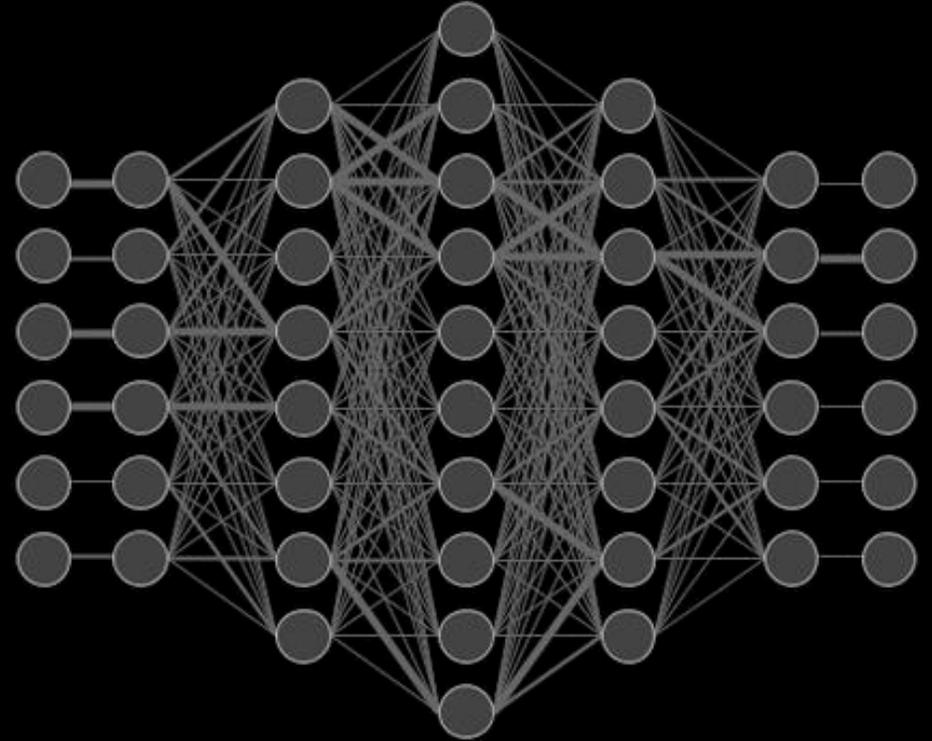


Chat-GPT / Open AI 가 제시하는
New Digital Experience



Agenda

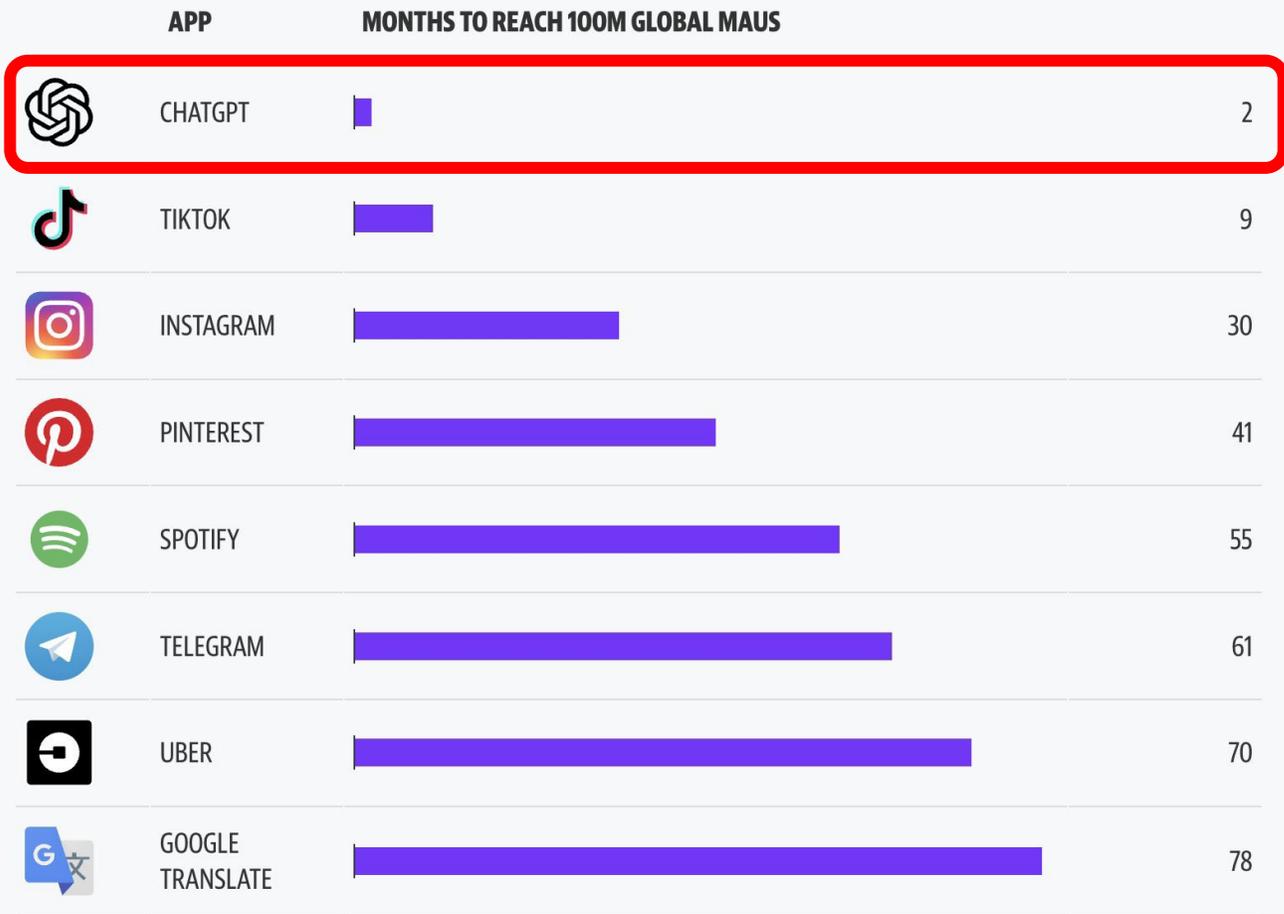
1. 생성 AI 의 시대
2. ChatGPT Overview
3. OpenAI Use Case
4. Microsoft & OpenAI



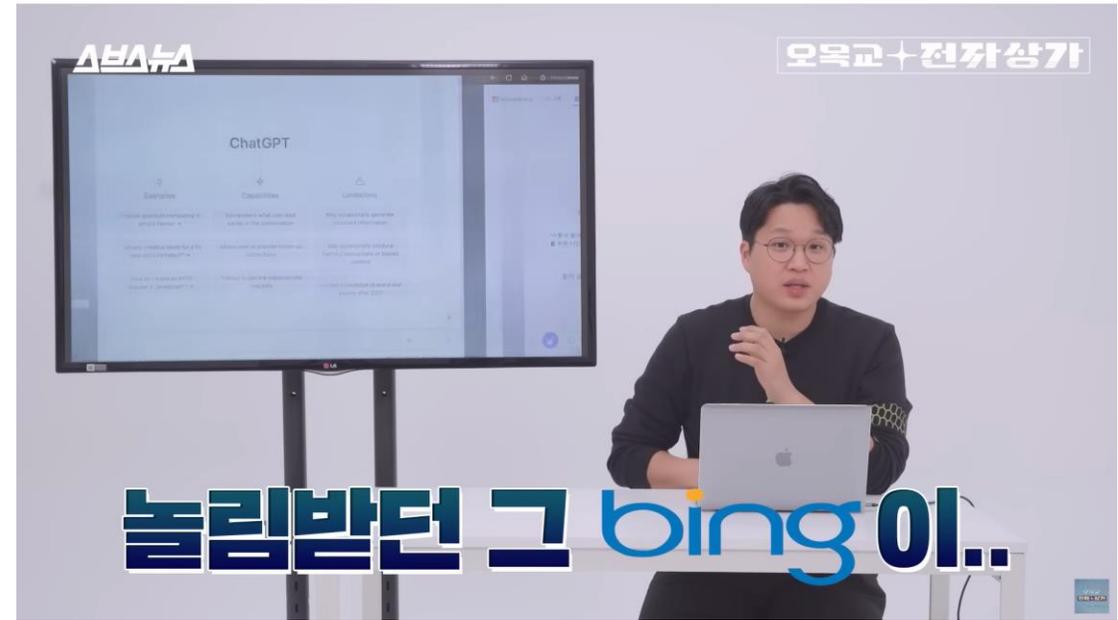
PART 1. 생성 AI의 시대

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

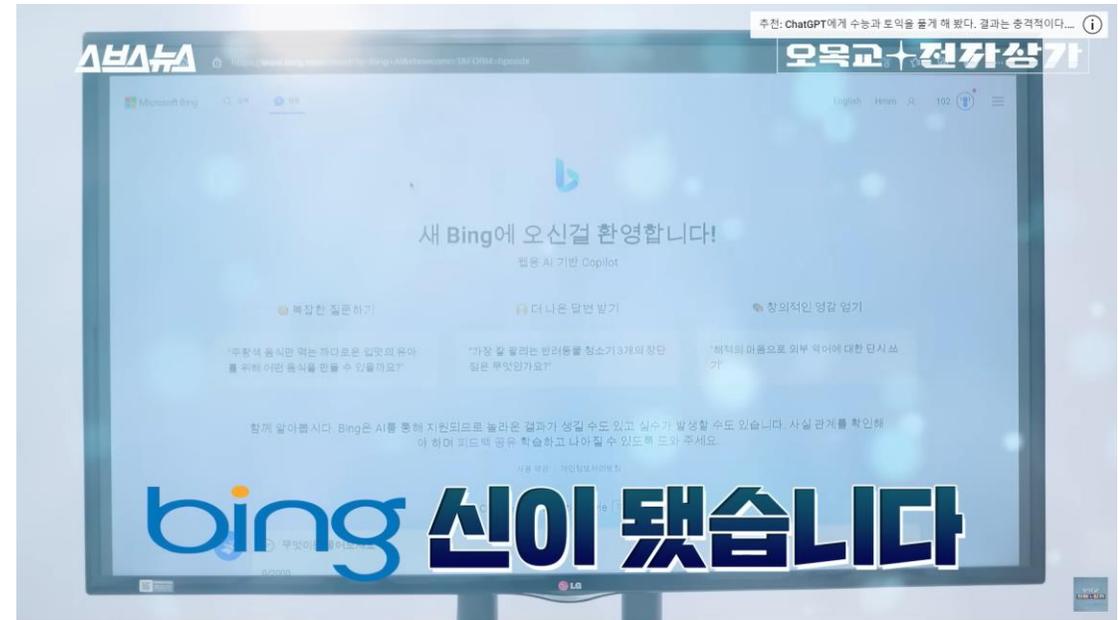
ChatGPT is estimated to have hit 100M users in January, 2 months after it's launch. Here's how long it took other top apps to reach that:



SOURCE: UBS



놀림받던 그 bing이..



Bing이 정말로 신이 됐습니다; ChatGPT는 애들 장난이고요, 구글은 정말 큰일났습니다 / 오목교 전자상가 EP128

챗GPT 품은 MS... 클라우드 시장도 흔드나

문화일보 | 입력 2023-03-06 08:56 업데이트 2023-03-06 11:15

프린트

이예린 기자

댓글 폰트 공유

새로운 Bing은 검색에 AI 채팅을 어떻게 결합했나

| 프로메테우스 모델과 새 UX로 검색 패러다임 교체 시도

홈 > 경제 > 기업·산업

'챗GPT'가 부른 AI 돌풍... 2000개 기업, 신기술 혁신의 향연

스브스뉴스 오목교 전자상가

놀림 받던 Bing과 엣지 브라우저의 반란, AI 챗봇으로 구글 제국을 뒤흔을까? [오목교 전자상가]

작성 2023.03.04 01:07 조회 4,997

프린트 글자 크기 +

이진우의 거리두기

챗GPT의 도발, '당신은 질문할 줄 아는가?'

입력 : 2023.03.08 03:00 | 수정 : 2023.03.08 09:30 | 이진우 포스텍 명예교수

The screenshot shows the Amazon Kindle Store search results for "chatgpt prompts". The top navigation bar includes the Amazon Prime logo, delivery location (New Castle 19720), and search filters. The search results list three books:

- ChatGPT Prompts Dictionary: Discover 2500+ AI Irre Quality Answers, Making Money Online And Generat**
Book 9 of 11: Transform Your Books into Profitable Cash Machines!
★★★★★ (19 reviews)
Kindle Edition
\$0.00 (Free with Kindle Unlimited membership)
Available instantly
- The Art of Prompt Engineering with ChatGPT: GPT- (Learn AI Tools the Fun Way! Book 3)**
Book 3 of 3: Learn AI Tools the Fun Way! | by Nathan Hunter |
★★★★☆ (59 reviews)
Kindle Edition
\$9.50 (was \$22.00)
Available instantly
- ChatGPT Prompts: 451+ AI Powered Prompts & Tec**
Improve Your Marketing Communications, Grow Yo
by Patrick Ejeke | Sold by: Amazon.com Services LLC | Feb 9, 2
★★★★★ (15 reviews)
Kindle Edition
\$0.00 (Free with Kindle Unlimited membership)
Available instantly

On the left side of the search results, there are filters for:

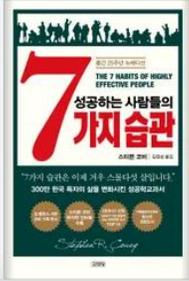
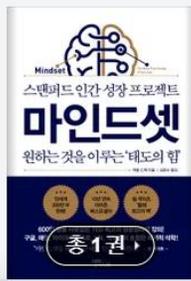
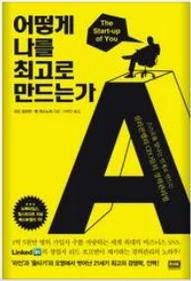
- Kindle Unlimited** (checked)
- Department**: Any Department, Kindle Store, Kindle Active Content, Kindle eBooks, Newspapers, Kindle Newsstand, Kindle Short Reads, Kindle Singles, Kindle Vella
- Customer Reviews**: 5 stars & up, 4 stars & up, 3 stars & up, 2 stars & up
- Language**: English, German, French, Spanish, Italian, Japanese, Portuguese
- Word Wise**: Word Wise Enabled (unchecked)
- New Releases**: Last 90 days

내 서재

RIDI | RIDI SELECT 

모든 책 책장 선호 작품

모든 책 검색



"어떻게 나를 최고로 만드는가?", "스탠퍼드 인간 성장 프로젝트 마인드셋 원하는 것을 이루는 태도의 힘", "성공하는 사람들의 7가지 습관" 책을 읽은 사람에게 추천해 줄 수 있는 다른 책은 어떤게 있을까?

- ✓ "'어떻게 나를 최고로 만드는가?', '스탠퍼드 인간 성장 프로젝트 마인드셋 원하는 것을 이루는 태도의 힘', '성공하는 사람들의 7가지 습관' 책 추천" 검색 중
- ✓ 답변을 생성하는 중...

"어떻게 나를 최고로 만드는가?", "스탠퍼드 인간 성장 프로젝트 마인드셋 원하는 것을 이루는 태도의 힘", "성공하는 사람들의 7가지 습관" 책을 읽은 사람에게 추천해 줄 수 있는 다른 책으로는 "책쓰기 제목을 잘 짓는 방법 10가지"가 있습니다¹. 또한, "성공하는 사람들의 7가지 습관"의 저자인 스티븐 R. 코비의 다른 책인 "효과적인 사람들의 7가지 습관"도 추천드립니다². 이 책은 "성공하는 사람들의 7가지 습관"과 마찬가지로 성공적인 사람들의 행동과 습관을 분석하고, 그들의 성공 비결을 알려줍니다². 또한, "어떻게 나를 최고로 만드는가?"와 관련된 책으로는 "최고의 인생을 만드는 법"이 있습니다³. 이 책은 자신의 인생을 살아가는 방법과 자신의 잠재력을 극대화하는 방법을 알려줍니다³.

