



쉽게 배우고 써먹는 Azure OpenAI 서비스

김도균 (MVP)



김도균

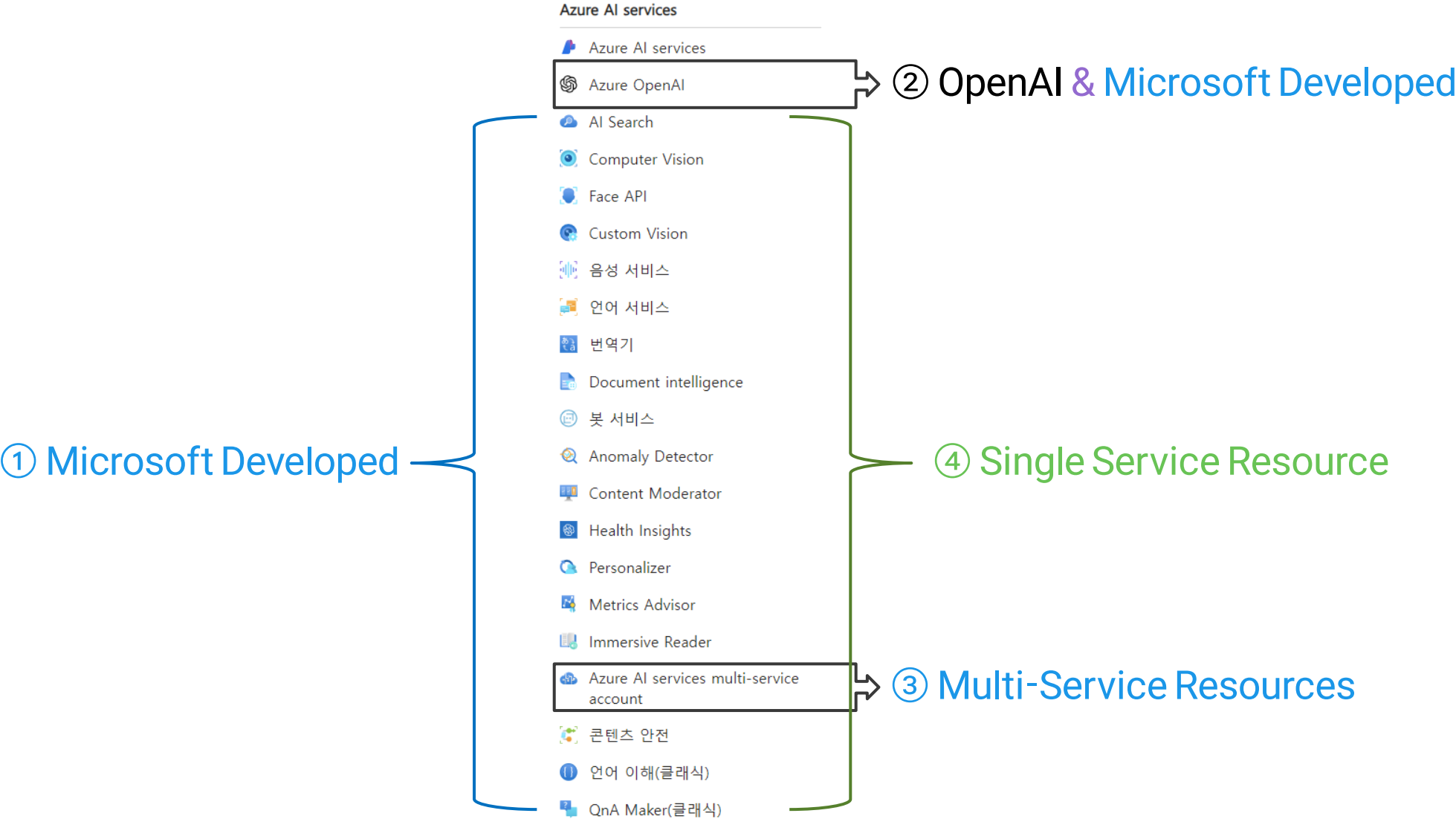
Microsoft MVP (Azure)

현. (주)늘품플러스 기술이사 · Microsoft Certified Azure Trainer

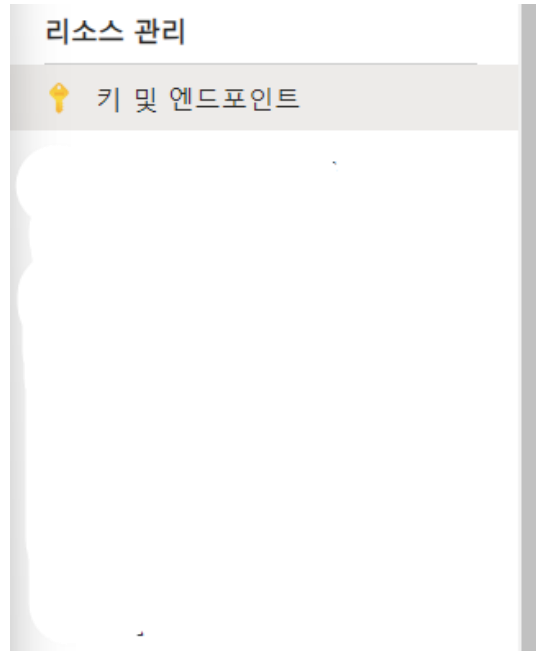
전. KT/KTDS, 도담시스템스, KAI, SK하이닉스

<처음 배우는 애저>, <처음 배우는 NHN 클라우드> 외 다수 번역 및 집필

Azure AI 서비스



서비스 액세스를 위한 정보 식별



키 표시

키 1
.....

키 2
.....

