.co.kr](#) [+2 더 보기](#)

3 / 20 ●

GitHub Trending



Auto-GPT

OpenAI
GPT-4

태스크 생성
및 관리 /
정보 요약



인터넷 검색



메모리

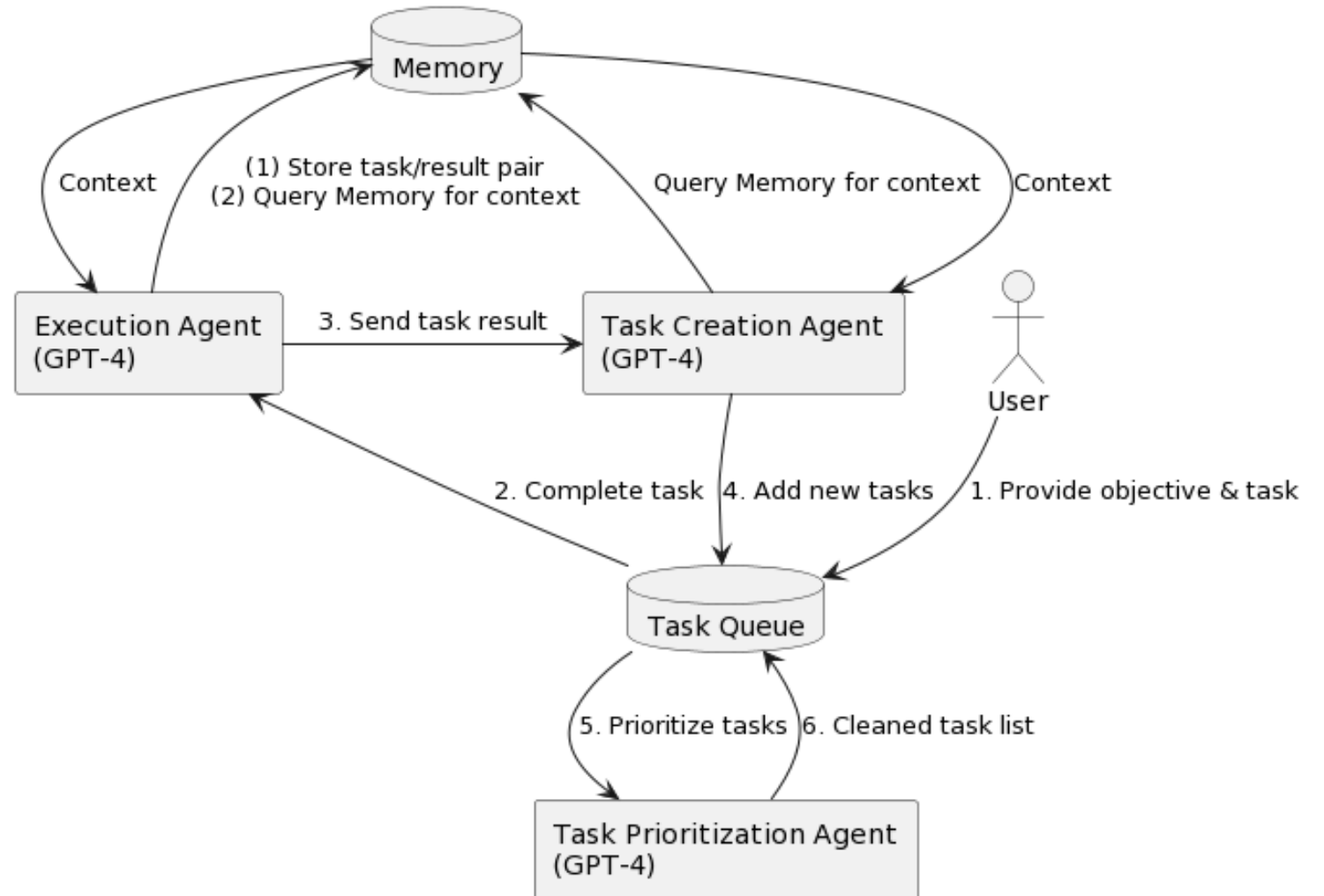
OpenAI
DALL·E



이미지 생성

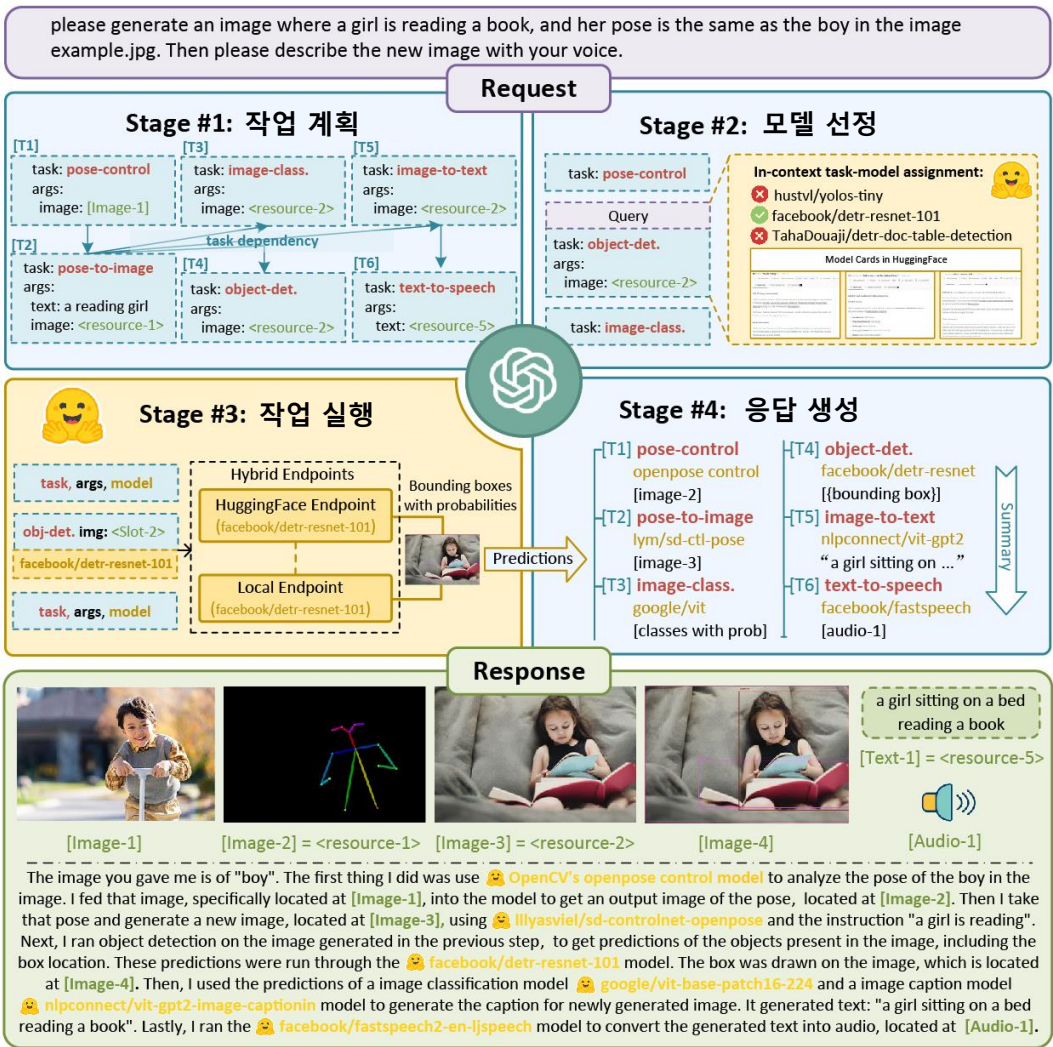
Eleven
Labs

음성 생성



Microsoft/Jarvis




OpenAI
ChatGPT
GPT-4



URL: <https://github.com/microsoft/JARVIS>

HuggingGPT


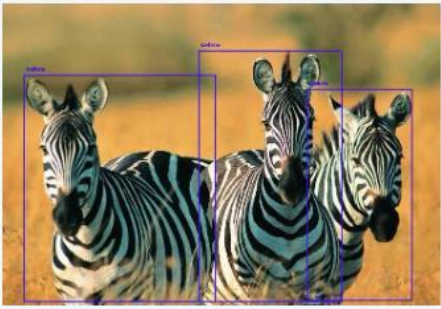
Human :
Given a collection of image A: 🖼️ [Image] , B: 🖼️ [Image] , C: 🖼️ [Image] , please tell me how many zebras in these picture?