자세한 정보: [1. brunch.co.kr](http://1.brunch.co.kr) [2. ko.wikipedia.org](http://2.ko.wikipedia.org) [3. blog.naver.com](http://3.blog.naver.com)

5 / 15

"최고의 인생을 만드는 법" 책을 RIDI에서 찾아줄 수 있어?

- ✓ "최고의 인생을 만드는 법 책 RIDI" 검색 중
- ✓ 답변을 생성하는 중...

네, "최고의 인생을 만드는 법" 책은 리디에서 구매하실 수 있습니다¹. 이 책은 뇌과학, 심리학, 철학 등 다양한 분야의 지식을 바탕으로 자신의 인생을 살아가는 방법과 자신의 잠재력을 극대화하는 방법을 알려줍니다¹.

자세한 정보: [1. ridibooks.com](http://1.ridibooks.com) [2. ridibooks.com](http://2.ridibooks.com) [3. ridibooks.com](http://3.ridibooks.com)

6 / 15



리디 고객센터

[채팅운영] 평일 10:00~17:00(점심 12:50~14:00)

ID/PW찾기

RIDI Bot

ID만 분실

아래 경로에서 가입했던 이메일을 입력하시면 ID 일부를 알 수 있고 메일로 전체 ID를 보내드려요 :)

[아이디 찾기]

PW 분실

아래 경로에서 가입했던 이메일을 입력하시면 비밀번호 재설정 정보를 메일로 보내드려요 :)

• • 도움이 되었나요?

패스워드 찾기

RIDI Bot

작품/콘텐츠 관련 문의하실 내용을 선택해주세요

• •

패스워드 분실

RIDI Bot

회원계정 관련 문의하실 내용을 선택해주세요. 📌

• •



Type a message

Assistant setup

Empty Example

Save changes 코드 보기

System message ①

너는 리디의 고객센터 챗봇 서비스야. 너는 고객의 문제를 도와줘야해. 리디 고객을 위해서 상냥하고 부드럽게 이야기 해줘. 리디 이외의 질문에 대해서는 답변하지 않아도 돼.

Few-shot examples ①

User:

아이디를 잃어 버렸어.

Assistant:

아래 경로에서 가입했던 이메일을 입력하시면 ID 일부를 알 수 있고 메일로 전체 ID를 보내드려요 :)
<https://ridibooks.com/account/find-id>

User:

패스워드를 잃어 버렸어.

Assistant:

아래 경로에서 ID, 가입했던 이메일을 입력하시면 비밀번호 재설정 정보를 메일로 보내드려요 :)
<https://ridibooks.com/account/find-password>

User:

포인트 무료충전소에 대해서 알려줘.

Assistant:

1. 포인트 무료충전소는 홈 우측 상단에서 확인할 수 있습니다.
2. 포인트 지급을 위해 이벤트 참여 전 앱 추적 허용 및 개인정보 이용동의가 필요합니다.

+ Add few-shot examples

Chat session

채팅 지우기 Show raw JSON

패스워드를 잃어 버렸어.



아래 경로에서 ID, 가입했던 이메일을 입력하시면 비밀번호 재설정 정보를 메일로 보내드려요 :)
<https://ridibooks.com/account/find-password>

아이디도 기억이 안나



가입했던 이메일로 아이디를 찾을 수 있어요. 아래 경로에서 이메일을 입력하시면 가입된 아이디를 확인하실 수 있습니다.
<https://ridibooks.com/account/find-id>

무료로 포인트를 모을 수 있어?



리디북스에서는 다양한 이벤트를 통해 무료 포인트를 지급해드리고 있어요. 포인트 무료충전소에서 이벤트를 확인하시면 더 많은 정보를 얻으실 수 있습니다. 또한, 리디셀렉트에서는 매일 무료로 책을 받을 수 있는 이벤트도 진행하고 있으니, 매일 리디셀렉트를 방문하시면 더 많은 포인트를 모으실 수 있습니다.

포인트 무료충전소는 어떻게 이용할 수 있어?



포인트 무료충전소는 리디북스 홈페이지 우측 상단에 위치하고 있습니다. 무료 포인트 이벤트에 참여하시려면, 먼저 앱 추적 허용 및 개인정보 이용동의를 하셔야 합니다. 그 후, 이벤트 참여 버튼을 누르시면 무료 포인트를 받으실 수 있습니다. 이벤트에 따라서 참여 조건과 방법이 다르니, 이벤트 페이지에서 자세한 내용을 확인해주세요. 이벤트 참여 후 포인트가 적립되면, 리디북스 내에서 책을 구매하거나 대여하는데 사용할 수 있습니다.

여기에 사용자 쿼리를 입력하세요. (줄 바꿈 입력: Ctrl + Enter)

보내기

Artificial Intelligence

Machine Learning

Deep Learning

Generative AI



Artificial Intelligence

인간의 지능을 복제하거나 능가할 수 있는 지능형 기계를 만들고자 하는 컴퓨터 과학 분야



Machine Learning

기계가 기존 데이터에서 학습하고 해당 데이터를 개선하여 의사 결정 또는 예측을 할 수 있도록 하는 AI의 하위 집합



Deep Learning

신경망 레이어를 사용하여 데이터를 처리하고 결정을 내리는 기계 학습 기술



Generative AI

프롬프트나 기존 데이터를 기반으로 새로운 문서, 시각, 청각 콘텐츠를 생성하는 기술

Deep Learning 시대

인프라 구축

인프라 운용

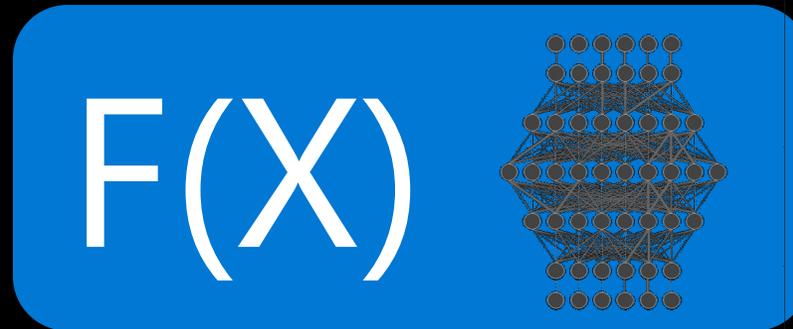
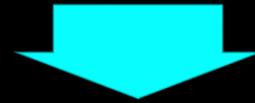
ML 환경 세팅

대량의 데이터 저장소

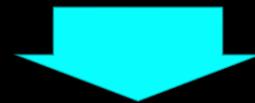
데이터 세트 모으기

양질의 데이터 세트 만들기

Input Data



Output Data



적합한 알고리즘

ML 학습 시간

ML 학습 비용

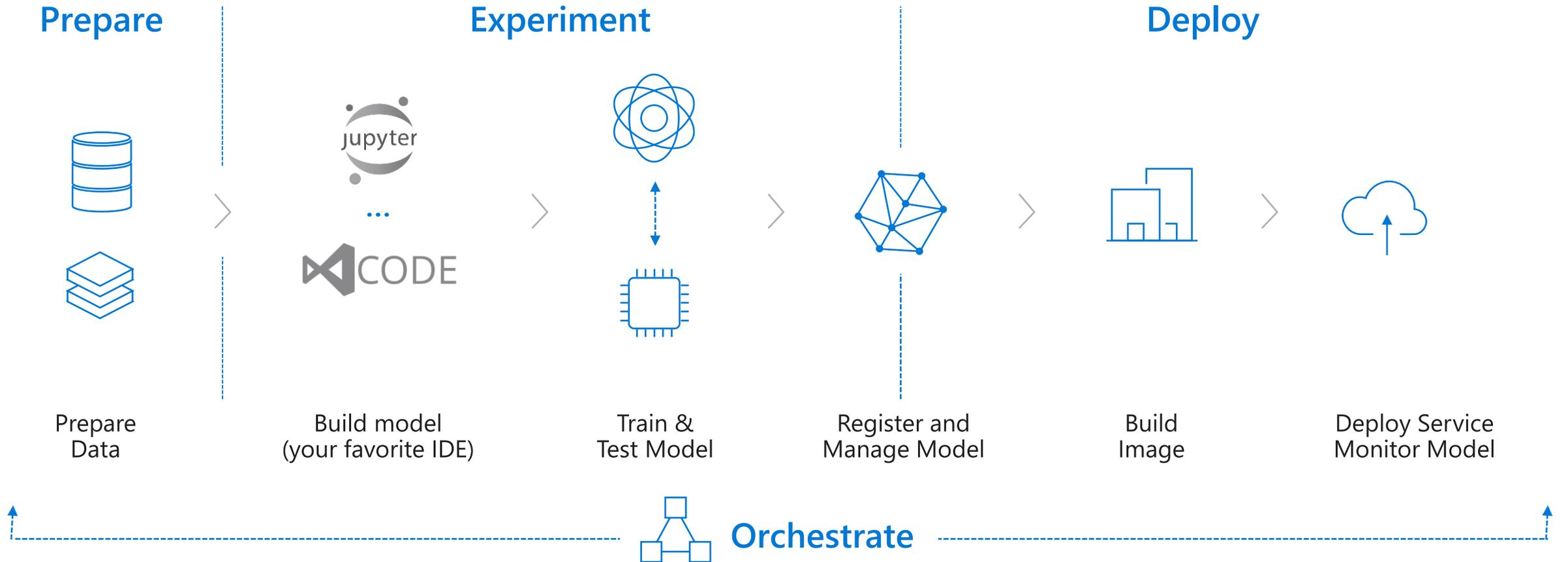
MLOps 환경 구축

추론 환경 AutoScaling

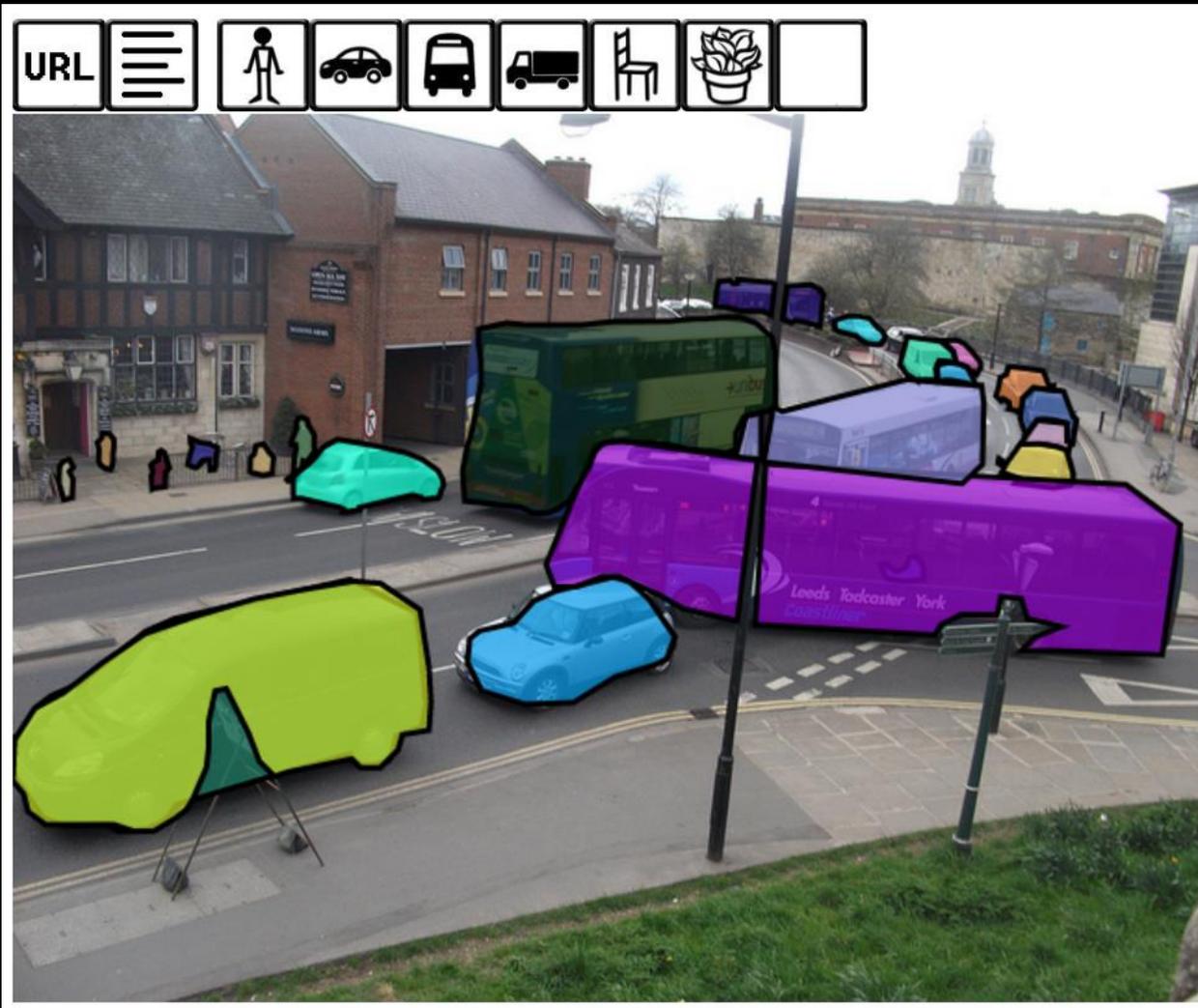
인력 수급

Machine Learning

Typical E2E Process



MSCOCO Dataset – 데이터에 진심!



- traffic tie up at intersection in city area.
- bus merging through traffic at a busy intersection.
- many busses drive down a busy street.
- a bus driving down a street next to a bunch of traffic.
- a bus pulling out onto a busy road.

Generative AI 시대

Generative : 생성 | Pre-trained : 사전 학습된 | Transformer : 변환

Do more with less!

인프라 구축

인프라 운용

ML 환경 세팅

대량의 데이터 저장소

데이터 세트 모으기

양질의 데이터 세트 만들기

적합한 알고리즘

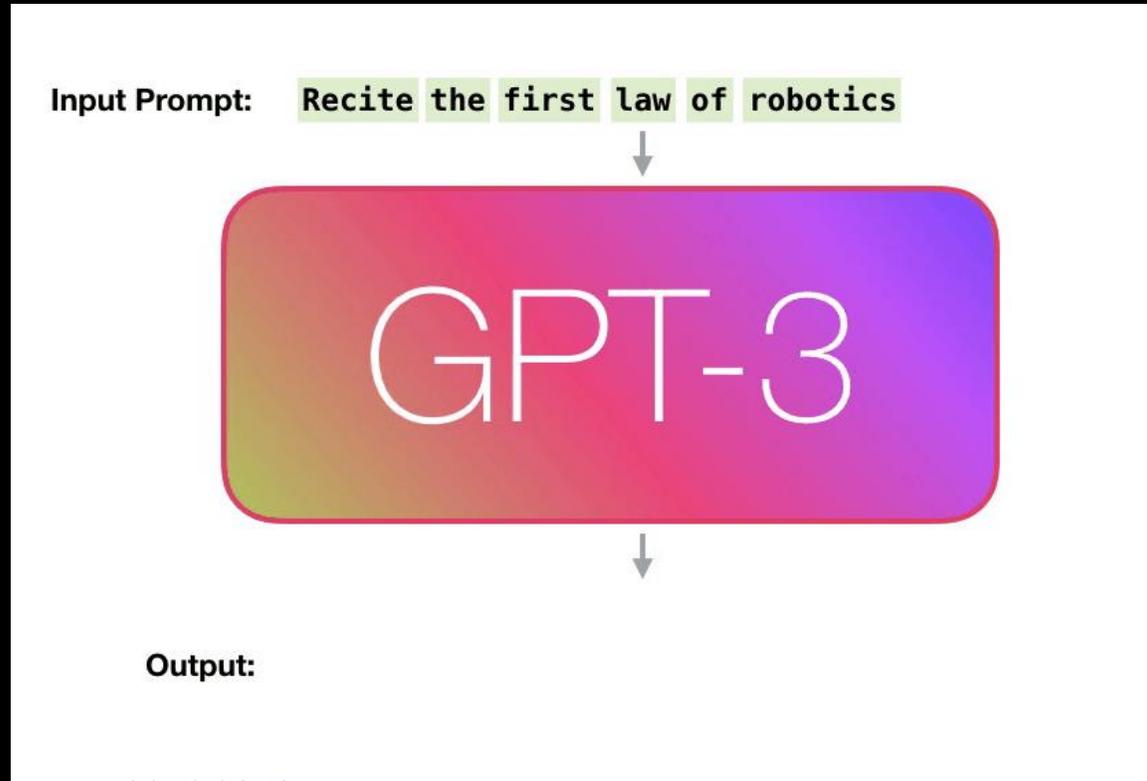
ML 학습 시간

ML 학습 비용

MLOps 환경 구축

추론 환경 AutoScaling

인력 수급



이진우의 거리두기

챗GPT의 도발, '당신은 질문할 줄 아는가?'