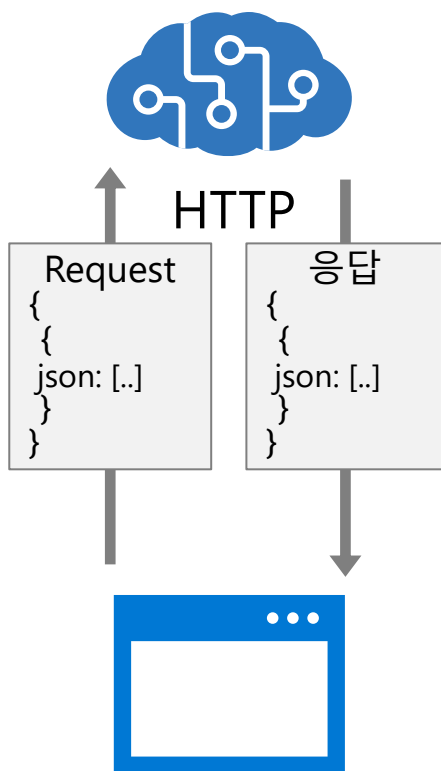
위치/지역 ⓘ
eastus

엔드포인트
https://ai-journal.openai.azure.com/

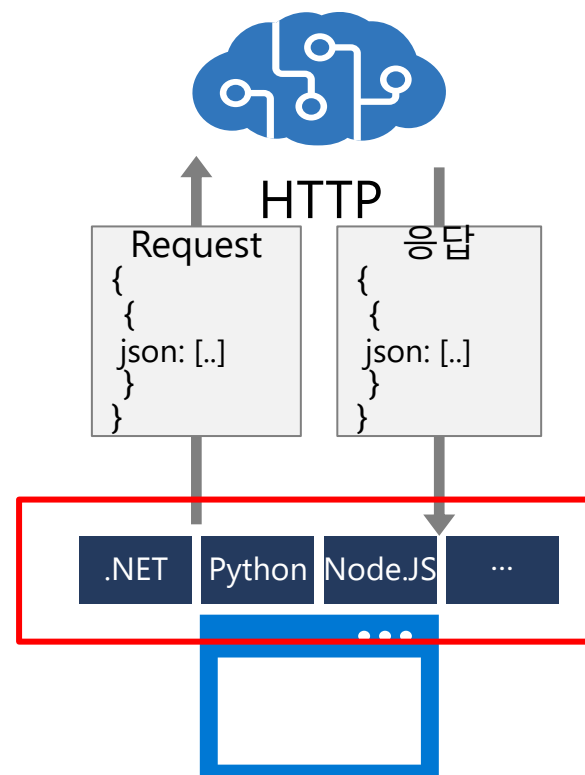
비밀 → Azure Key Vault

Azure AI Service 요청

REST API: 직접 HTTP 요청



SDK: 언어별 REST I/F 추상화



Gen. AI?



1950
년대

인공 지능

인간 지능을 복제본(replica) 초과할 수 있는
인텔리전트 머신을 만들려고 하는 컴퓨터 과학 분야

1990
년대

Machine Learning

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

2010
년대

딥 러닝

신경망 계층을 사용하여 데이터를 처리하고 의사
결정을 내리는 기계 학습 기술

2020
년대

생성형 AI

프롬프트 또는 기존 데이터가 제공되면 새로운 글,
시각적 개체 및 청각 콘텐츠 생성.

Azure OpenAI 리소스 배포

액세스 신청: <https://aka.ms/oaiapply>

홈 > Azure AI services | Azure OpenAI >

Create Azure OpenAI ...

① 기본 사항 ② 네트워크 ③ 태그 ④ 검토 + 제출

GPT-3 모델에서 제공하는 OpenAI의 언어 생성 기능을 사용하여 새로운 비즈니스 솔루션을 사용하도록 설정합니다. 이러한 모델은 수조 개의 단어로 미리 학습되었으며 유추 시 제공되는 몇 가지 간단한 예제를 사용하여 시나리오에 쉽게 적응할 수 있습니다. 요약에서 콘텐츠 및 코드 생성에 이르기까지 다양한 시나리오에 적용합니다.

[자세한 정보](#)

프로젝트 세부 정보

구독 * ⓘ

리소스 그룹 * ⓘ

[새로 만들기](#)

인스턴스 세부 정보

지역 ⓘ

이름 * ⓘ

가격 책정 계층 * ⓘ

지역에 따라 제공되는 모델이 다를 수 있음

Azure OpenAI Studio

- Azure OpenAI 모델 작업용 웹 포털
 - ✓ <https://oai.azure.com>
- 기본 모델 확인 및 배포
- 사용자 지정 데이터 원본 커넥터
- 사용자 지정 모델의 미세 조정 및 데이터 파일 관리
- 모델 테스트를 위한 플레이그라운드
 - ✓ 채팅 (GPT 3 및 이전)
 - ✓ 완료 (GPT3.5 – Turbo 이상)

플레이그라운드

채팅

완료

DALL-E

도우미(미리 보기)

관리

배포

모델

데이터 파일

활당량

콘텐츠 필터(미리 보기)

Azure OpenAI 서비스 시작

생성 AI 모델을 살펴보고, 사용 사례에 대한 고유한 프롬프트를 만들고, 선택한 모델을 미세 조정합니다.

시작

텍스트 생성



채팅 플레이그라운드

ChatGPT를 사용하여 사용자 지정된 AI 도우미를 설정합니다. GPT-3.5-Turbo 및 GPT-4 모델을 시험 사용해 보세요.

[지금 사용해 보기](#)

텍스트 생성



도우미 플레이그라운드

미리 보기

미리 발드된 대화 상태 관리 및 사용자 지정 도구를 사용하여 GPT 기반 AI 도우미의 개발 속도를 높이세요.

[지금 사용해 보기](#)

텍스트 생성



사용자 데이터 가져오기

고급 AI 모델에 사용자 데이터를 업로드하여 사용자의 이해와 작업 완료 의사 결정을 돕는 대화형 Copilot을 만드세요.

[지금 사용해 보기](#)

텍스트 생성



완성 플레이그라운드

요약, 콘텐츠 생성, 분류와 같은 사용 사례에 대한 완성 모델을 실험합니다.

[지금 사용해 보기](#)

이미지 생성



DALL-E 플레이그라운드

미리 보기

자연어로 설명을 작성하여 고유한 이미지를 생성합니다.

[지금 사용해 보기](#)

Gen AI. 모델 유형

모델 패밀리	설명
GPT-4	자연어 및 코드 생성을 위한 최신 채팅 기반 모델 지원 - 다중 모달 버전 GPT-4o(제한된 지역)
GPT-3	자연어 및 코드 생성 모델
Embedding	텍스트를 숫자 벡터 형식으로 변환해 텍스트 유사성 쉽게 파악 가능 모델
DALL-E	자연어 이미지 생성 모델(제한된 지역)

Azure AI | Azure OpenAI Studio

김도균
ai-journal (미국 동부, 50)

Azure OpenAI

플레이그라운드

채팅

완료

DALL-E

도우미(미리 보기)

관리

배포

모델

데이터 파일

할당량

콘텐츠 필터(미리 보기)

Azure OpenAI Studio > 모델

개인정보처리방침

모델

Azure OpenAI는 다양한 기능과 가격대를 가진 모델로 구성됩니다. 제공된 기본 모델 중 하나를 배포하여 Playground에서 시험해 보거나 더 나은 성능과 더 정확한 결과를 위해 특정 사용 사례 및 데이터에 맞게 사용자 지정 모델을 학습시키세요.
다양한 유형의 기본 모델에 대해 자세히 알아보기

