

目次

- 第1部: イントロ
- 第2部: 利用料金（の可能性）について
- 第3部: 利用申請
- 第4部: 利用料金発生時の流れ
- 第5部: 代表的サービスの使い方
 - 所有者の追加
 - OpenAIのサービスを利用しよう
 - 仮想マシンを作ろう
 - ノートブックを使おう

OpenAIのサービスを利用しよう

このスライドは 2025年2月21日 UTokyo Azure 利用説明会のために作成したものです。

Azureの仕様は頻繁に変わりますので、以下で動作しない場合は
Azure OpenAI Serviceのドキュメントを参照してください。

目的

- 学習済みモデルを利用する2つの選択肢
 - OpenAIのモデルをすぐに使いたい → Azure OpenAI Service
 - OpenAIを含め様々なモデルを使いたい → Azure AI Foundry
- Azure OpenAI Serviceのロードマップ
 1. リソースグループを作成
 2. モデルカタログからモデルを選択（たとえば”gpt-4o”を選ぶ）
 3. チャットプレイグラウンドからチャットを実行
 4. 外部のプログラムからWeb APIを通じてサービスを使う
 5. その他の話題
 - 5.1 料金
 - 5.2 1分当たりのトークン制限
 - 5.3 データ、プライバシー、セキュリティ
- その他の学習済みモデルの利用 (Azure AI Foundryの利用)
 - DeepSeek R1

1. リソースグループを作成

① Azure Portalを開く

<https://portal.azure.com/#home>

② “OpenAI”で検索

③ “Azure OpenAI”
をクリック

The search results for "openai" are displayed. The results are categorized into "すべて" (All) and "サービス (2)" (Services). The "Azure OpenAI" service is listed under "サービス". Below the services, there are sections for "リソースグループ" (Resource Groups), "Marketplace", "ドキュメント" (Documents), and "Microsoft Entra ID". A note at the bottom says "すべてのサブスクリプションを検索しています。" (Searching all subscriptions).

- すべて
- サービス (2)
 - Azure OpenAI
 - Azure AI services

リソースグループ

- [openai_test1]

Marketplace

- Azure OpenAI
- API Management
- Azure AI Foundry
- Azure AI services

ドキュメント

- Azure AI Foundry アーキテクチャ - Azure AI Foundry
- Azure Kubernetes Service (AKS) で OpenAI を使うアプリケーションをデプロイする - ...
- Azure OpenAI Service を使用して生成 AI ソリューションを開発する - Training
- Azure AI Foundry ポータルにおける Azure OpenAI - Azure AI Foundry

Microsoft Entra ID

- Azure Machine Learning OpenAI

Microsoft Entra ID で検索を続行してください

すべてのサブスクリプションを検索しています。



最終表示日

21 時間前

21 時間前

1 日前

1 日前

1 日前

2 日前

4 日前

2 週間前

2 週間前

4 週間前



サブスクリプション



リソース グループ



すべてのリソース



ダッシュボード

ホーム > Azure AI services

Azure AI services | Azure OpenAI

Azure AI services

+ 作成 削除されたリソースの管理 ビューの管理 更新 CSV にエクスポート クエリを開く タグの割り当て 削除

任意のフィールドのフィルター... サブスクリプション 次の値と等しい すべて 種類 次の値と等しい すべて フィルターの追加 増やす (2)

0 件中 0 ~ 0 件のレコードを表示しています。

名前 ↑↓	サブタイプ ↑↓	場所 ↑↓	カスタム ドメイン名 ↑↓	価格レベル ↑↓	状態 ↑↓
-------	----------	-------	---------------	----------	-------

表示する Azure OpenAI がありません
さまざまな自然言語タスクを実行します。

Azure OpenAI の作成 詳細情報 ↗ フィードバックの送信

検索

概要 All Azure AI services Azure AI services Azure AI services Azure OpenAI AI Search Computer Vision Face API Custom Vision 音声サービス 言語サービス 翻訳 Document intelligence ポット サービス Anomaly Detector Content Moderator Health Insights

ホーム > Azure AI services | Azure OpenAI >

Azure OpenAI の作成

⚠️ このステップを変更すると、後で選択する内容がリセットされる可能性があります。

[詳細情報](#)

プロジェクトの詳細

サブスクリプション * ⓘ

1923306880-Yama...