입력 : 2023.03.08 03:00 | 수정 : 2023.03.08 09:30 | 이진우 포스텍 명예교수



PART 2. ChatGPT Overview

AI의 발전 방향 (Generative AI : Language)

A powerful new class of **large language models** is making it possible for machines to **“write, code, draw and create”** with credible and sometimes superhuman results

Wave #1 : 텍스트의 이해 및 처리

순환 신경망 기반, 소규모 문장 처리
단순 분류 중심의 역할 수행

RNN / LSTM

핵심어 검출 / 텍스트 생성

Wave #2 : 자연어 처리방식의 혁신

긴 문장의 맥락을 이해할수있는 모델
“Attention is All You Need” - Transformers

BERT / GPT

맥락 / 의미 이해 - 독해 / 문장생성

Wave #3 : 최적화/대형화를 통한 생성 AI 탄생

클라우드 기반 ‘대형 언어 모델’, ‘컴퓨팅 자원 최적화’, ‘알고리즘 최적화’, ‘속도 최적화’와 더불어 Chat GPT와 같은 Killer App 등장 (파라미터 개수 규모에 따라 성능 증가)

GPT-3 / MT-LNG

자연어 복합 문장에 대한 이해 / 자연어 생성

What Next?

- 모델 : 대형화 !! vs 최적화 !!
- 대상 : Text/ Code/ Image → Speech/ Video/ 3D

ChatGPT

대화에 최적화된 언어 모델

대화식으로 상호 작용합니다. 대화 형식을 통해 ChatGPT는 후속 질문에 답하고, 실수를 인정하고, 잘못된 전제에 이의를 제기하고, 부적절한 요청을 거부할 수 있습니다.

DALL·E

자연어로 표현할 수 있는 다양한 개념에 대해 **텍스트 캡션에서 이미지를 생성**하는 DALL·E라는 신경망입니다.

Codex

자연어를 코드로 변환
(Python, JavaScript, Go, Perl, PHP, Ruby, Swift, TypeScript, Shell 등)

Image GPT

A large transformer model trained on pixel sequences can **generate coherent image completions and samples**. By establishing a correlation between sample quality and image classification accuracy, our best generative model also contains features competitive with top convolutional nets in the unsupervised setting.

Jukebox

A neural net that generates music, including rudimentary singing, as raw audio in a variety of genres and artist styles. We're releasing the model weights and code, along with a tool to explore the generated samples.

- AI 연구 및 개발 회사
- 미션: 인공지능으로 인류에게 도움을 주자

 AI

ChatGPT란?



ChatBot +

Generative : '생성하는' 문장

Pre-trained : '사전 학습된' 사전 학습된 정보를 기반한

Transformer : '변환' 글의 맥락을 이해할 수 있는 언어 모델

A screenshot of the ChatGPT interface. At the top, it says "ChatGPT". Below that, there are three columns: "Examples", "Capabilities", and "Limitations". Each column has a list of items. The "Examples" column includes: "Explain quantum computing in simple terms" →, "Get any creative ideas for a 10-year old's birthday?" →, and "How do I make an HTTP request in Javascript?" →. The "Capabilities" column includes: "Remembers what user said earlier in the conversation", "Allows user to provide follow-up corrections", and "Trained to decline inappropriate requests". The "Limitations" column includes: "May occasionally generate incorrect information", "May occasionally produce harmful instructions or biased content", and "Limited knowledge of world and events after 2021". At the bottom, there is a text input field and a send button. A small footer at the very bottom reads: "ChatGPT Feb 13 2023: Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve."

사용 예

- 주제 요약 / 설명
- 브레인스토밍
- 코드 작성

주요 기능

- 이전 대화를 기억하여 답변
- 적절하지 않은 질문은 거르고 답변
- 추가 질의를 통해 답변을 개선

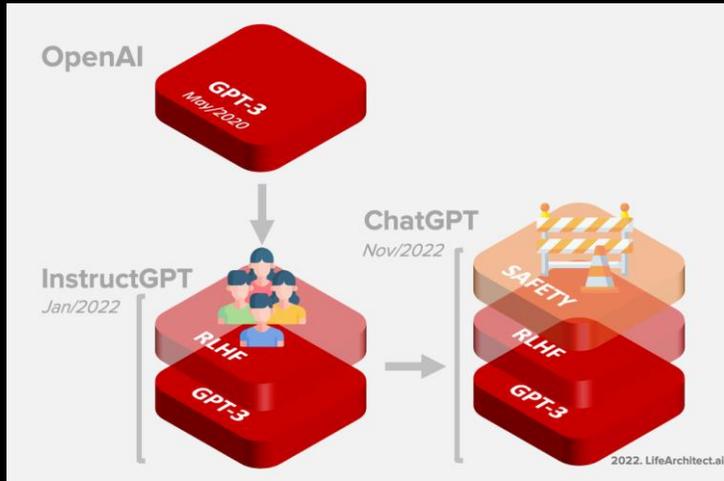
한계

- 부정확한 데이터를 답변할 수 있다
- 부적절한 콘텐츠를 생성할 수도 있다
- 2021년 이후 데이터에 대한 지식 미보유

ChatGPT 는 어떻게 만들었나?



사람의 피드백을 통한 강화 학습
(RLHF + Safety)



Step 1. 샘플 데이터를 기반으로 지도학습 수행

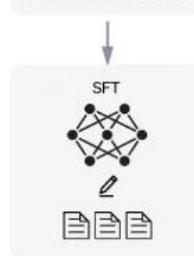
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.

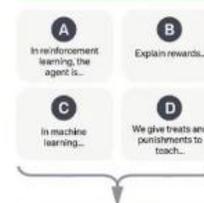
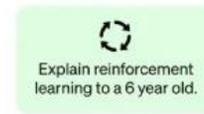


This data is used to fine-tune GPT-3.5 with supervised learning.



Step2. 비교 데이터 수집 / 보상 모델 훈련

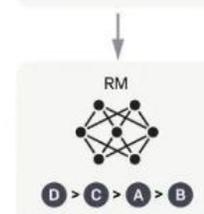
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.

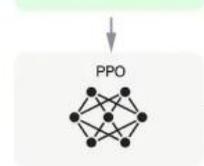


Step 3. 강화학습 모델을 활용하여, 보상모델 최적화

A new prompt is sampled from the dataset.



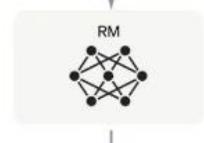
The PPO model is initialized from the supervised policy.



The policy generates an output.



The reward model calculates a reward for the output.

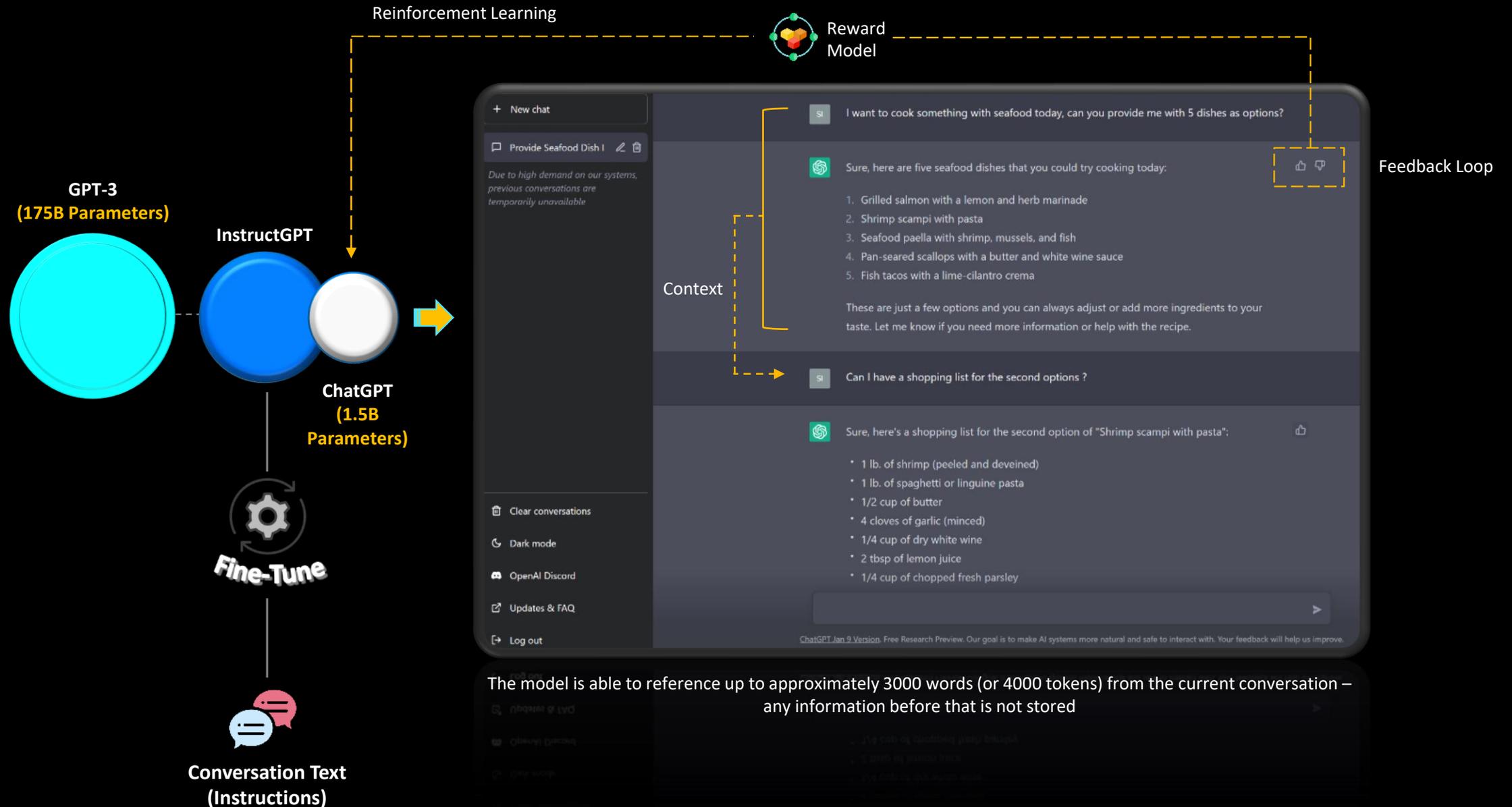


The reward is used to update the policy using PPO.



GPT-3라고 하는 Large Language Model 을 기반으로 함. (1750억개의 파라미터)

ChatGPT



GPT-3: Make smarter apps



On-demand 인텔리전스를 위해 API를 통해 액세스할 수 있는 AI 모델

ChatGPT: Get more productive



상호 작용하고 작업 수행을 요청할 수 있는 챗봇을 구축하는 데 사용되는 GPT-3 모델

언어 생성

대용량 텍스트 코퍼스

1750억 개의 파라미터

느림

필요함

목적

훈련 데이터

모델 크기

추론 속도

파인튜닝 필요 여부

대화형 인공지능

대화 데이터

15억 개의 파라미터

빠름

필요 없음

Can be used for:



Semantic 텍스트 이해 구현



내부 정보 검색 및 추출



Copilot과 같은 애플리케이션 구축



Can be used for a lot more including what ChatGPT can do



콘텐츠 제작을 위한 아이디어 구상



일반 질문 답변



코드 생성 및 대화 지원

ChatGPT 에 DALL E를 추가 활용하여 동화 제작 예시



한걸 추위를 싫어하는 백곰의 북극 탈출기를 동화로 만들어줘, 1000자 이내로

생성

옛날 옛적, 북극에는 추위를 싫어하는 백곰이 살고 있었습니다. 매년 겨울이 오면, 백곰은 추운 날씨 때문에 우울해지곤 했습니다. 그러던 어느 날, 폭설과 추위로 인해 백곰의 마을은 위험한 상황에 놓였습니다.

이에 대처하기 위해 백곰은 북극 여우, 북극 곰, 북극 물개, 그리고 북극 토끼와 함께 모여 마을을 구하기로 결심했습니다. 첫 번째로 백곰은 바람을 잡는 기술이 뛰어난 북극 여우를 만났습니다. 북극 여우는 빠른 속도로 바람을 잡아 마을로 전해주어 모든 동물들이 따뜻해질 수 있었습니다.

다음으로 백곰은 추위를 이기기 위한 몸과 인내력을 가진 북극 곰을 만났습니다. 북극 곰은 마을 주변에 불을 피워서 모든 동물들이 따뜻해질 수 있도록 도와주었습니다.

백곰은 또한 바다에 사는 북극 물개를 만났습니다. 북극 물개는 바다 속에서 먹이를 잡는 데에 뛰어나며, 그의 능력은 마을의 음식 공급에 큰 도움이 될 것이라 생각했습니다.

마지막으로, 백곰은 스스로 눈을 쌓는 기술이 뛰어난 북극 토끼를 만났습니다. 북극 토끼는 눈을 쌓아서 모든 동물들이 길을 따라 마을로 돌아갈 수 있도록 도와주기로 약속했습니다.

이들의 협력으로 백곰은 모든 동물들이 따뜻하고 안전하게 지낼 수 있는 마을을 만들었습니다. 이제부터는 백곰도 겨울에 대해 두려워하지 않게 되었고, 그는 북극에 사는 다른 동물들과 함께 겨울을 이겨낼 수 있는 힘을 얻었습니다. 마을은 평화롭게 지내고, 백곰과 친구들은 영원한 우정을 맺었습니다.