기본 모델

배포 사용자 지정 모델 만들기 열 옵션 새로 고침

검색

모델 이름	모델 버전	만든 날짜	상태	배포 가능
dall-e-2	2.0	2024. 4. 15. 오전 9:00	성공	예
dall-e-3	3.0	2023. 8. 11. 오전 9:00	성공	예
gpt-35-turbo	0125	2024. 2. 15. 오전 9:00	성공	예
gpt-35-turbo	1106	2023. 11. 15. 오전 9:00	성공	예
gpt-35-turbo	0613	2023. 6. 19. 오전 9:00	성공	예
gpt-35-turbo	0301	2023. 3. 9. 오전 9:00	성공	예
gpt-35-turbo-16k	0613	2023. 6. 19. 오전 9:00	성공	예
gpt-35-turbo-instruct	0914	2023. 9. 14. 오전 9:00	성공	예
gpt-4	turbo-2024-04-09		성공	예
gpt-4	0613	2023. 6. 19. 오전 9:00	성공	예
gpt-4	1106-Preview	2023. 11. 15. 오전 9:00	성공	예
gpt-4	0125-Preview	2024. 1. 25. 오전 9:00	성공	예
gpt-4-32k	0613	2023. 6. 19. 오전 9:00	성공	예
gpt-4o	2024-05-13		성공	예
text-embedding-3-large	1	2024. 1. 25. 오전 9:00	성공	예

Gen.AI 모델 배포

Azure AI | Azure OpenAI Studio

Azure OpenAI

플레이그라운드

채팅

완료

DALL-E

도우미(미리 보기)

관리

배포

모델

데이터 파일

할당량

콘텐츠 필터(미리 보기)

Azure OpenAI Studio > 배포

배포

배포는 콘텐츠 조정 모델, 버전 처리, 배포...

+ 새 배포 만들기

배포 편집

배포 이름

gpt-journal

Dalle3

개인정보처리방침

Azure AI Studio 살펴보기

검색

모델 사용 ...

콘텐츠 필터

속도 제한(...)

업데이트 시간

25. 3. 21.	기본값	10000	2024. 6. 11. 오전 9:24
24. 10. 31.	기본값	-	2024. 6. 13. 오후 9:45

모델 배포

제공된 기본 모델 또는 사용자 지정 모델에 대해 API 호출을 만들도록 배포를 설정합니다. 완성된 배포를 사용할 수 있습니다. 배포가 완료되고 사용할 준비가 되면 배포 상태가 성공으로 전환됩니다.

모델 선택

gpt-35-turbo

모델 버전

기본값으로 자동 업데이트

배포 유형

표준

배포 이름

고급 옵션

콘텐츠 필터

Default

배포에 사용할 수 있는 분당 240K 토큰 할당량

분당 토큰 속도 제한(수천)