リソース グループ * ⓘ

openai_test1

[新規作成](#)

インスタンスの詳細

リージョン ⓘ

Japan East

名前 * ⓘ

gpt-4o-test-1

価格レベル * ⓘ

Standard S0

[価格の詳細を表示](#)

コンテンツ レビュー ポリシー

Azure OpenAI Service の有害な使用を検出して軽減するために、Microsoft は、お客様が完了 API およびイメージ生成 API に送信したコンテンツと、それが返送するコンテンツをログに記録します。サービスのフィルターによってコンテンツにフラグが設定されている場合は、Microsoft の常勤従業員によってレビューされる可能性があります。

モデルを稼働させるリソースだけでなく、パブリックIPやディスクなど、いろいろなリソースがこのモデルに紐づきます。
課金を完全に止めるなら、リソースグループごと削除するのが簡単です。
逆に言えば、不要になったら丸ごと消していく単位でリソースグループを分けておくことをお勧めします。

そのサービスのリソース
があるリージョンのみ
選べる
(たとえばJapan Westは
選べない)

< 前へ

次へ

フィードバックの送信

個人 Azure OpenAI の作成 - Microsoft Product Availability by Region

https://azure.microsoft.com/en-us/explore/global-infrastructure/products-by-region/table

Product Availability by Region

With 60+ announced regions, more than any other cloud provider, Azure makes it easy to choose the datacenter that's right for you.

© Microsoft 2024. All rights reserved. Release dates, features, and requirements are subject to change prior to final commercial availability. MICROSOFT MAKES NO WARRANTIES, EXPRESSED, IMPLIED, OR STATUTORY, AS TO THE INFORMATION ON THIS PAGE.

* Reserved Access Regions: Certain regions are access restricted to support specific customer scenarios, for example in-

openai 1/1

Azure OpenAI Serviceに限らず、どのリージョンにどのようなリソースがあるかはこのページで確認できる。

SKUとはStock Keeping Unitの略。

Search Products

Select a Geography

All

Preview

Products	Geographies	Japan		Korea	
		Japan East	Japan West	Korea Central	Korea South
Language Understanding (LUIS)		●	●	●	
Video Indexer		●	●	●	
Azure AI Metrics Advisor		●		●	
Azure OpenAI Service		●			●
Speaker recognition					
Azure AI Speech		●			
Language service		●			
Azure Video Analyzer				●	
Azure AI Video Indexer		●	●	●	
Azure API Center					
Azure Active Directory B2C					
Azure Advisor					

Japan WestはOpenAI Serviceのリソースがない

ホーム > Azure AI services | Azure OpenAI >

Azure OpenAI の作成 ...

✓ 基本 2 ネットワーク 3 Tags 4 レビューおよび送信

Configure network security for your Azure AI services resource.

種類 *

- インターネットを含むすべてのネットワークがこのリソースにアクセスできます。
- Selected networks, configure network security for your Azure AI services resource.
- 無効になっています。どのネットワークもこのリソースにアクセスできません。このリソースにアクセスする唯一の方法となるプライベート エンドポイント接続を構成できます。

< 前へ

次へ

フィードバックの送信

ホーム > Azure AI services | Azure OpenAI >

Azure OpenAI の作成

 基本 ネットワーク Tags レビューおよび送信

タグは名前と値のペアで、同じタグを複数のリソースやリソース グループに適用することでリソースを分類したり、統合した請求を表示したりできるようにします。 [タグに関する詳細情報](#)

タグを作成してから別のタブでリソースの設定を変更すると、タグは自動的に更新されることにご注意ください。

名前 <small>①</small>	値 <small>①</small>	リソース
<input type="text"/> :	<input type="text"/> Azure AI services	

< 前へ

次へ

フィードバックの送信

ホーム > Azure AI services | Azure OpenAI >

Azure OpenAI の作成

✓ 基本 ✓ ネットワーク ✓ Tags ④ レビューおよび送信☞ Automation のテンプレートをダウンロードする

使用条件

"作成" をクリックすることで、お客様は (a) 上記の Marketplace のオファリングに関連する法律条項とプライバシーに関する声明に同意し、(b) Microsoft より、そのオファリングに関連する料金が、現在の支払い方法に対して Azure サブスクリプションと同じ請求頻度で請求されることを認め、かつ、(c) Microsoft がお客様の連絡先情報、使用量情報、取引に関する情報を、サポート、請求、その他の取引上のアクティビティを目的として、オファリングのプロバイダーと共有する可能性があることに同意するものとします。Microsoft は、サードパーティのオファリングに対する権利は提供しません。その他の詳細については、[Azure Marketplace 使用条件](#)を参照してください。

基本

サブスクリプション	1923306880-YamakataLab
リソース グループ	openai_test1
リージョン	Japan East
名前	gpt-4o-test-1
価格レベル	Standard S0

ネットワーク

< 前へ

次へ

作成

フィードバックの送信

ホーム >



Microsoft.CognitiveServicesOpenAI-20250129153123 | 概要

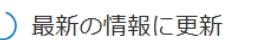
削除 キャンセル



再デプロイ



ダウンロード



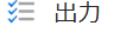
最新の情報に更新



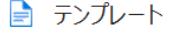
概要



入力



出力



テンプレート

... デプロイが進行中です

- デプロイ名 : Microsoft.CognitiveServ... 開始日時 : 2025/1/29 16:09:37
サブスクリプション : 1923306880-Yamakata... 関連付け ID : 48828a5b-d1e5-4953-9b8f-...
リソース グループ : openai_test1