한걸 제목을 만들어줘

요약

"추위를 이기는 우정의 힘: 백곰의 북극 탈출기"

👍 🗨

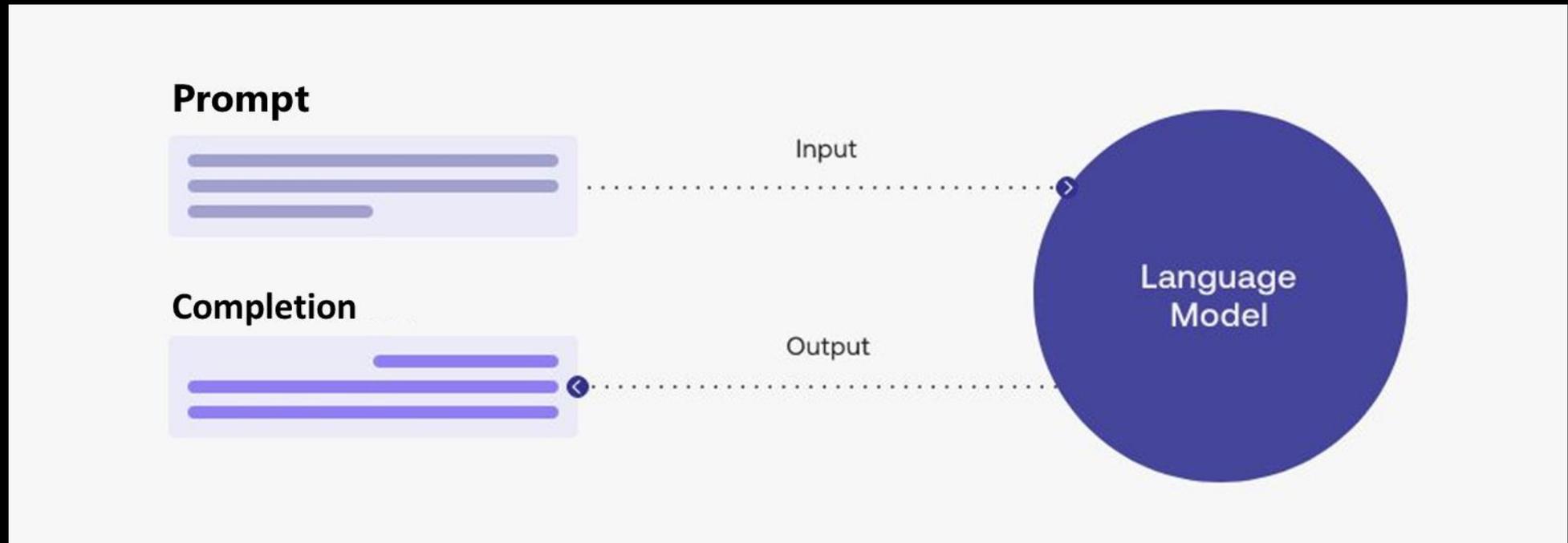
White bear who hates the cold in the ice field, a fairy tale illustration
(빙원에서 추위를 싫어하는 백곰, 동화의 삽화)



ChatGPT 활용 방법



프롬프트(Prompt)는 거대 언어 모델(Large Language Model; LLM)로부터 응답을 생성하기 위한 입력값을 의미
답변되는 텍스트 값을 **Completion** 이라고 명명함.



※ 향후 Prompt Engineering이라는 영역이 하나의 전문영역으로 각광 받을 것으로 예상됩니다.

Azure OpenAI | 언어 모델을 어떻게 모델 최적화를 할수 있나?

특별한 예시 없이 자연어 기반
작업 지시 - Zero shot learning

Prompt

너가 기자라면, 제목을 어떻게 만들지
제안해줘

마이크로소프트가 현지시간 21일 애저 오픈AI 서비스(Azure OpenAI Service)에 GPT-4 모델을 추가했다. 최근 발표된 새로운 Bing(Bing) 및 코파일럿(Microsoft 365 Copilot)은 이미 GPT-4로 구동되고 있다. 이날부터 오픈AI 서비스를 사용하고 있는 고객과 기업은 GPT-4을 활용하여 그 어느 때보다 더 진보된 방식으로 자체 앱과 서비스를 구축할 수 있다. 예를 들어, 오픈AI 서비스를 통해 봇 개발자가 파워 가상 에이전트(Power Virtual Agents)의 코파일럿에서 자연어를 사용하여 단 몇 분 만에 인공지능 가상 도우미를 만들 수 있다.

Sample response

기사 제목
1."마이크로소프트, 애저 오픈AI 서비스에 GPT-4 모델 추가"
2."GPT-4를 활용한 자체 앱 및 서비스 구축 가능해진다"
3."봇 개발자들을 위한 마이크로소프트의 새로운 인공지능 가상 도우미 출시"

작업 지시 및 샘플 제공
- One / Few shot learning

Prompt

게임 해설을 하이라이트로 바꿔줘

(예시)

해설: What a pickup she has

하이라이트: New York has domina

해설

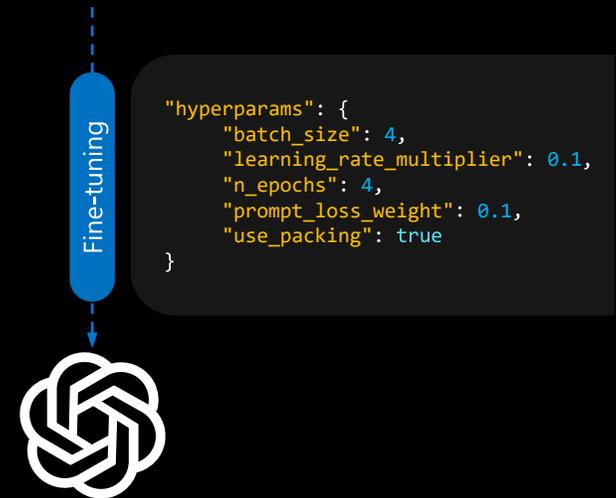
- 1. Turner is so important defensively to
- 2. Griner pulled way out, Hartley with
- 3. At 1:54 remaining in the quarter, Pho

Sample response

하이라이트:
1. New York has had a strong run in the
2. Phoenix leading by 1 point, 24-23
3. New York Liberty's comeback has be

다량의 작업지시와 샘플을 입력하여,
모델 최적화- Fine Tuning

Prompt and completion examples



{"prompt": "<prompt text>", "completion": "<ideal generated text>"
{"prompt": "<prompt text>", "completion": "<ideal generated text>"
{"prompt": "<prompt text>", "completion": "<ideal generated text>"

← 사용자 할수 있는 최적화 (Prompt Engineering)

← 모델 관리자의 최적화 영역(Finetuning)

ChatGPT 200% 활용 예시



새로운 프로젝트를 맞춤형으로 제작

프롬프트:

Financial Times의 제품 관리자로 활동하십시오. 속보 및 심도 있는 뉴스에 대한 선호도와 온라인 뉴스에 대한 비용 지불 의향을 기준으로 청중을 분류하는 것을 목표로 하는 연구 프로젝트를 계획합니다.

- 10~20대를 대상으로 한 온라인 방송을 어떻게 만들 수 있을까?
- 방송국의 프로듀서 역할일 때, IT 속보 혹은 심층뉴스에 대한 선호도를 가지는 계층에 대해 비용지불을 유도할 수 있는 방송 계획을 수립해줘
- 참가자의 독서습관, 심층뉴스 속보 선호도 및 온라인 뉴스에 대한 비용 지불 의지를 조사할 수 있는 설문 문항을 10개 만들어줘



AI in action
How ChatGPT would research this new product?

We asked an AI model to plan a study to segment audiences of the Financial Times

Analysis by ChatGPT prompted by G. Piechota for INMA 2023



- 01 Develop a survey. Ask about habits, needs
- 02 Recruit current and prospect readers
- 03 Segment readers using statistical methods
- 04 Identify segments and characterise them
- 05 Validate the segments using qualitative research
- 06 Develop a strategy to target each segment
- 07 Implement the strategy and monitor results

[INMA: What happened when we asked ChatGPT to develop a new product for Fina...](#)

| 모델 최적화 1. Prompt / Tokens / Hyper Parameter 활용 예시

Prompt Instruction

Suggest three names for an animal that is a superhero.

Animal: Cat

Names: Captain Sharpclaw, Agent Fluffball, The Incredible Feline

Animal: Dog

Names: Ruff the Protector, Wonder Canine, Sir Barks-a-Lot

Animal: Horse

Names:

Tokens (모델에서 인식하는 텍스트 단위)

I have an orange cat named Butterscotch.

I have an orange cat named Butterscotch.

<https://platform.openai.com/tokenizer>

Hyper Parameter (사용자가 조정할 수 있는 기능)

- Temperature : 창의성 범주를 설정 (0 ~ 1)
- Top-p : 다음 단어 생성시 연관 단어 후보군 범위 설정 (0 ~ 1)
- Beam width : 모델이 가장 적합한 문장을 선택하는 방법 (0 ~ 1)

Horses are my favorite

animal 49.65%

animals 42.58%

\n 3.49%

! 0.91%

IF TEMPERATURE IS 0

Horses are my favorite animal

IF TEMPERATURE IS 1

Horses are my favorite animal

Horses are my favorite animals

Horses are my favorite !

Horses are my favorite animal

튜닝2 모델 최적화 Fine-Tuning 이란?

나의 데이터를 추가로 학습시켜 커스텀 모델을 생성하여 활용

- GPT-3의 튜닝을 수행할 경우, 최소 수천개의 예제를 제공해야 하며, MAX 2.5M → 2B 개의 토큰, 120 → 720시간의 훈련시간을 허용
- 프롬프트보다 더 나은 품질의 결과
- 프롬프트보다 더 많은 예를 훈련
- 프롬프트를 줄여 토큰을 절약
- 더 빠른 응답 성능
- 모델의 가중치와 편향이 반영됨 (few-shot learning과 대조적임)

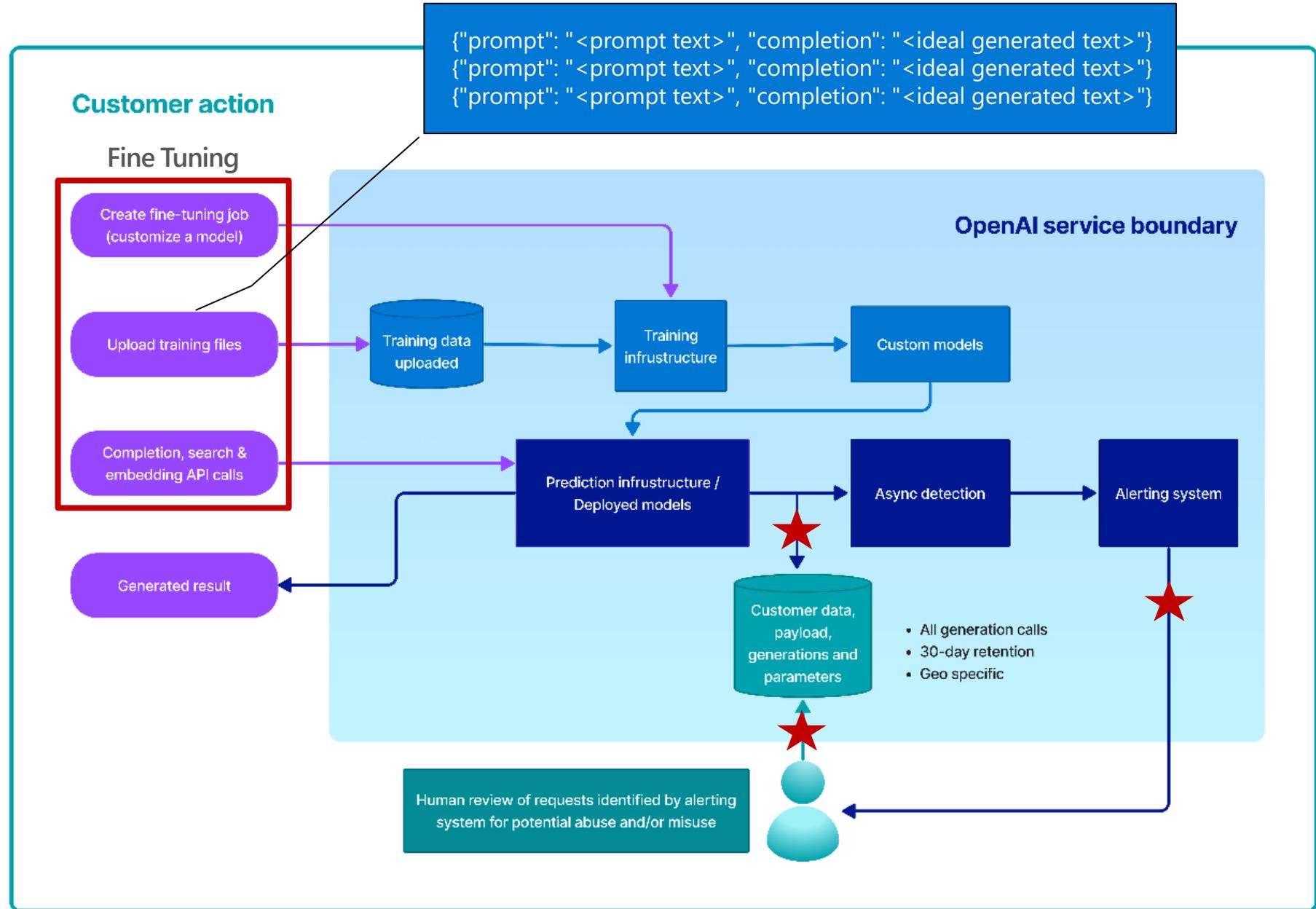
※Data Privacy / Security 정책



Responsible AI 정책에 의거,
로그/ 데이터 감사를 수행합니다.

단, 사용시나리오 협의 및 법적 요건
리뷰 후 해당 기능을 비활성화할 수 있습니다.

[Azure OpenAI Limited Access Review](#)



AI 사용자의 활용 역량이 중요해집니다.



- 어떤 Prompt를 작성해야, 왼쪽과 같은 그림을 만들 수 있을까요?
- 오른쪽 그림에 해당되는 Completion을 얼마나 빠르게 생성할수있을까요?



Prompt

a painting of a fox sitting in a field at sunrise in the style of cartoon



PART 3. Open AI Use Case

Azure OpenAI | Top 4 Capabilities & Use Cases



Content Generation

컨택 센터 : 고객 질의에 대한 맞춤형 답변 생성

웹 : 맞춤형 UI 생성 (개인화)



Summarization

컨택 센터 : 콜 로그에 대한 요약

전문 영역 주제에 대한 요약 (회계/ 마케팅 보고서, SNS 外)



Code generation

개발자 / 데이터 관리자 : 코드 및 SQL 쿼리 등 작성

운영자 : 각종 스크립트 작성



Semantic Search

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 / 콘텐츠 생성

예시 1

Cognitive Services | Azure OpenAI Studio

Azure OpenAI Studio > GPT-3 플레이그라운드

GPT-3 플레이그라운드

배포 예

hs-gpt-model Summarize Text

코드 보기

텍스트 요약 예시는 text-davinci-002 모델에서 가장 잘 작동합니다. 배포 만들기

상점에서 캐릭터를 구매하였는데, 선택하지 않았던 다른 캐릭터가 구매되었습니다. 게임 업데이트 직후 새로운 캐릭터가 추가되면서 캐릭터가 밀리는 증상이 발생한 것 같습니다. 화면에서는 영희라는 캐릭터를 구매했는데, 구매한 캐릭터 내역을 보니 철수가 들어와 있습니다. 철수 캐릭터는 원하지 않았으며, 이를 취소하거나 원하는 캐릭터인 영희로 보상해 주기를 희망합니다. 그리고 이 문제가 재발하지 않도록 내부 프로세스를 개선해 주세요.

고객의 대화에서 다음을 추출해.

- 이 대화의 주된 이유:
- 고객의 감정:
- 문제가 발생한 시점:
- 고객이 원하는 보상:
- 고객 응대 방법:
- 고객 응대를 위한 Email 콘텐츠를 HTML 포맷으로 작성:

1. 이 대화의 주된 이유: 상점에서 잘못된 캐릭터가 구매한 것에 대한 불만사항
2. 고객의 감정: 불만
3. 문제가 발생한 시점: 게임 업데이트 직후
4. 고객이 원하는 보상: 잘못된 캐릭터를 취소하거나 원하는 캐릭터인 영희로 보상
5. 고객 응대 방법: 내부 프로세스를 개선하여 문제가 발생하지 않도록 하고, 고객이 원하는 보상을 지급하는 방법
6. 고객 응대를 위한 Email 콘텐츠를 HTML 포맷으로 작성:
<p>안녕하세요. 최근 귀하께서 상점에서 구매하신 캐릭터가 잘못된 것으로 판명되었습니다. 이에 대해 귀하께 깊이 사과 드리며, 해당 문제가 발생하지 않도록 내부 프로세스를 개선하고 있습니다. 또한, 귀하께서 원하는 보상이 필요할 경우 철수 캐릭터를 취소하거나 원하신 영희 캐릭터로 보상해 드리도록 하겠습니다.<p> 감사합니다.