해당 RPM(분당 요청) = 720

동적 할당량 사용

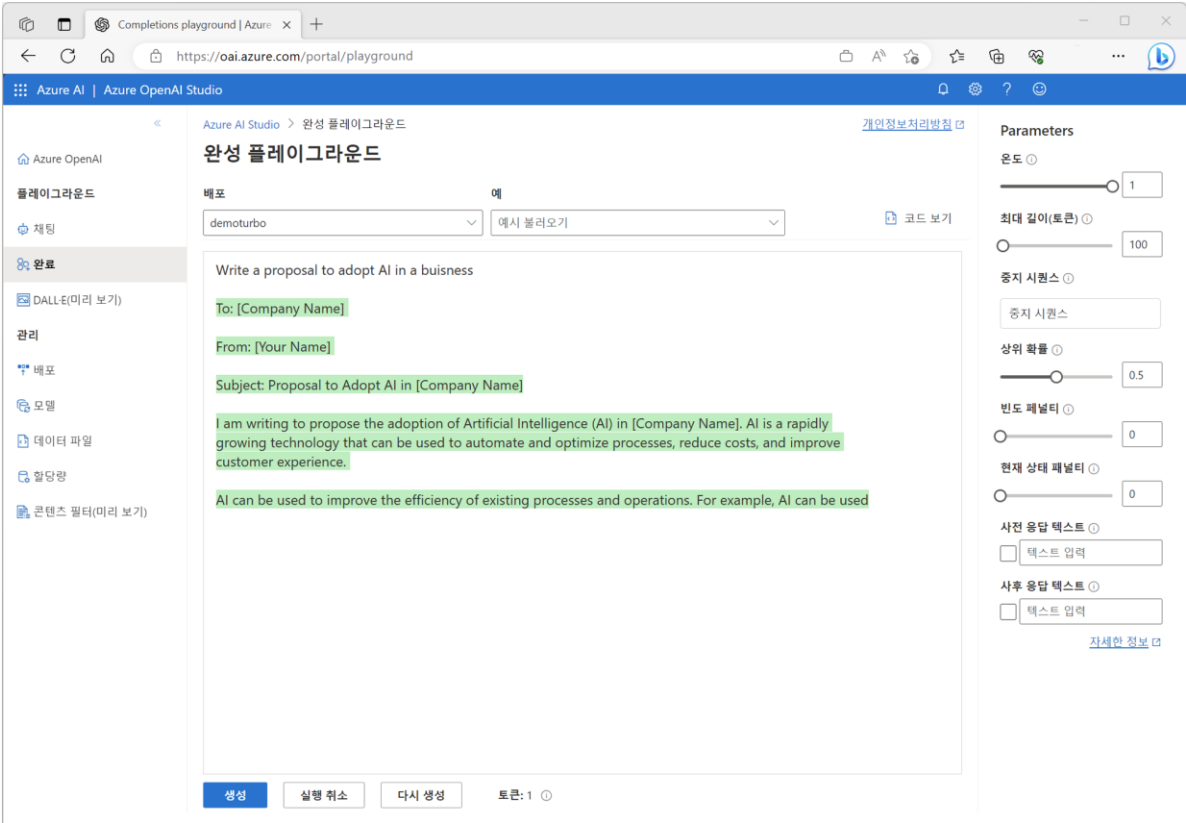
사용

만들기

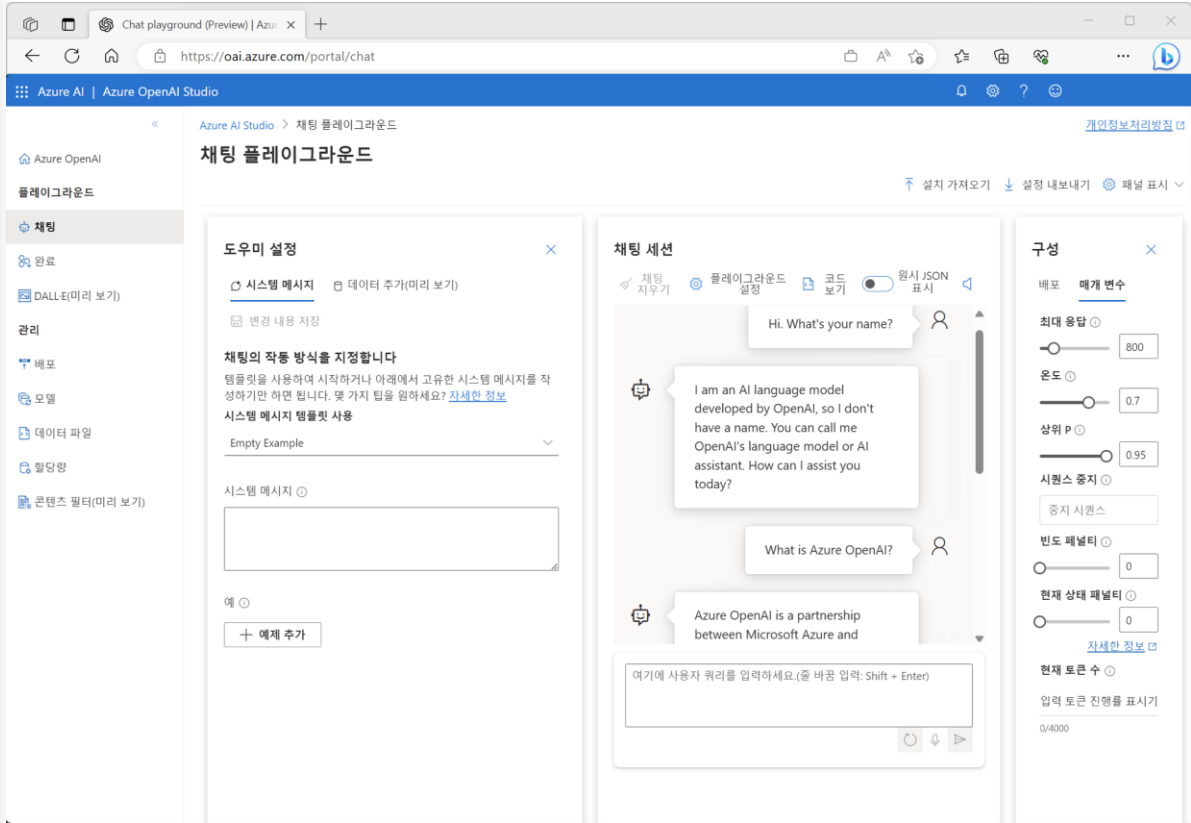
취소

플레이그라운드에서 모델 테스트

완성 플레이그라운드



채팅 플레이그라운드



도우미 (Preview)

- Copilot 기능을 갖춘 애플리케이션 개발 간소화
- 고급 도구로 강화된 사용자 지정 AI 도우미
 - ✓ 사용자 지정 함수
 - ✓ 코드 인터프리터
 - ✓ 파일 검색
- 특정 지침을 사용해 모델의 성격과 기능 조정
- 스레드를 사용한 상태 저장 지원
- 사용 사례
 - ✓ AI 기반 제품 추천
 - ✓ 영업 분석 앱
 - ✓ 코딩 도우미
 - ✓ 직원 Q&A 챗봇

도우미 플레이그라운드

Azure OpenAI

플래그라운드

채팅

완료

DALL-E

도우미(미리 보기)

관리

배포

모델

데이터 파일

할당량

콘텐츠 필터(미리 보기)

채팅 플래그라운드

배포 대상

설치 가져오기

설정 내보내기

패널 표시

설정

확인 데이터 추가

변경 내용 적용

안전 시스템 메시지

시스템 메시지 템플릿 사용

서식 파일 사용

템플릿을 사용하여 시작하거나 아래에서 고유한 시스템 메시지를 작성하기만 하면 됩니다. 몇 가지 힌트를 원하세요? [자세한 정보](#)

Default

시스템 메시지

You are an AI assistant that helps people find information.

예

+ 추가

채팅 지우기

플래그라운드 설정

코드 보기

JSON 표시

figure.ai에 대해 알려줘

죄송합니다, 현재 제가 알고 있는 정보로는 "figure.ai"에 대한 구체적인 내용을 찾지 못하였습니다. 이것이 특정 기업명, 웹사이트 또는 서비스를 가리키는 것일 수 있으나 추가적으로 확인해야 합니다. 좀 더 자세한 맥락이나 상세 정보를 주시면 도움을 드릴 수 있을 것 같습니다.

FIGURE 01의 몸무게는?

죄송합니다, "FIGURE 01"이라는 특정 객체나 개체에 대한 정보를 제공하는 것은 어렵습니다. 이것이 특정 상품명인지, 캐릭터 이름인지 또는 다른 종류의 항목을 가리키고 있는 지 명확하지 않기 때문입니다. 좀 더 구체적으로 설명해주시면 도움을 드릴 수 있을 것 같습니다.

구성

배포 매개 변수

배포

gpt4t-demo

세션 설정

지난 메시지 포함

10

현재 토큰 수

입력 토큰 진행률 표시기

287/32768

Azure OpenAI에서 사용자 데이터 적용 방법



데이터 세트 업로드

- *.jsonl 형식 텍스트 파일
- 학습 데이터와 유효성 검사 데이터 세트
- 데이터 원본
 - ✓ 로컬 파일 업로드
 - ✓ Azure Blob 또는 기타 공유 웹 위치



사용자 지정 모델

- 모델 사용자 지정 시 학습 데이터 직접 지정
 - ✓ 로컬 파일 업로드
 - ✓ Azure Blob 또는 기타 공유 웹 위치
- 미리 업로드한 데이터 세트 지정
 - ✓ 기존 파일에서 선택



배포된 모델에 사용자 고유 데이터 추가

- 채팅 플레이그라운드에서 데이터 추가
- 데이터 원본
 - ✓ Azure AI Search
 - ✓ Azure Blob Storage
 - ✓ Azure Cosmos DB for MongoDB vCore
 - ✓ Elasticsearch
 - ✓ URL/Web Address
 - ✓ Upload files

사용자 지정 모델을 위한 미세조정

사용자 데이터 파일 준비

Azure OpenAI

플레이그라운드

채팅

완료

DALL-E

도우미(미리 보기)

관리

배포

모델

데이터 파일

할당량

콘텐츠 필터(미리 보기)

Azure OpenAI Studio > 데이터 파일

개인정보처리방침

Azure AI Studio 살펴보기

데이터 세트 테스트 및 학습

사용 사례에 맞게 모델을 사용자 지정할 수 있도록 학습과 유효성 검사 데이터 세트를 업로드하세요. 미세 조정 작업은 여기에 표시되는 미세 조정 결과 파일도 출력합니다. 여기에서 '도우미' 관련 파일도 찾을 수 있습니다.
[미세 조정을 위해 데이터 세트를 준비하는 방법에 대해 자세히 알아보기](#)

새 데이터 세트 업로드

삭제

다운로드

열 옵션

새로 고침

검색

데이터 업로드

로컬 파일을 업로드하거나 Azure Blob 또는 다른 공유 웹 위치에서 가져옵니다.