▼ デプロイの詳細

リソース	種類	状態
		表示するリソースがありません。

フィードバックの送信

デプロイに関するエクスペリエンスについてお聞かせください



Microsoft Defender for Cloud

アプリおよびインフラストラクチャをセキュリティで保護します

[次に移動: Microsoft Defender for Cloud >](#)

無料の Microsoft チュートリアル

[今すぐ学習を開始する >](#)

エキスパートとの共同作業

Azure のエキスパートは、Azure での資産の管理を支援し、サポートの最前線となるサービス プロバイダー パートナーです。

[Azure Expert の検索 >](#)

ホーム >



Microsoft.CognitiveServicesOpenAI-20250129153123 | 概要

[削除](#) [キャンセル](#) [再デプロイ](#) [ダウンロード](#) [最新の情報に更新](#)

概要

[入力](#)[出力](#)[テンプレート](#)

✓ デプロイが完了しました

デプロイ名 : Microsoft.CognitiveServ... 開始日時 : 2025/1/29 16:09:37
サブスクリプション : 1923306880-Yamakata... 関連付け ID : 48828a5b-d1e5-4953-9b8f-...
リソース グループ : openai_test1

[デプロイの詳細](#)

▼ 次の手順

[リソースに移動](#)

フィードバックの送信

[デプロイに関するエクスペリエンスについてお聞かせください](#)

Cost Management

予算内に収め、予期しない料金が請求されないように通知を受け取ります。

[コストのアラートを設定 >](#)

Microsoft Defender for Cloud

アプリおよびインフラストラクチャをセキュリティで保護します

[次に移動: Microsoft Defender for Cloud >](#)

無料の Microsoft チュートリアル

[今すぐ学習を開始する >](#)

エキスパートとの共同作業

Azure のエキスパートは、Azure での資産の管理を支援し、サポートの最前線となるサービス プロバイダー パートナーです。

[Azure Expert の検索 >](#)

ホーム > Microsoft.CognitiveServicesOpenAI-20250129153123 | 概要 >

gpt-4o-test-1

Azure OpenAI

検索

Go to Azure AI Foundry portal 削除

JSON ビュー

概要

アクティビティ ログ

リソース グループ ([移動](#))

API Kind

OpenAI

アクセス制御 (IAM)

openai_test1

価格レベル

Standard

タグ

状態

エンドポイント

[エンドポイントを表示するには、ここをクリックします](#)

問題の診断と解決

アクティブ

キーの管理

[キーを管理するにはここをクリック](#)

> リソース管理

場所

Japan East

> Security

サブスクリプション ([移動](#))[1923306880-YamakataLab](#)

> 監視

サブスクリプション ID

a3b0857d-8c96-43a1-b2ee-d787b83bbe5e

> オートメーション

タグ ([編集](#))[タグの追加](#)

> ヘルプ

開始する Develop Explore and deploy

Build your own secure copilot and generative AI applications with Azure OpenAI Service

Deploy an Azure OpenAI model and start making API calls.
Connect your own data, call functions, and improve workflow with
Azure OpenAI language, image and speech models. You can
access the service through REST APIs, Python SDK, or our web-
based interface in the Azure OpenAI Studio.
[Learn More](#)

ホーム > Microsoft.CognitiveServicesOpenAI-20250129153123 | 概要 >

gpt-4o-test-1

Azure OpenAI

検索



Go to Azure AI Foundry portal



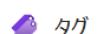
削除

1923306880-YamakataLab

[キーを管理するにはここをクリック](#)

アクティビティ ログ

アクセス制御 (IAM)



問題の診断と解決

> リソース管理

> Security

> 監視

> オートメーション

> ヘルプ

開始する Develop Explore and deploy

Build your own secure copilot and generative AI applications with Azure OpenAI Service

Deploy an Azure OpenAI model and start making API calls. Connect your own data, call functions, and improve workflow with Azure OpenAI language, image and speech models. You can access the service through REST APIs, Python SDK, or our web-based interface in the Azure OpenAI Studio.

[Learn More](#)

Monitor your Azure OpenAI usage

Monitor, assess and optimize your usage of the Azure OpenAI service API with an out-of-the-box metrics dashboard

[Azure OpenAI Metrics Dashboard](#)

チャット プレイグラウンド - Azure OpenAI

https://ai.azure.com/resource/playground?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/acc... A ⚡ 🌐 ⚙️ ⚪ ... 🎨

Azure AI Foundry | Azure OpenAI Service / gpt-4o-test-1 / チャットプレイグラウンド

すべてのリソース 5 🌐 ⚙️ 😊 gpt-4o-test-1 (japaneast, S0) 🔍

AI ソリューションを構築するためのより多くのモデル、機能、ツールをお試しになるか、追加の手順なしで既存のデプロイを導入してください。

チャットプレイグラウンド

コードの表示 デプロイ インポート エクスポート プロンプトのサンプル ...