생성 토큰: 1609

Parameters

온도 0.82

최대 길이(토큰) 1257

중지 시퀀스

중지 시퀀스

상위 확률 1

빈도 페널티 0

현재 상태 페널티 0

최상의 1

사전 응답 텍스트

☑ 텍스트 입력

사후 응답 텍스트

☑ 텍스트 입력

자세한 정보

File Edit Selection View Go Run Terminal Help summarization.py - openai - Visual Studio Code

```
1 #Note: The openai-python library support for Azure OpenAI is in preview.
2 import os
3 import openai
4 openai.api_type = "azure"
5 openai.api_base = "https://hsopenai.openai.azure.com/"
6 openai.api_version = "2022-12-01"
7 openai.api_key = os.getenv("OPENAI_API_KEY")
8
9 # prompt_text = "Convert movie titles into emoji. \n\n Back to the Future: 🍌🍌🍌🍌 \n Batman: 🦇🦇🦇🦇 \n Transformers:
10 # prompt_text = "이 자식을 위해 이것을 요약하십시오. \n\n 육성은 태양에서 다섯 번째 행성이며 태양계에서 가장 큰 행성입니다. 그것은
11 # prompt_text = "다음 트윗의 감정을 분류하세요. \n\n1. "방학은 끝나가고 속제는 밀려왔을 뿐..." \n2. "이거 짜증나. 심심해" \n
12 prompt_text = "상점에서 캐릭터를 구매하였는데, 선택하지 않았던 다른 캐릭터가 구매되었습니다. 게임 업데이트 직후 새로운 캐릭터가
13
14 response = openai.Completion.create(
15     engine="hs-gpt-model",
16     prompt=prompt_text,
17     temperature=0.8,
18     max_tokens=1500,
19     top_p=1,
20     frequency_penalty=0,
21     presence_penalty=0,
22     best_of=1,
23     stop=None)
24
25 response_text = response.choices[0].text
26
27 print("Prompt: " + prompt_text + "\n\nResponse:" + response_text)
28
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL

PS Microsoft.PowerShell.Core\FileSystem:\wsl.localhost\Ubuntu\home\studydev\openai> python3 .\summarization.py
Prompt: 상점에서 캐릭터를 구매하였는데, 선택하지 않았던 다른 캐릭터가 구매되었습니다. 게임 업데이트 직후 새로운 캐릭터가 추가되면서 캐릭터가 밀리는 증상이 발생한 것 같습니다. 화면에서는 영희라는 캐릭터를 구매했는데, 구매한 캐릭터 내역을 보니 철수가 들어와 있습니다. 철수 캐릭터는 원하지 않았으며, 이를 취소하거나 원하는 캐릭터인 영희로 보상해 주기를 희망합니다. 그리고 이 문제가 재발하지 않도록 내부 프로세스를 개선해 주세요.

고객의 대화에서 다음을 추출하여 JSON 포맷으로 출력해.

- 이 대화의 주된 이유:
- 고객의 감정:
- 문제가 발생한 시점:
- 고객이 원하는 보상:
- 고객 응대 방법:
- 고객 응대를 위한 Email 콘텐츠를 HTML 포맷으로 작성:

Response:

```
{
  "이 대화의 주된 이유": "상점에서 캐릭터가 밀린 문제",
  "고객의 감정": "불만",
  "문제가 발생한 시점": "게임 업데이트 직후",
  "고객이 원하는 보상": "철수 캐릭터 취소 및 영희로 보상",
  "고객 응대 방법": "내부 프로세스 개선",
  "고객 응대를 위한 Email 콘텐츠를 HTML 포맷으로 작성": "<html><body><p>안녕하세요, 고객님!</p><p>상점에서 캐릭터가 밀린 문제에 대한 고객님의 불만에 대해 사과드립니다. 게임 업데이트 직후 새로운 캐릭터가 추가되었는데, 캐릭터가 밀린 증상이 발생한 것 같습니다. 그리고 고객님의 원하는 보상으로 철수 캐릭터를 취소하고, 영희로 변경해드리겠습니다. 또한, 이 문제가 반복되지 않도록 내부 프로세스를 개선해 드리겠습니다.</p></body></html>"
}
```

PS Microsoft.PowerShell.Core\FileSystem:\wsl.localhost\Ubuntu\home\studydev\openai>

예시 2

Cognitive Services | Azure OpenAI Studio KH

Azure OpenAI Studio > ChatGPT 플레이그라운드 [Privacy & cookies](#)

ChatGPT 플레이그라운드

[Show panels](#)

Assistant setup

Xbox customer support agent

[Save changes](#) [코드 보기](#)

System message ①

You are an Xbox customer support agent whose primary goal is to help users with issues they are experiencing with their Xbox devices. You are friendly and concise. You only provide factual answers to queries, and do not provide answers that are not related to Xbox.

Few-shot examples ①

User:

How much is a PS5?

Assistant:

I apologize, but I do not have information about the prices of other gaming devices such as the PS5. My primary focus is to assist with issues regarding Xbox devices. Is there a specific issue you are having with your Xbox device

[+ Add few-shot examples](#)

Chat session

채팅 지우기 Show raw syntax

PS5 얼마야?

저는 Xbox 고객 지원 대표로서 주로 Xbox 기기에 관련된 문제를 해결하는 데 초점을 맞추고 있습니다. 다른 게임 기기인 PS5의 가격 정보는 알 수 없으며, 해당 정보를 제공할 수 없습니다. Xbox 기기에 관련된 문제가 있으시면 도움을 드릴 수 있습니다.

Xbox 컨트롤러가 여러가지가 있던데 어떤 차이점이 있어?

네, Xbox 컨트롤러에는 여러 가지 종류가 있습니다. 다음은 일부 Xbox 컨트롤러의 차이점입니다.

1. Xbox Wireless Controller: Xbox One 및 Xbox Series X/S와 호환되는 표준 컨트롤러입니다. 무선 및 유선 모두 사용 가능합니다.
2. Xbox Elite Wireless Controller: Xbox One 및 Xbox Series X/S와 호환되는 고급 컨트롤러입니다. 사용자 정의 가능한 버튼 및 패들, 교체 가능한 촉각 패드 및 스틱, 그리고 무선 및 유선 모드를 지원합니다.
3. Xbox Adaptive Controller: 장애를 가진 사용자를 위한 컨트롤러로, 다양한 액세서리 및 장비를 연결할 수 있습니다.
4. Xbox Design Lab Controller: 사용자 지정 가능한 컨트롤러로, 사용자가 원하는 색상 및 디자인으로 만들 수 있습니다.

이러한 컨트롤러는 각각의 고유한 기능과 장점이 있으며, 사용자의 요구에 따라 선택할 수 있습니다.

여기에 사용자 쿼리를 입력하세요. (줄 바꿈 입력: Ctrl + Enter) [보내기](#)

매개 변수

배포

chatgpt35

Max response ① 350

온도 ① 0

Top P ① 0.95

Stop sequence ①

<lim_end> x

[자세한 정보](#)

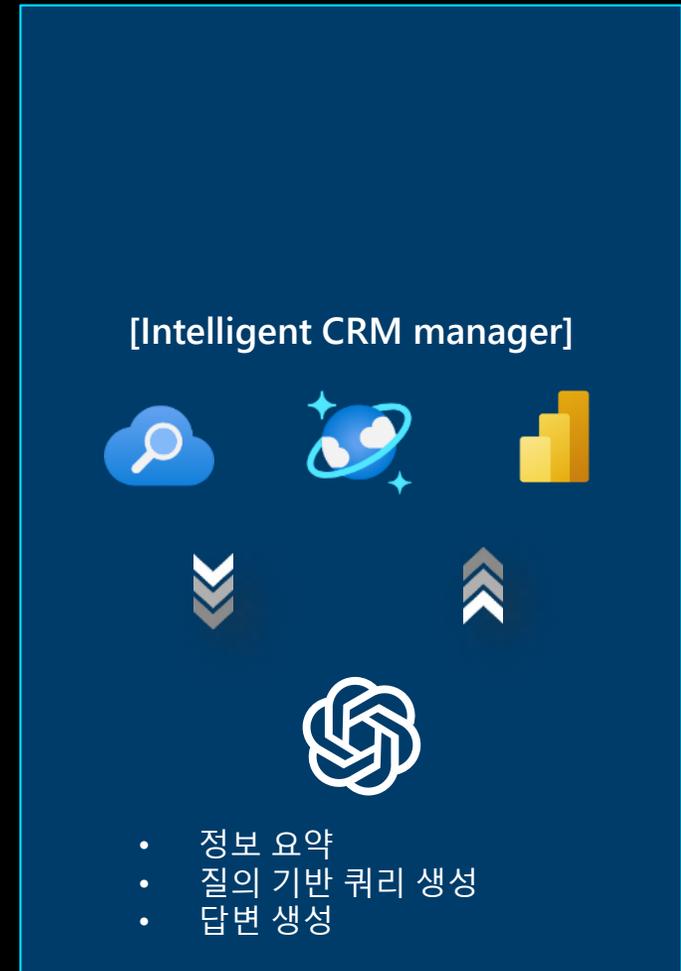
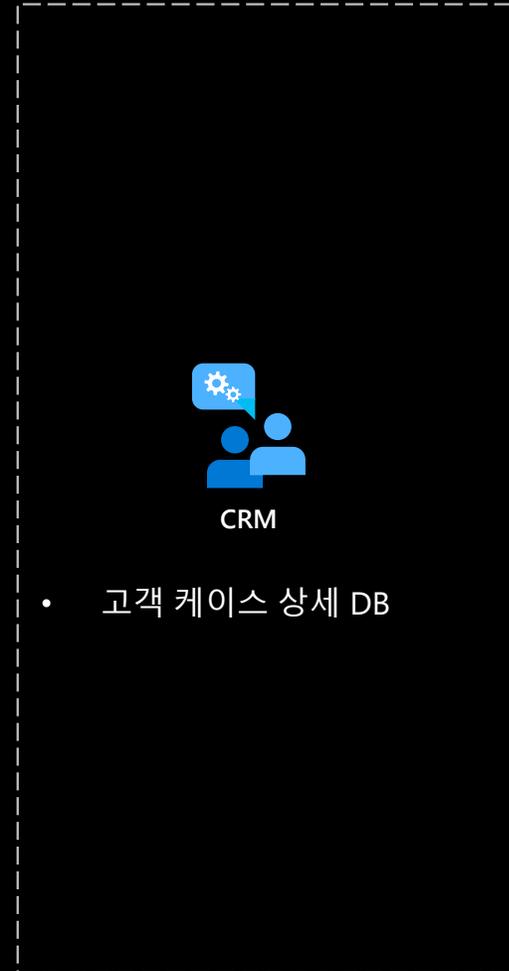
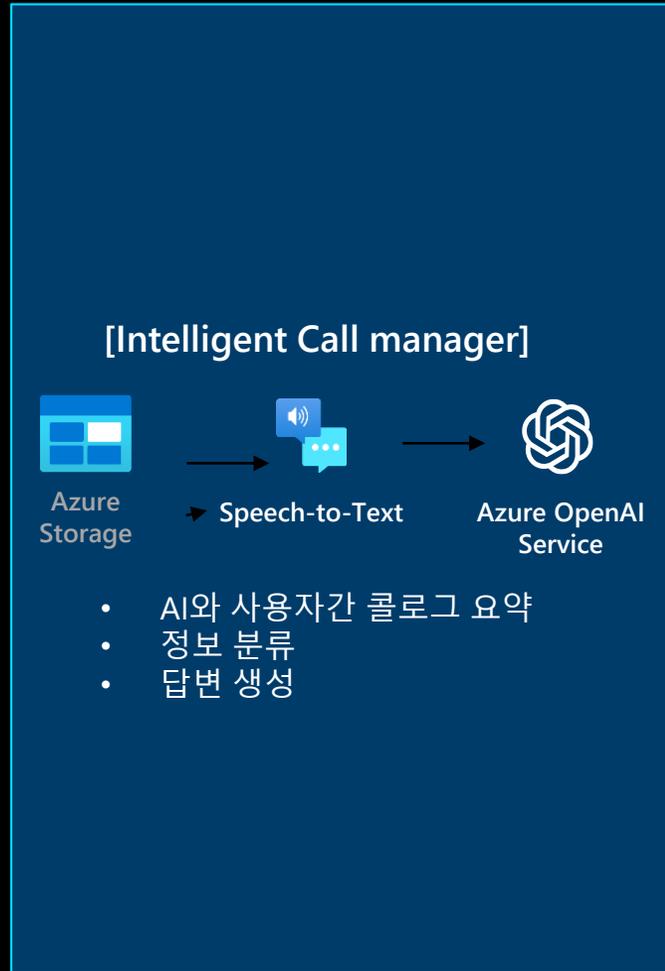
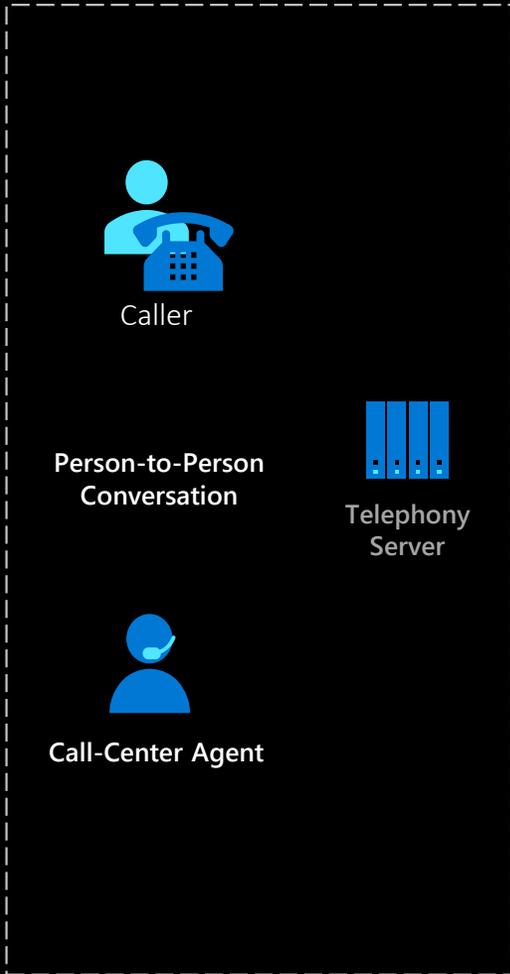
Session settings

Past messages included ① 10

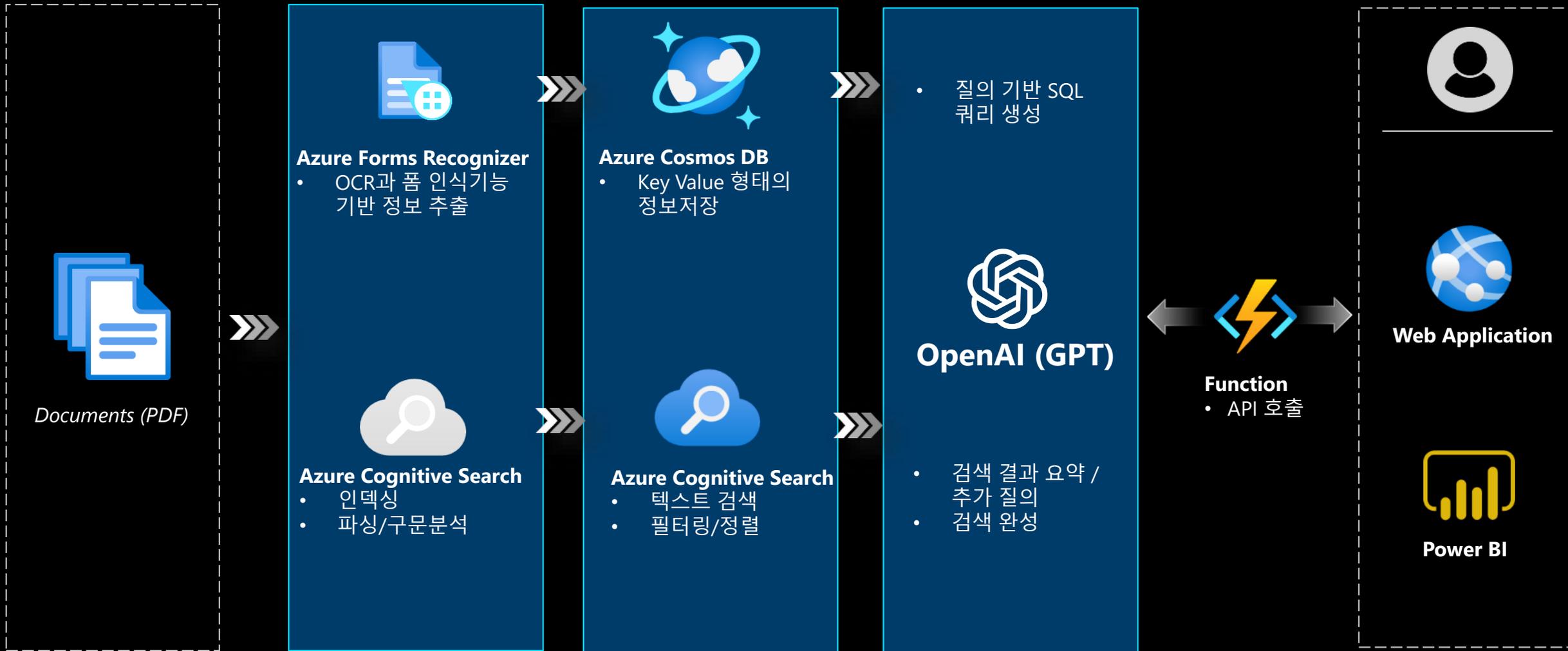
Current token count ①

1666/4096

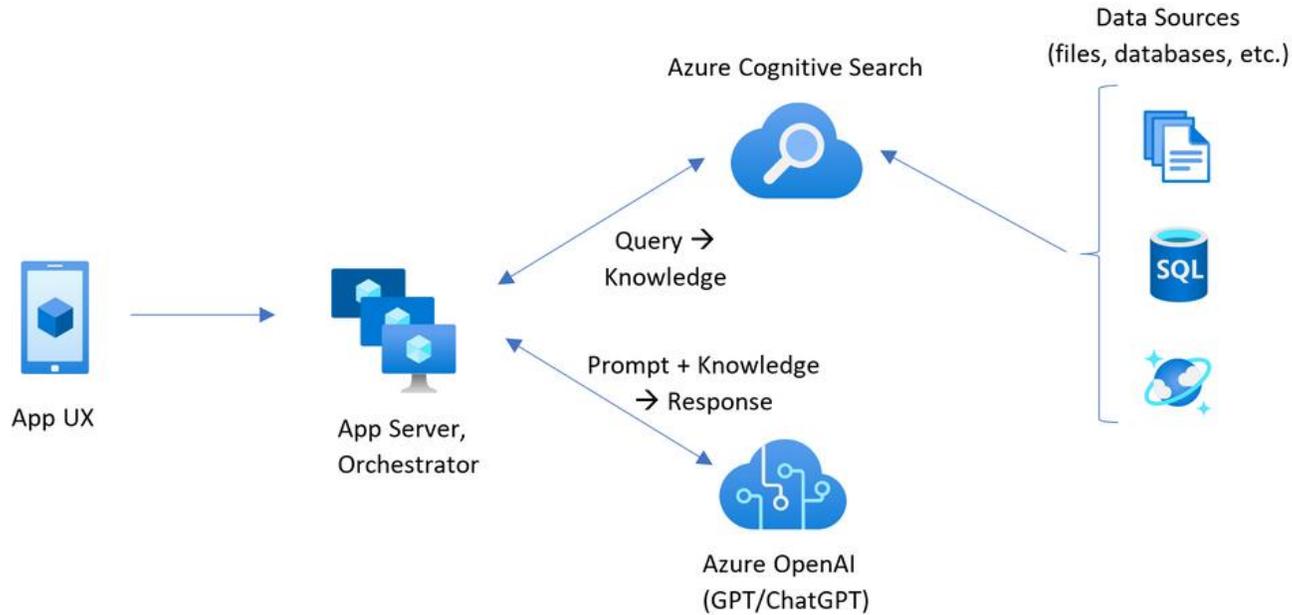
I 컨텍센터 콜 정보 요약 및 분류



I 문서에서 정보를 추출하고, 요약 및 맞춤형 쿼리 생성



ChatGPT를 활용한 QnA 서비스



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

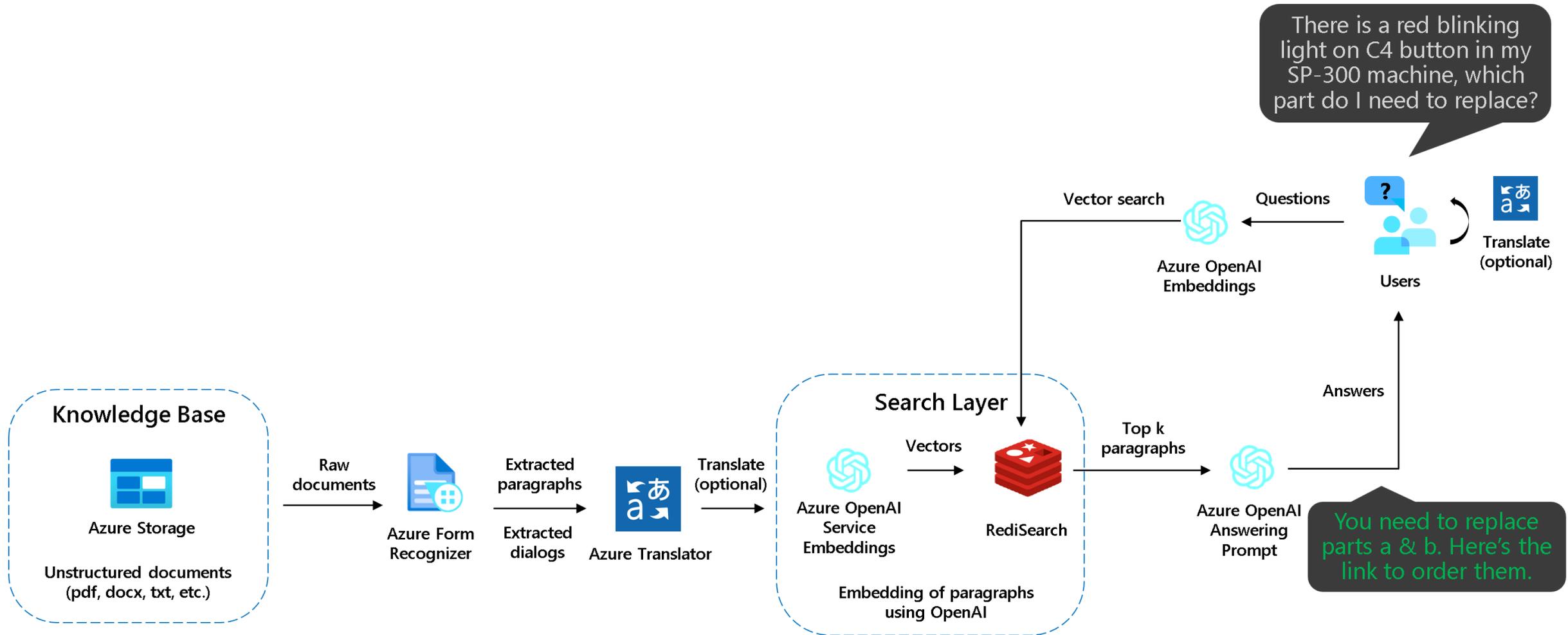
The screenshot shows the user interface of the QnA service. At the top, it displays "GPT + Enterprise data | Sample" and "Azure OpenAI + Cognitive Search". The chat area contains two messages:

1. User question: "Does my plan cover annual eye exams?"
AI response: "Both Northwind Health Plus and Standard plans offer coverage for vision exams ¹. 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

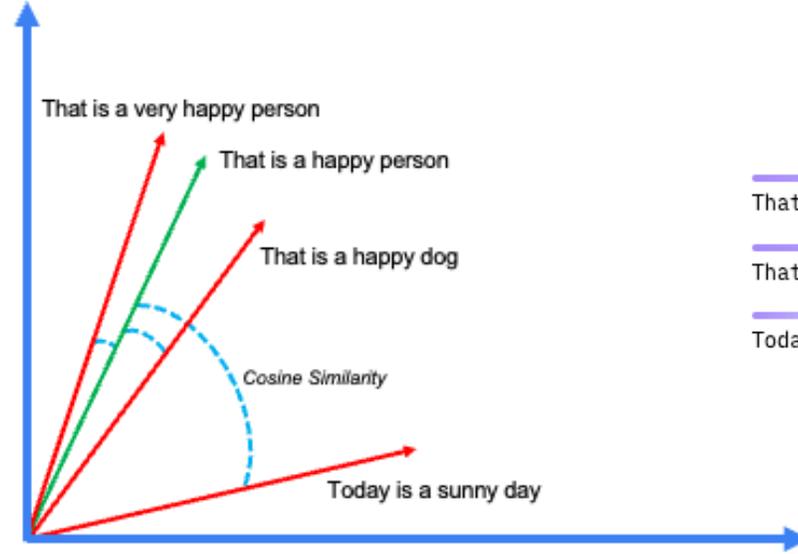
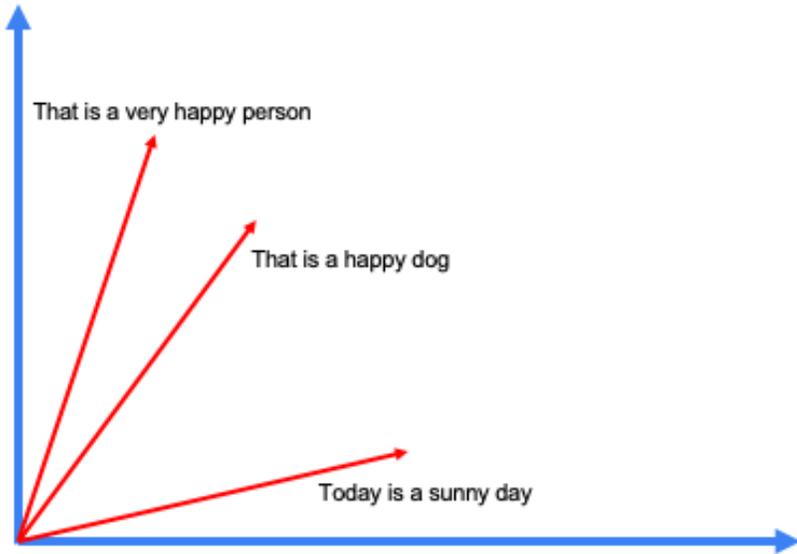
2. User question: "Hearing too?"
AI response: "Both Northwind Health Plus and Standard plans offer comprehensive hearing care coverage, including hearing tests and evaluations, hearing aids, and other associated services. You can receive hearing care services from any in-network provider and enjoy comprehensive 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'?"

At the bottom, there is a text input field with the placeholder "Type a new question (e.g. does my plan cover annual eye exams?)" and a send button.

임베딩을 활용한 지식 관리 서비스를 QnA 서비스로 활용



데이터 → 벡터!



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

“That is a very happy person”

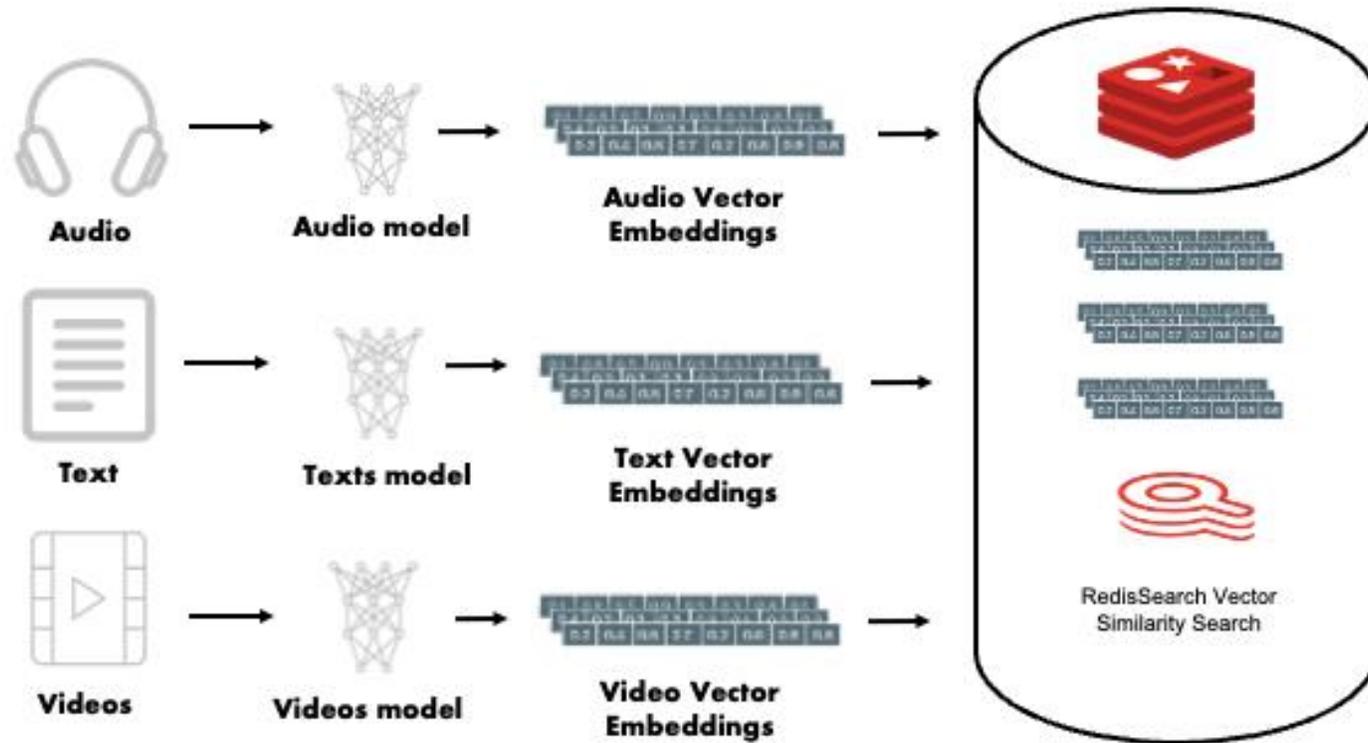
Data를 Vector(수치)화 하여 공간에 끼워넣는(임베딩) 것

데이터 → 임베딩 → 벡터



text-embedding-ada-002

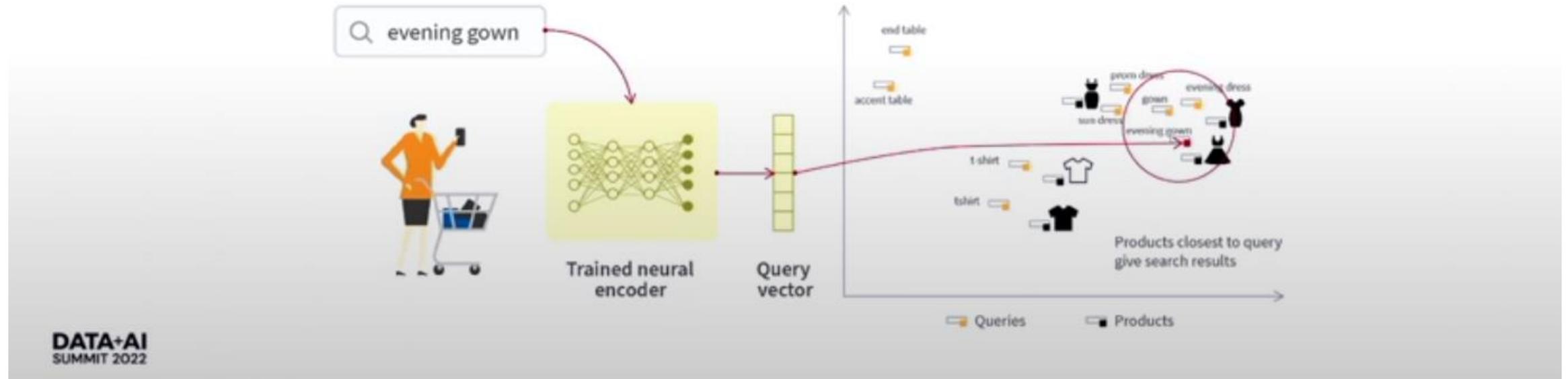
벡터 임베딩

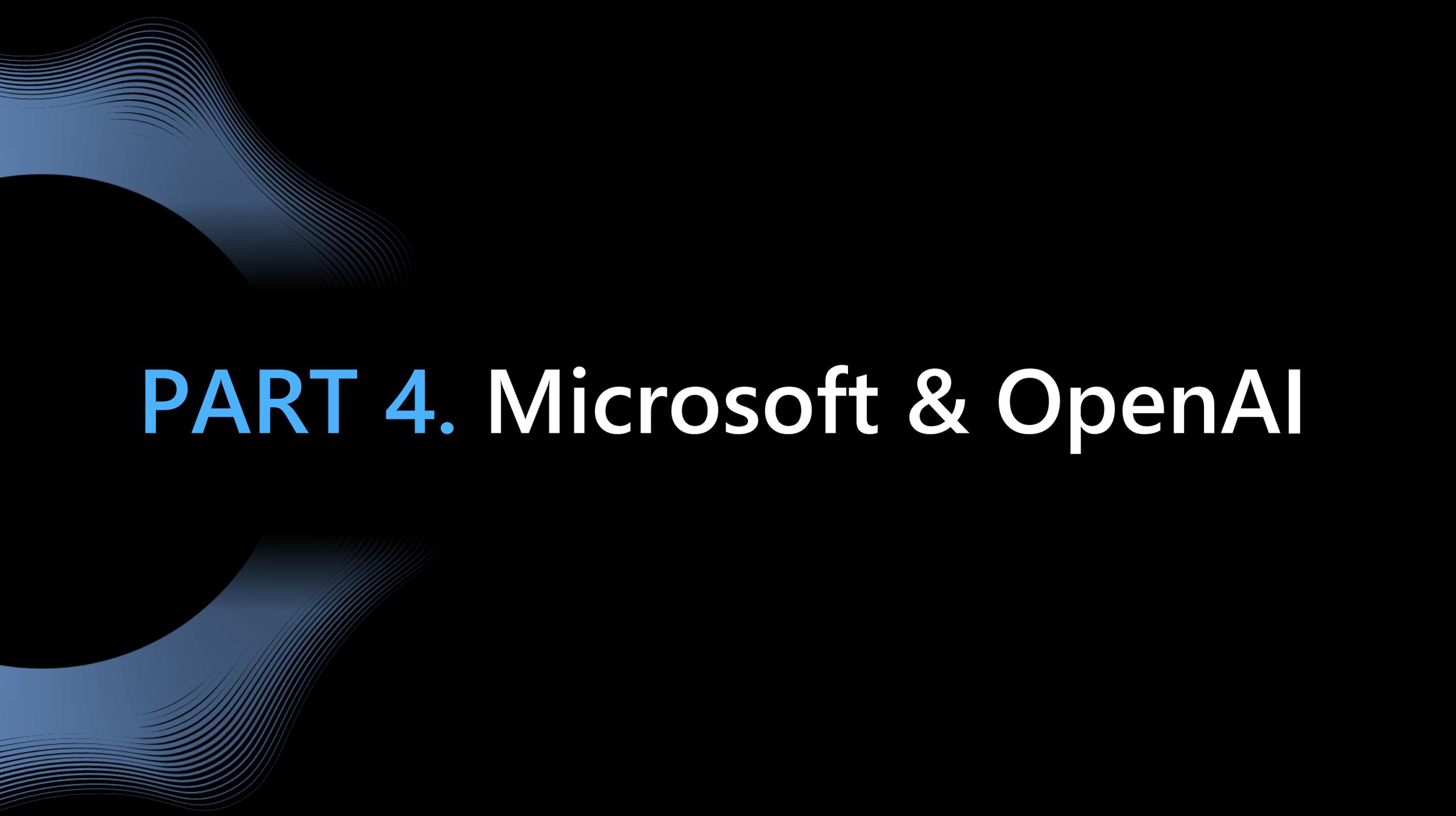


Vector Similarity Search