로컬 파일 업로드

Azure Blob 또는 기타 공유 웹 위치

끌어서 놓기
파일 찾아보기

.jsonl (<200MB, UTF-8 BOM 텍스트 파일)

데이터 요구 사항에 대한 자세한 정보

파일 업로드

취소

데이터 업로드

로컬 파일을 업로드하거나 Azure Blob 또는 다른 공유 웹 위치에서 가져옵니다.

로컬 파일 업로드

Azure Blob 또는 기타 공유 웹 위치

파일 이름 *

파일 이름을 입력합니다.

파일 위치 *

Azure Blob 퍼블릭 액세스 URL, SAS 또는 기타 공유 웹 링크 입력

.jsonl (<512MB, UTF-8 BOM 텍스트 파일)

Azure Blob에 퍼블릭 액세스에 대한 자세한 정보

Azure Blob SAS(공유 액세스 서명)에 대한 자세한 정보

파일 업로드

취소

15

사용자 지정 모델을 위한 미세조정

더 나은 성능과 정확한 결과를 위해 특정 사용 사례와 데이터에 맞게 지정 모델 학습

배포한 모델 미세조정

채팅 플레이그라운드의 데이터 원본 추가

The screenshot displays the Azure OpenAI Studio Chat Playground interface. A red arrow points from the '+ 데이터 원본 추가' button in the '설정' (Settings) tab to the '데이터 추가' (Add Data) dialog box. The dialog box has a sidebar with three options: '데이터 원본' (selected), '데이터 관리', and '검토 및 완료'. The main content area is titled '데이터 원본 선택 또는 추가' (Select or Add Data Source). It contains a text block explaining that data sources are used to ground generated results and lists supported sources: Azure Blob Storage, databases, URLs, or local files. Below this is a section '데이터 원본 선택 *' (Select Data Source *) with a dropdown menu currently showing '데이터 원본 선택'. The dropdown list includes: 'Azure AI Search', 'Azure Blob Storage (preview)', 'Azure Cosmos DB for MongoDB vCore', 'Elasticsearch (preview)', 'URL/web address (preview)', and 'Upload files (preview)'. At the bottom of the dialog are '다음' (Next) and '취소' (Cancel) buttons. The background interface shows the '채팅 플레이그라운드' (Chat Playground) title, a sidebar with navigation links, and a '구성' (Configuration) panel on the right with settings for deployment, session, and tokens.

Azure OpenAI

플레이그라운드

채팅

완료

DALL-E

도우미(미리 보기)

관리

배포

모델

데이터 파일

할당량

콘텐츠 필터(미리 보기)

<<

Azure OpenAI Studio

모델

Azure OpenAI는
및 데이터에 맞게
다양한 유형의 기

사용자 지정 모

배포

모델 이

gpt-35

Azure OpenAI

플레이그라운드

채팅

완료

DALL-E

도우미(미리 보기)

관리

배포

모델

데이터 파일

할당량

콘텐츠 필터(미리 보기)

gpt-35-turbo-0125: ftjob-3ea7dc9b732d420f84b0e2b0e1533496

모델 배포 삭제 새로 고침

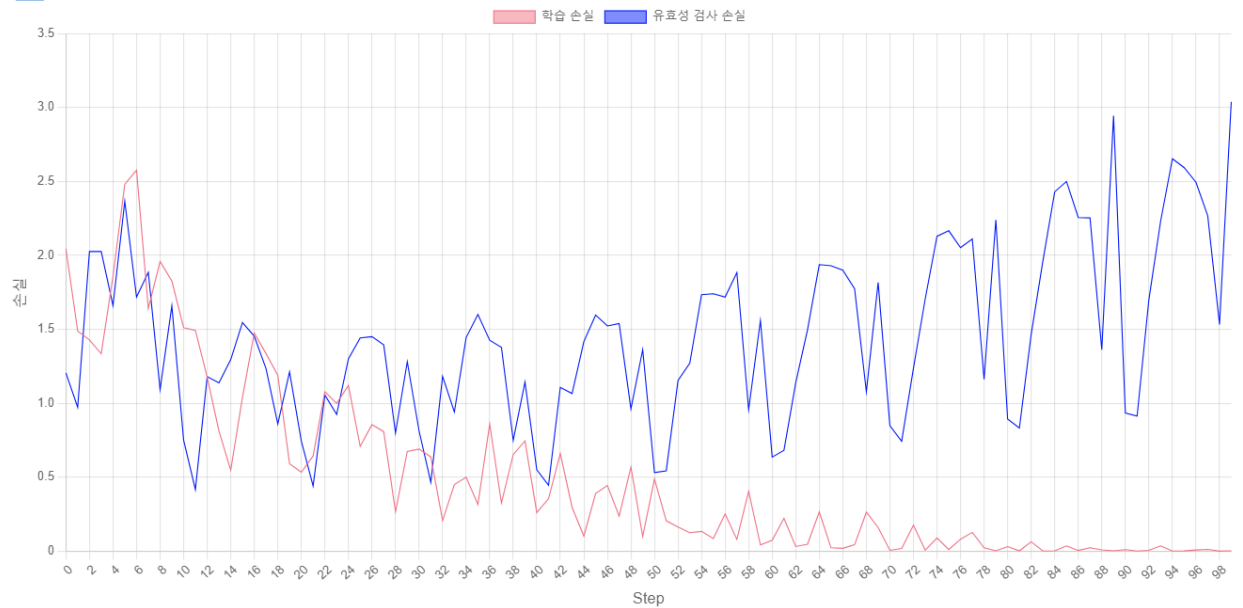
상태: ☺ 학습 성공

학습 완료 날짜: 2024. 6. 14. 오후 1:36
학습 파일: training_set.jsonl
기본 모델: gpt-35-turbo-0125
총 학습 시간: 51 분, 49 초
Training tokens billed: 5000