default Submit

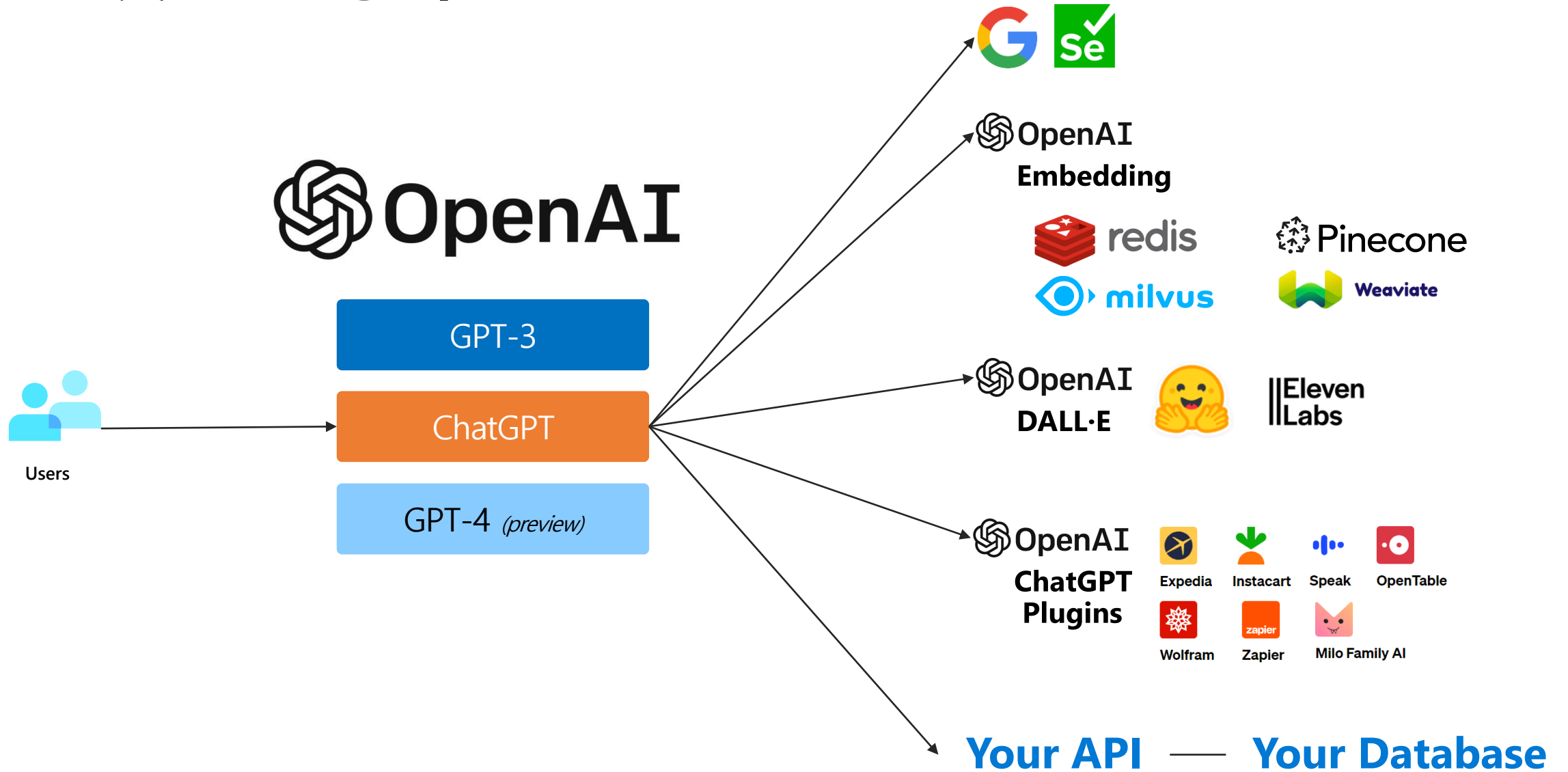
HuggingGPT

HuggingGPT :
Sure, I can do that. For your request, I have used two tasks: image-to-text and object-detection. For the image-to-text task, I have used the Salesforce/blip-image-captioning-large model, which is a Bootstrapping Language-Image Pre-training model designed for unified vision-language understanding and generation. For the object-detection task, I have used the facebook/detr-resnet-101 model, which is an End-to-End Object Detection model with ResNet-101 backbone. For the image A: 🖼️ [Image] , the inference result of image-to-text task is a close up of a cat sitting on a window sill. The inference result of object-detection task is a generated image with predicted box at 🖼️ [Image] , which contains two predicted objects: a cat with score 0.9992075562477112 and a potted plant with score 0.9939723610877991. For the image B: 🖼️ [Image] , the inference result of image-to-text task is zebras are standing in the grass and looking at the camera. The inference result of object-detection task is a generated image with predicted box at 🖼️ [Image] , which contains three predicted objects: three zebras with score 0.999295711517334, 0.995798170565588 and 0.9943941831588745 respectively. For the image C: 🖼️ [Image] , the inference result of image-to-text task is a zebra standing in a field of tall grass. The inference result of object-detection task is a generated image with predicted box at 🖼️ [Image] , which contains one predicted object: a zebra with score 0.9995033740997314. Therefore, there are four zebras in these pictures. Is there anything else I can help you with?



default Submit

AGI 첫걸음 정리



왜 Azure OpenAI로 시작해야 하는가?

Azure  OpenAI

GPT-3

Codex

ChatGPT

DALL·E *(preview)*

GPT-4 *(preview)*



Azure 구독 내에 배포, 보호, 사용자만 액세스할 수 있으며 사용자의 애플리케이션과 연결됨
프로덕션 및 솔루션 스케일링 (SLA, BCDR)



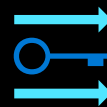
새로운 고객 시나리오를 지원할 수 있는
대규모 사전 교육 AI 모델



추가 학습 데이터로 미세 조정된 맞춤형 AI모델
Fine-tuning 모델 지원 (720시간, 20억개 토큰)



유해한 사용을 감지하고 완화하기 위한 콘텐츠
필터링 및 Responsible AI



데이터 소유권
역할 기반 액세스 제어(RBAC) 및 사설
네트워크를 통한 엔터프라이즈급 보안



Thank you