Vector Similarity Search Demo

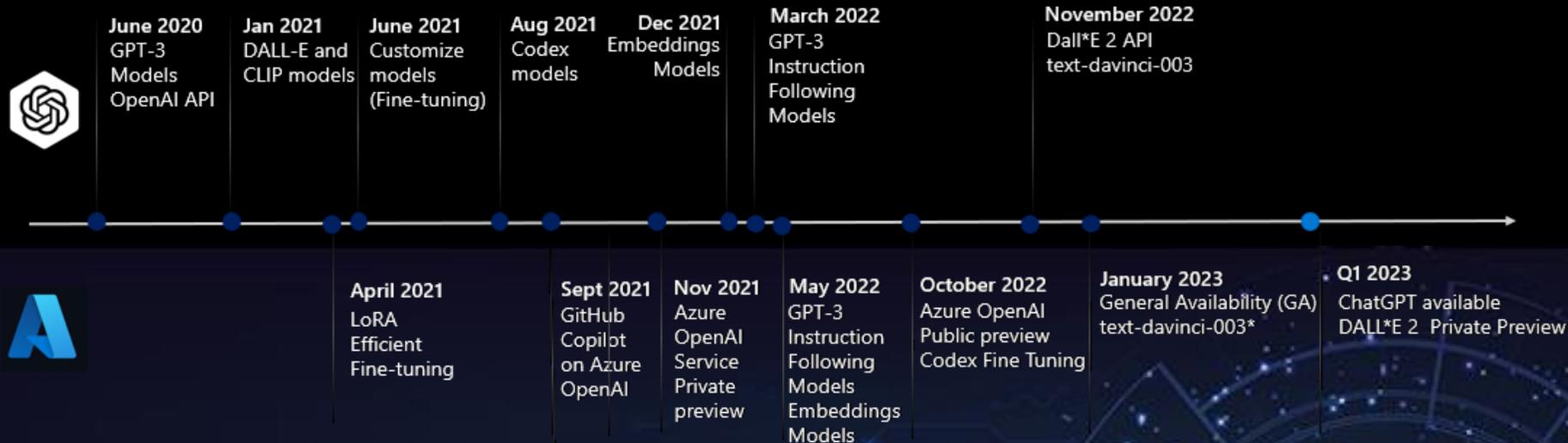
Find similar products with visual and semantic vector search





PART 4. Microsoft & OpenAI

Microsoft & Open AI 협력이 가지는 의미

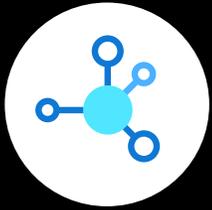


- ❖ **Supercomputing at scale** – Microsoft will increase our investments in the development and deployment of **[PART I] specialized supercomputing systems** to accelerate OpenAI’s groundbreaking independent AI research. We will also continue to build out Azure’s leading AI infrastructure to help customers **build and deploy their AI applications on a global scale.**
- ❖ **New AI-powered experiences** – Microsoft will deploy **[PART II] OpenAI’s models across our consumer and enterprise products and introduce new categories of digital experiences** built on OpenAI’s technology. This includes Microsoft’s **Azure OpenAI Service**, which empowers developers to build cutting-edge AI applications through **direct access to OpenAI models** backed by Azure’s trusted, enterprise-grade capabilities and AI-optimized infrastructure and tools.
- ❖ **Exclusive cloud provider** – **[PART III] As OpenAI’s exclusive cloud provider**, Azure will power all OpenAI workloads across research, products, and API services.
- ❖ **More Announcements** at <https://openai.com/blog>

AI

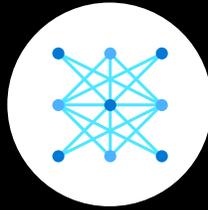
PART I. Specialized supercomputing systems

AI Cloud Infrastructure



NVidia GPUs
Intel FPGAs

High Bandwidth Network



NVLink
Infiniband

Systems Software



DeepSpeed library
ONNX Runtime

Platform



Azure
Machine Learning
Azure Cognitive
Services

Models



Turing family of
models
OpenAI Models

All available on Azure for everyone!

AI Supercomputer on Azure

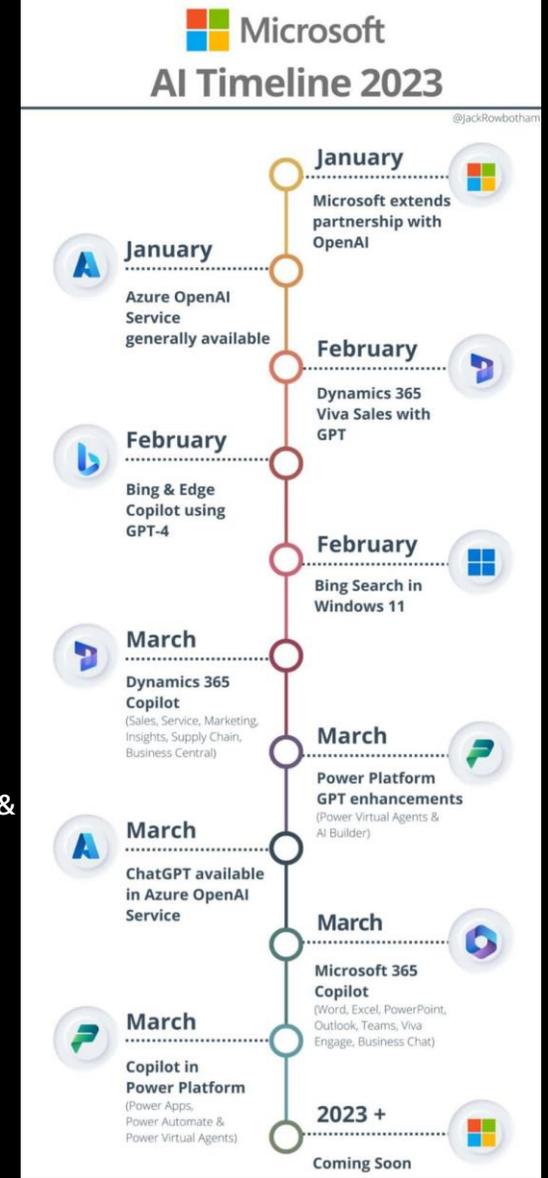
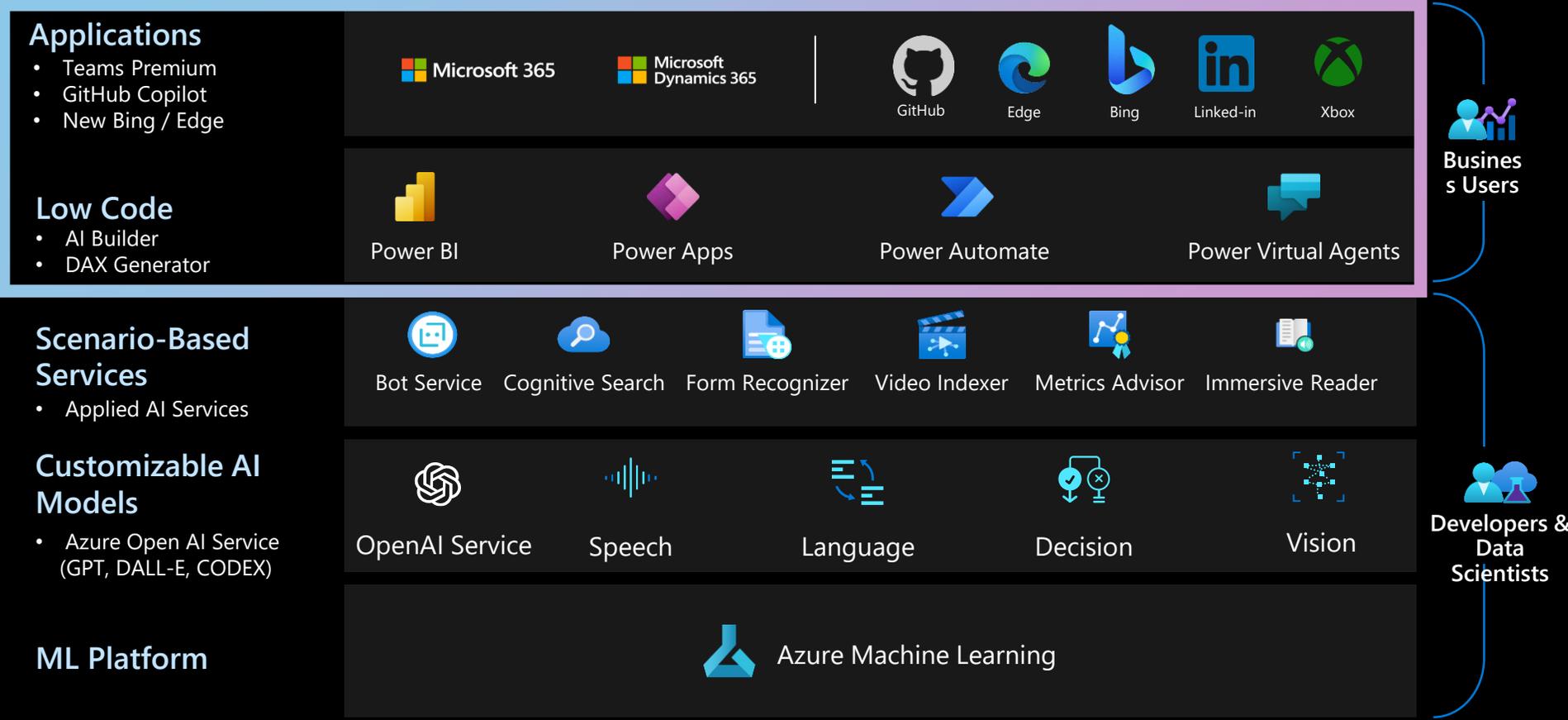
- The 5th largest supercomputer in the world
 - 300,000 CPU cores, 10,000 GPUs
 - **1st cloud based, designed for massive AI training**
 - Built in partnership with OpenAI
-

Highly distributed Frameworks

- Open-Source frameworks for massive model training **(20x bigger/10x faster)**
- DeepSpeed: Distributed framework enabling massive models training **(200B parameters)**
- Azure Machine Learning: Orchestrator via new ONNX distributed training. Azure Machine Learning recipes for training

PART II. Microsoft의 차별화된 디지털 경험 제공

1



Introducing



Microsoft 365 Copilot



Microsoft 365 Copilot

Natural Language



Large Language
Models

+



Microsoft Graph
- Your Data -

+



Microsoft 365
Apps



Microsoft 365 Copilot

Embedded across Microsoft 365 apps





The next
revolution in
computing

Consumer and enterprise products and introduce new categories of digital experiences

Microsoft Service Innovation

Microsoft AI

- Turing**
Rich language understanding
- Z-Code**
100 languages translation
- Florence**
Breakthrough visual recognition

OpenAI

- GPT**
Human-like language generation
- DALL-E**
Realistic image generation
- Codex**
Advanced code generation

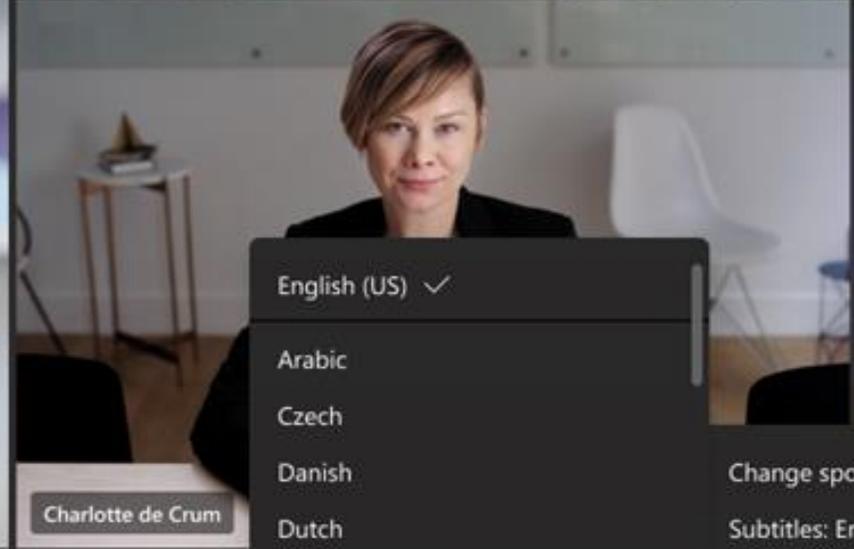
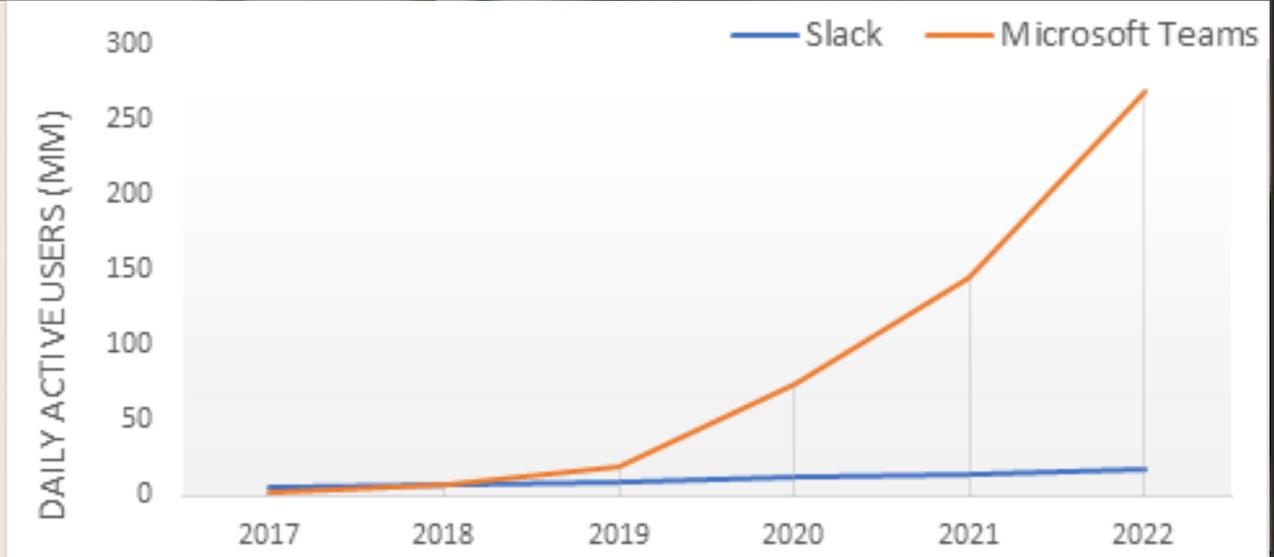
Azure Cognitive Services

- Vision Service
- Speech Service
- Language Service
- Decision Service
- OpenAI Service

Azure Applied AI Services

- Cognitive Search
- Form Recognizer
- Immersive Reader
- Bot Service
- Video Analyzer

-  Better search and Q&A
-  Better customer engagement and support
-  Better matching and content moderation
-  Better email management and meeting preparation
-  Better knowledge management
-  Better meeting management
-  Better reading and writing assistance
-  Better content moderation



Beth Davies



Kadji Bell



Babak Shammass

Babak S: So what did you all think of today's kickoff meeting for the new building project? I was gonna mention it in the group chat.

동시 통역 / 미팅 분석 / 요약 / Action Item 도출



Power BI DAX (Data Analysis eXpressions) | 자연어 기반 쿼리

Store Inspections - Power BI Desktop

File Home Insert Modeling View Help Format Data / Drill

Get data New visual More visuals New measure Sensitivity (preview) Publish

T-MOBILE DASHBOARD

Last Refresh: Wednesday, May 05, 2021

Clay Wesener
Principal GPM