학습 파일 다운로드

학습 결과

손실 평균 토큰 정확도 토큰 정확도



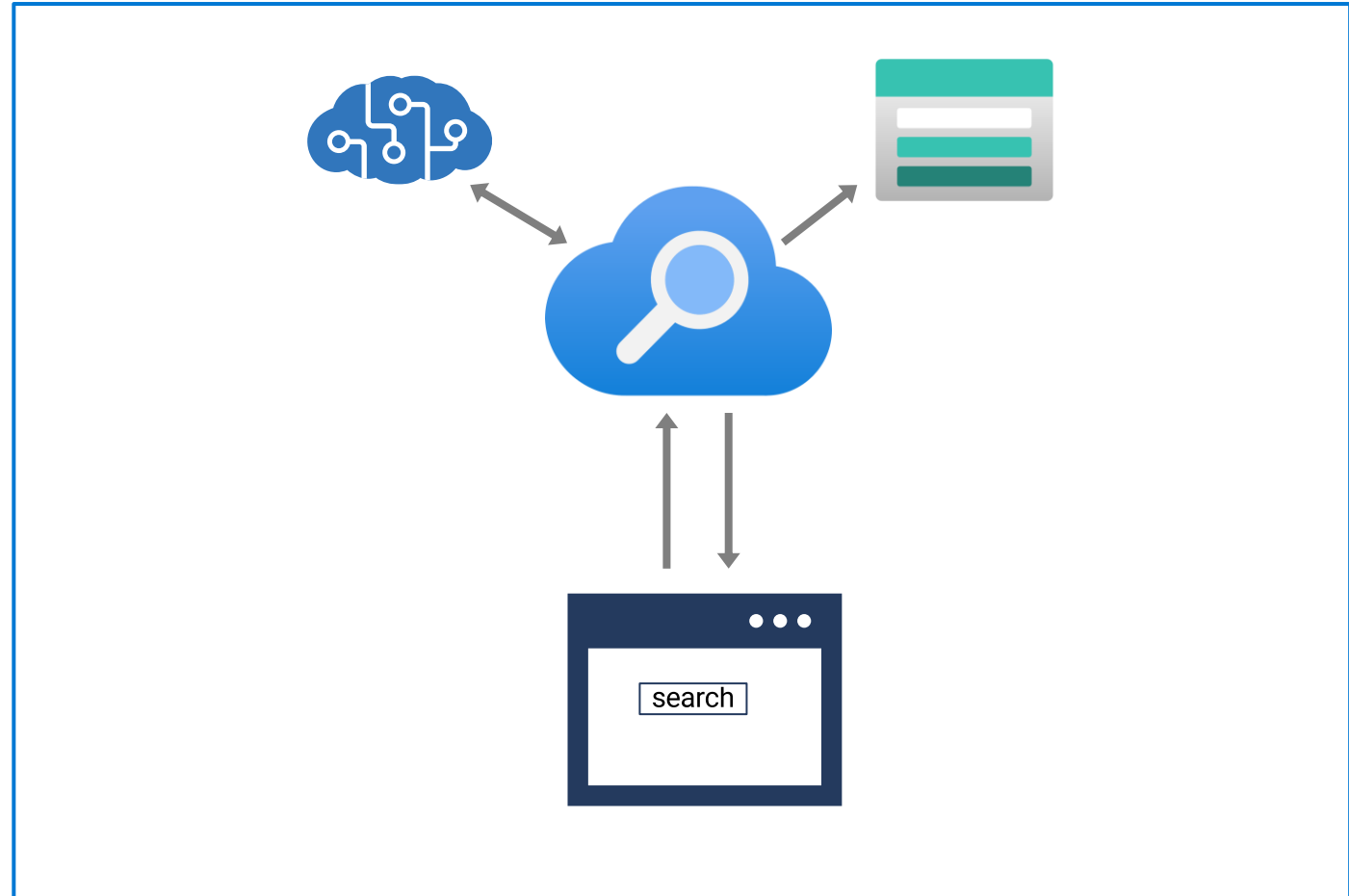
Azure AI Search

AI 기반 지식 마이닝

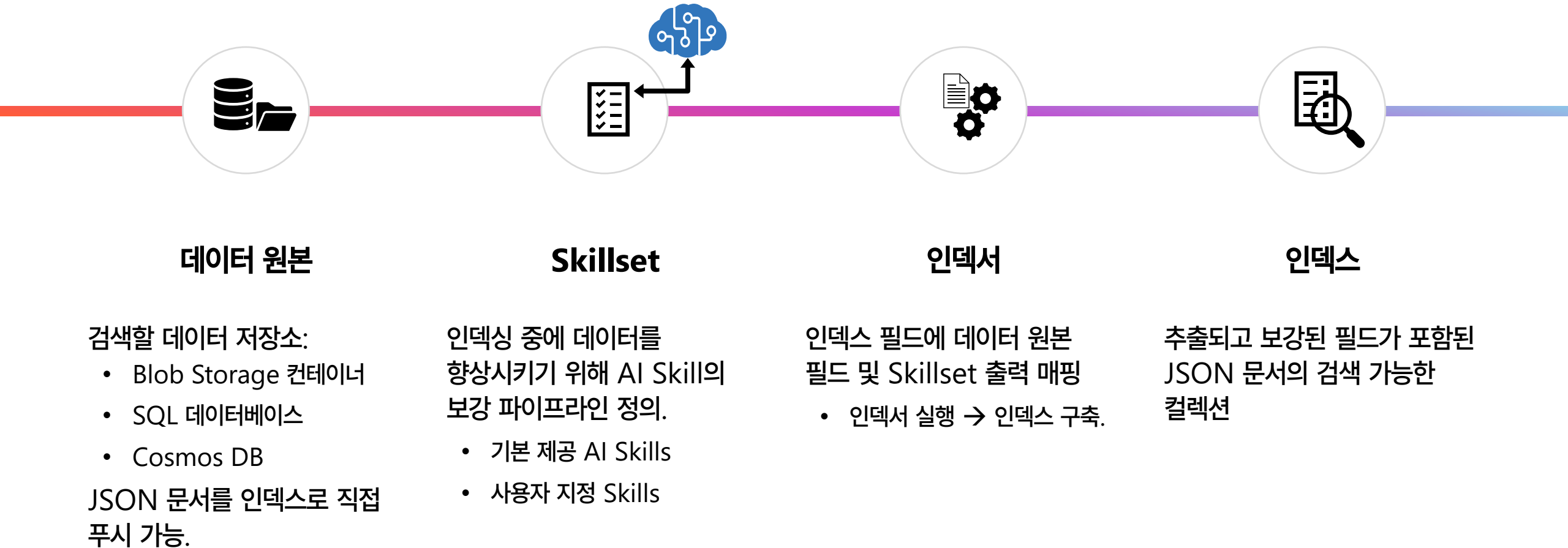
- 다양한 원본의 문서 및 데이터 인덱싱
- Skills를 사용한 인덱스 데이터 보강
- 분석 및 통합을 위해 추출된 인사이트를 지식 저장소에 저장

Azure 리소스:

- 핵심 인덱싱 및 쿼리: **Azure AI Search**
- 인덱스 보강: **Azure AI Services**
- 지식 저장소 지속성: **스토리지 계정**



AI 검색 솔루션의 핵심 구성 요소



데이터 추가

- ☒ 데이터 원본
- ☐ 데이터 관리
- ☐ 검토 및 완료

데이터 원본 선택 또는 추가

데이터 원본은 생성된 결과를 데이터에 근거하는 데
로 Azure Blob Storage, 데이터베이스, URL 또는 로
[Azure AI 데이터 개인 정보 보호 및 보안에 대해 자](#)

데이터 원본 선택 *

데이터 원본 선택

다음

설정

확인 데이터 추가

자신의 데이터 원본에 대한 인사이트를 얻으세요. 귀
하의 데이터는 Azure 구독에 안전하게 저장됩니다. 데
이터를 보호하는 방법에 대해 자세히 알아보세요.