Overview **Inspections** Stores Agents List View

Analysis
Area Inspected, Overtime and Store

Inspections by Response

Response	Count	Percentage
OK	7K	31.47%
Issue	15K	68.53%

Inspections by Date and Response

Inspections by Checklist Item and Response

Checklist Item	OK	Issue
Clutter free & accessible	10.6K	4.8K
Clean and hygienic	~1.5K	1.8K
Filled with available Products	2.7K	~0.2K
Devices are confirmed to be operational	~0.5K	~0.2K

Percentage Issues, Inspections and Inspection Score by Area Inspected

Area	Percentage Issues	Inspections
Store Entrance	~25%	~12K
Waiting Area	~30%	~6K
Phones	~35%	~2K
Information Kiosk	~38%	~2K
Front Tables	~40%	~2K
Tablets	~55%	~4K
Back Tables	~60%	~1K
Register	~85%	~1K

Overview **Inspections** Stores Agents List View

Easy measures

Easily create measures by writing an expression in natural language or selecting a built-in calculation.

Write your own Built-in calculations

Use natural language, like "Total sales for Canada"

Submit

PART III. As OpenAI's exclusive cloud provider

고객 및 파트너 서비스의 혁신을 위한 OpenAI on Azure 서비스 제공

Applications

- Teams Premium
- GitHub Copilot
- New Bing / Edge

 Microsoft 365

 Microsoft Dynamics 365



GitHub



Edge



Bing



LinkedIn



Xbox

Low Code

- AI Builder
- DAX Generator



Power BI



Power Apps



Power Automate



Power Virtual Agents

Scenario-Based Services

- Applied AI Services



Bot Service



Cognitive Search



Form Recognizer



Video Indexer



Metrics Advisor



Immersive Reader

2

Customizable AI Models

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



OpenAI Service



Speech



Language



Decision



Vision

ML Platform



Azure Machine Learning



Business Users



Developers & Data Scientists



인공 일반 지능(AGI)이
인류에게 도움을 주자



지구상의 모든 사람과 조직이 더
많은 것을 달성할 수 있도록

GPT-3

텍스트 생성 및 이해

Codex

코드 생성 및 이해

DALL·E

텍스트 프롬프트에서 이미지 생성
(preview)

ChatGPT

챗봇형 AI 모델
(preview)

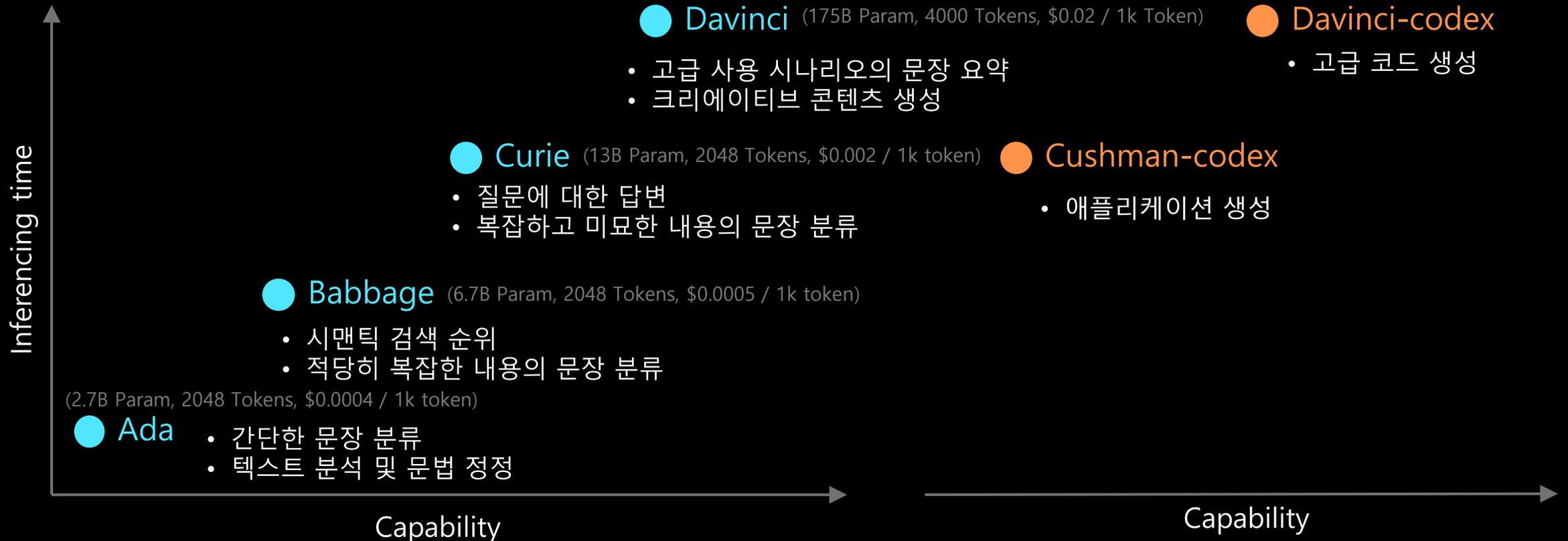
GPT-4

(preview)

Azure OpenAI | GPT-3 Family of Models

GPT-3 models

Codex models



Azure OpenAI 가격 시뮬레이션

모델 사용료 (Standard / Fine-tuned)

Instance	Series	Models	Inferencing per 1,000 tokens	Hosting per hour
Standard	Base Series	Ada	\$0.0004	
		Babbage	\$0.0005	
		Curie	\$0.002	
		Davinci	\$0.02	
	Base Series Fine-tuned	Ada	\$0.0004	\$0.05
		Babbage	\$0.0005	\$0.08
		Curie	\$0.0002	\$0.24
		Davinci	\$0.02	\$3
	Codex Series	Code-Cushman	\$0.024	
		Code-Davinci	\$0.10	
	Codex Series Fine-tuned	Code-Cushman	\$0.024	\$0.54/hour
	Embeddings	Ada	\$0.0004	
Babbage		\$0.005		
Curie		\$0.02		
Davinci		\$0.20		

Fine-tuning 학습 비용

Instance	Series	Models	Training per compute hour
Standard	Base Series Fine-tuned	Ada	\$20
		Babbage	\$22
		Curie	\$24
		Davinci	\$84
	Codex Series Fine-tuned	Code-Cushman	\$26

CASE #1 : Fine-Tune 미적용 비용 시뮬레이션
 텍스트 생성 목적으로 기본 모델 두가지를 사용 시
 (davinci-002 : 10,000 token / curie-001 : 20,000 token)

Item	Description
text-davinci-002 / Token Processed Cost	\$0.02 X 10K
text-curie-001 / Token Processed Cost	\$0.002 X 20K
Total	\$0.24

CASE #2 : Fine-Tune 적용 시 비용 시뮬레이션
 Davinci모델에 대해 Fine-Tune 을 1시간동안 수행 및 모델 배포
 5일 (120hr) 동안 14.5M 토큰생성

Item	Description
Fine-tune computing Cost	Davinci \$84 X 1hr
Model Deployed Cost	\$3 X 120 hr
Inferencing Cost (Token)	\$0.02 X 14.5K
Total	\$444.29

- ※ Azure에서 API 호출시, 'Inferencing per 1000 tokens' 항목 할인 적용
- ChatGPT와 Code-davinci는 최대 35%
 - Text-ada, text-babbage, text-curie, text davinci, code-cushman, dall-e는 최대 20%

Why Azure OpenAI?

Azure OpenAI Service

GPT-3

Codex

DALL·E *(preview)*

ChatGPT

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 AI 모델
Fine-tuning 모델 지원 (720시간, 20억개 토큰)



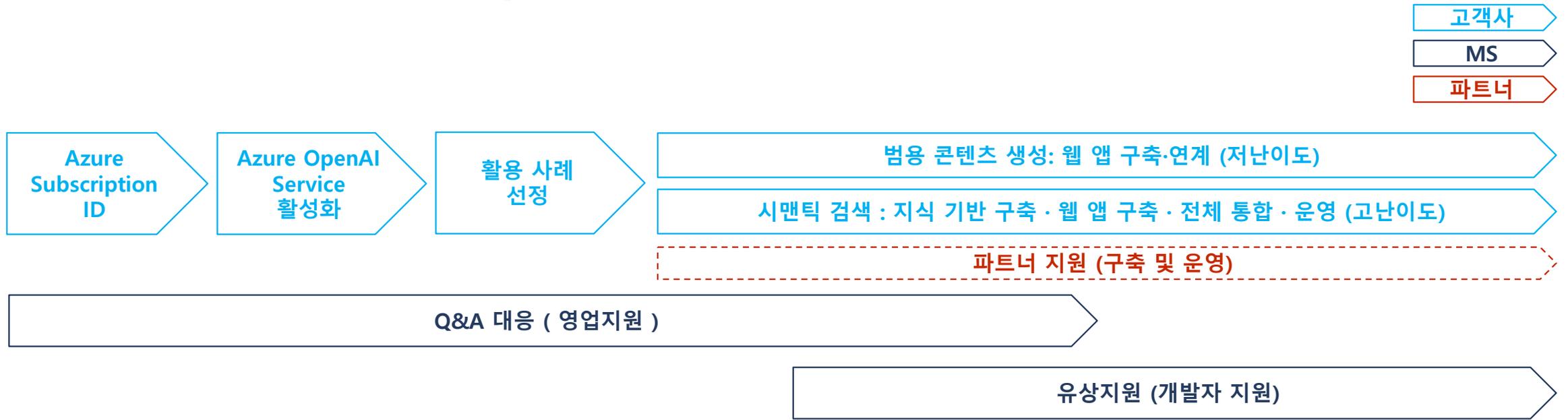
유해한 사용을 감지하고 완화하기 위한 콘텐츠
필터링 및 Responsible AI



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

비용 혜택 with Azure Services

Next Action - Azure OpenAI Service 사용 순서



OpenAI 서비스 도입을 위한 확인 사항

- Azure Open AI Service를 사용하려면 서비스 **사용 신청**이 필요합니다.
- 신청하려면 **Azure 구독 ID가 필요**합니다.
- 프론트 엔드 웹 앱을 제공함으로써 안전하고 격리된 환경에서 서비스를 제공할 수 있습니다.
- 또한 서비스 출시 전에는 책임있는 AI 기본 원칙의 (질적) 평가를 권장합니다. Azure OpenAI 서비스의 개인정보 보호, 보안 정책에 대한 자세한 내용은 참고 자료를 확인하세요.



Thank you