데이터 원본: 리소스 검색:
Web page aisearch-demo-swc
인덱스: 청크 크기:
figureaiweb 1024

고급 설정 >

⚠ 데이터가 다음 경고와 함께 연결되었습니다.
• Truncated extracted text to '65536'
characters. (1 item(s) impacted)

🗑 데이터 원본 제거

채팅 지우기

플레이그라운드 설정

코드 보기

☐ JSON 표시

figure.ai에 대해 알려줘

Figure는 일반적인 목적의 인간형 로봇을 만드는 최초의 AI 로봇 기업입니다¹. 이 회사는 인간의 형태를 가진 로봇, Figure 01을 개발하였으며, 이 로봇은 제조, 물류, 창고, 소매 등 다양한 분야에서 지원을 제공할 수 있습니다¹. Figure 01은 인간의 형태와 최첨단 AI를 결합하여 단일 기능 로봇을 넘어서는 능력을 가지고 있습니다¹.

Figure 01의 사양은 다음과 같습니다:

- 키: 5'6"
- 페이로드: 20kg
- 무게: 60kg
- 작동 시간: 5시간
- 속도: 1.2M/S
- 시스템: 전기¹.

Figure는 또한 세계적인 로봇 공학 팀을 모집하였으며, 이 팀은 AI와 인간형 경험을 합쳐 100년 이상의 경험을 가지고 있습니다¹. 이 회사는 더 나은 삶의 질을 창출하고, 공급망을 수리하며, 신혼 경제를 지원하는 능력을 갖추게 하는 인간형 로봇을 개발하고 있습니다¹. 이를 통해 인력 부족을 해결하고, 위험한 직업에서 근로자의 수를 줄이는 등 인간이 더 많은 일을 할 수 있게 하는 것이 목표입니다¹.

1 참조

1 [www.figure.html](#) - Part 1

배포된 모델의 웹 앱 배포

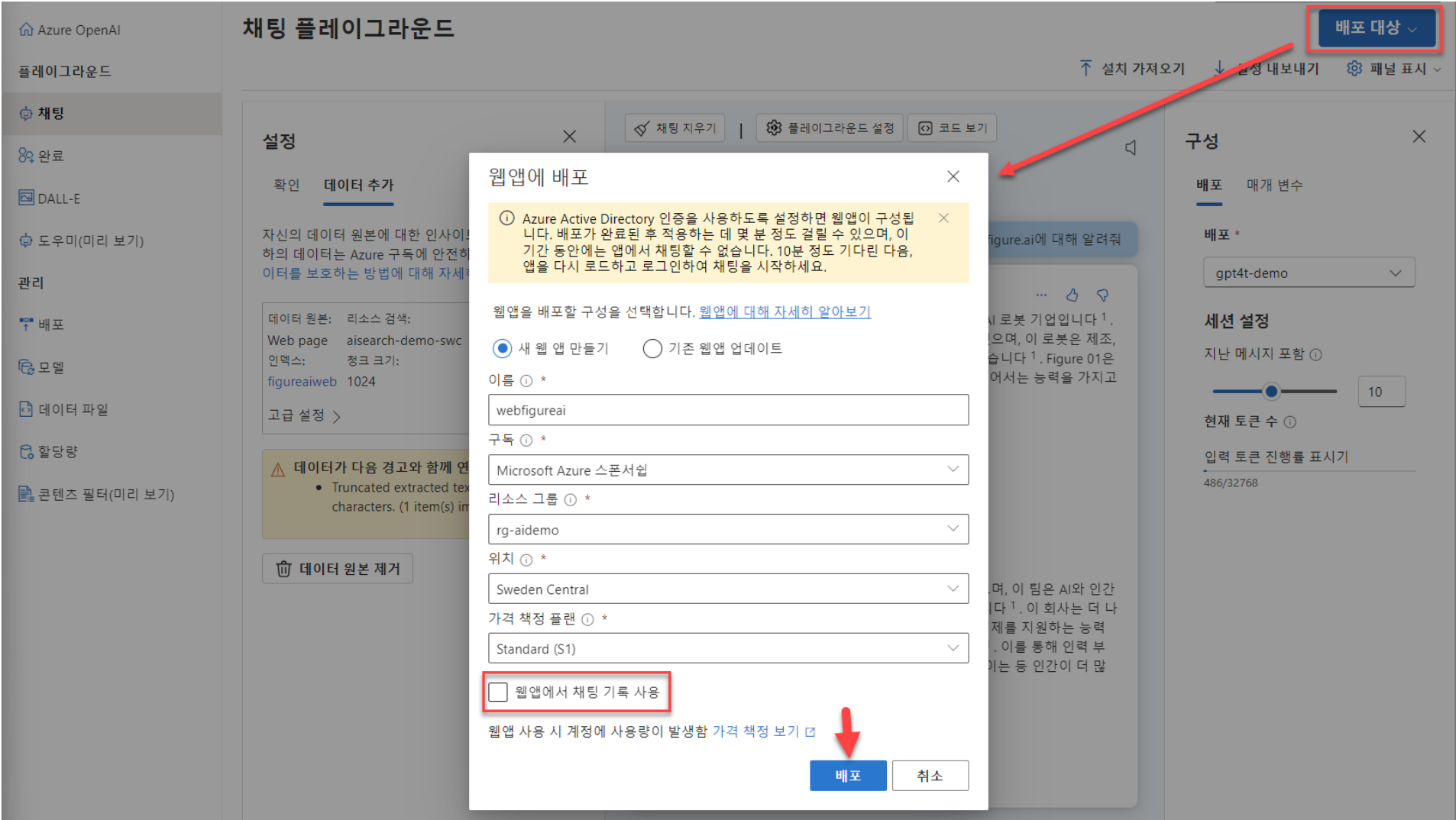


FIGURE 01의 몸무게는?

Figure 01, the world's first commercially-viable autonomous humanoid robot, weighs 60kg ¹.

1 reference ▾

AI-generated content may be incorrect

1 [www.figure.html - Part 1](#)

Figure.ai의 기업 평가 가치는 얼마인가?

Figure AI는 투자자들로부터 6억 7천 5백만 달러를 모집하며, 이를 통해 26억 달러의 기업 가치를 확보하였습니다 ¹.

1 reference ▾

AI-generated content may be incorrect

1 [robot-startup-figure...s-amazon-nvidia.html - Part 1](#)

Type a new question...



Citations



Skip Navigation Markets Pre-Markets U.S. Markets Europe Markets China Markets Asia Markets World Markets Currencies Cryptocurrency Futures & Commodities Bonds Funds & ETFs

6 billion as Bezos, OpenAI, Nvidia join funding

Published Thu, Feb 29 2024 8:00 AM EST Updated Thu, Feb 29 2024 11:38 AM EST

Annie Palmer@in/annierpalmer/

WATCH LIVE

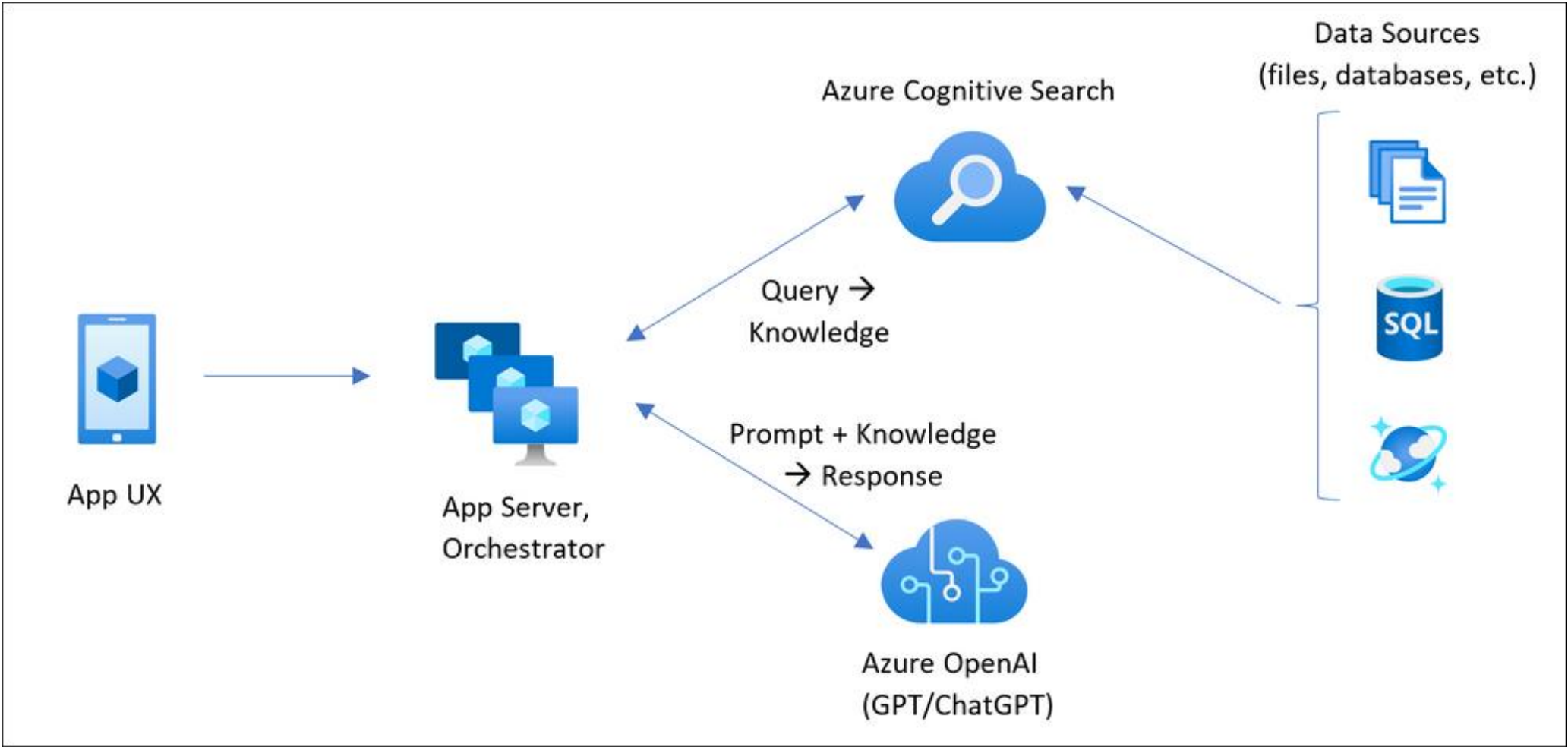
Key Points Figure AI raised \$675 million from investors including Jeff Bezos, Nvidia, Microsoft and OpenAI. The startup says it will use the money to accelerate development of its humanoid robot, which is intended for commercial use. Founded in 2022, Figure AI has developed a general-purpose robot, called Figure 01, that looks and moves like a human.

Startup Figure AI is developing general-purpose humanoid robots. Figure AI

Figure AI, a startup working to build humanoid robots that can perform dangerous and undesirable jobs, just got a big vote of confidence from some of the largest names in artificial intelligence. The company said Thursday that it raised \$675 million at a \$2.6 billion valuation from investors including Jeff Bezos, Nvidia, and Microsoft.

Founded in 2022, Figure AI has developed a general-

검색 증강 생성(RAG) 패턴 샘플 아키텍처



Margie's Travel

Enter a search term

London hotel

Search

Search:London hotel

Results:17

Filter by author:

☐ Reviewer

Sort by: Relevance

Refine Results

[201819.pdf](#)

Great *Hotel The City Hotel, London, UK 9/29/2018* The *hotel* is beautifull and the price is great!

A little far away from downtown *London*, but really close to a subway station.

- Author: Reviewer
- Size: 98503 bytes
- Modified: 2024-06-14T11:54:35Z
- Language: en
- Sentiment: positive
- Key Phrases:
 - The City Hotel
 - Great Hotel
 - subway station
 - Federico Pedroza
 - downtown London
- Locations:
 - The City
 - London, UK
 - London
 - UK

[201823.pdf](#)

Peaceful *The City Hotel, London, UK 10/5/2018* We stayed in this *hotel* in late September.

Although in busy part of *London* it was quiet at the weekend.

Margie's Travel



스스

Enter a search term

...

Skyscraper

Search

[SON 보기](#)

Search:Skyscraper

Results:4

Filter by author:

☐ Margies Travel

스도

Sort by: Relevance

Refine Results

[San Francisco Brochure.pdf](#)

Image caption: {"tags":["outdoor","sky","building","skyscraper","tall","tower","high","clear"],"captions":[{"text":"a tall building with many windows","confidence":0.44366756081581116}]}

- Author: Margies Travel
- Size: 385331 bytes
- Modified: 2024-06-14T11:54:31Z
- Language: en
- Sentiment: positive
- Key Phrases:
 - The Lombard Hotel Friendly, family-run hotel
 - The Golden Gate Bridge
 - San Francisco Bay Area
 - The Wharf Hotel
 - The Transamerica Pyramid
- Locations:
 - San Francisco San Francisco
 - Northern California
 - San Francisco Peninsula
 - United States
 - New York City
- Image Tags:
 - bridge
 - sky
 - suspension bridge
 - outdoor
 - cable-stayed bridge

감사합니다.