セットアップ

デプロイ * + 新しいデプロイの作成

デプロイが存在しません

モデルに指示とコンテキストを与える ⓘ

情報を見つけるのに役立つ AI アシスタントです。

変更の適用 プロンプトを生成する

+ セクションの追加

① ソースデータが追加されている場合、一部のショットの例は使用されません。以前の例はすべてクリアされました。データを削除すると、例を追加できます。

> データを追加する
> パラメーター

チャットの履歴

メッセージアイコン

サンプル プロンプトを使って始める

Dialogue creation Create a conversation between two characters who are meeting for the first time in a mysterious place.

Character development Describe a day in the life of a superhero who has lost their powers.

Recipe creation Invent a recipe for a dish that combines flavors from two different cuisines.

ここにユーザー クエリを入力します。 (Shift + Enter で新しい行)

マイク アイコン

チャットプレイグラウンドのヘルプ

関連リソース チャットのクイックスタート

2. カタログからモデルを選ぶ

チャット プレイグラウンド - Azure OpenAI

https://ai.azure.com/resource/playground?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/acc... A ⚡ 🌐 ⚙️ ⚖️ ⚓ ⚔️

Azure AI Foundry | Azure OpenAI Service / gpt-4o-test-1 / チャットプレイグラウンド

すべてのリソース 5 🌐 ⚙️ 😊 gpt-4o-test-1 (japaneast, S0) 🔍

AI ソリューションを構築するためのより多くのモデル、機能、ツールをお試しになるか、追加の手順なしで既存のデプロイを導入してください。

チャットプレイグラウンド

コードの表示 デプロイ インポート エクスポート プロンプトのサンプル ...

セットアップ

デプロイ * + 新しいデプロイの作成

デプロイが存在しません

モデルに指示とコンテキストを与える ⓘ

情報を見つけるのに役立つ AI アシスタントです。

変更の適用 プロンプトを生成する

+ セクションの追加

① ソースデータが追加されている場合、一部のショットの例は使用されません。以前の例はすべてクリアされました。データを削除すると、例を追加できます。

> データを追加する
> パラメーター

チャットの履歴

新規 フィルタ 🎧 ⚙️ 🔍

サンプル プロンプトを使って始める

Dialogue creation Create a conversation between two characters who are meeting for the first time in a mysterious place.

Character development Describe a day in the life of a superhero who has lost their powers.

Recipe creation Invent a recipe for a dish that combines flavors from two different cuisines.

ここにユーザー クエリを入力します。 (Shift + Enter で新しい行)

マイク リセット

チャットプレイグラウンドのヘルプ

関連リソース チャットのクイックスタート

カスタム AI ソリューションを構築するための適切なモデルを見つける

アナウンス

o1 モデルを体験する



o1 シリーズは、科学とコーディングの問題を解決するために強化された推論機能を備えています。

[モデルのチェックアウト](#) [ブログを読む](#)

プロジェクトを作成して、より多くのモデルを取得する

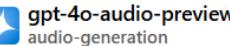
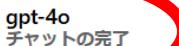
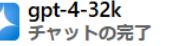
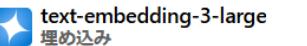
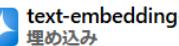
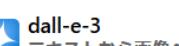


Meta、Mistral などのモデルを探索し、Azure AI サービスなどの新機能にアクセスできます。

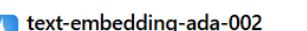
[その他のモジュールを探索](#) [Learn more](#)[推論タスク](#) [タスクの微調整](#)

検索

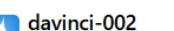
モデル 22

o1
チャットの完了o1-mini
チャットの完了gpt-4o-audio-preview
audio-generationgpt-4o-realtime-preview
audio-generationgpt-4o
チャットの完了gpt-4o-mini
チャットの完了o1-preview
チャットの完了gpt-4
チャットの完了gpt-4-32k
チャットの完了text-embedding-3-large
埋め込みtext-embedding-3-small
埋め込みtts
テキスト読み上げtts-hd
テキスト読み上げwhisper
音声認識dall-e-3
テキストから画像へ

dall-e-2



text-embedding-ada-002



davinci-002

モデルのヘルプ

How to use the model catalog

Search by name, filter, or browse to find the right model for your use case. Click to see model details and specifications from the publisher and to deploy a model.

Relevant resources

2025年2月現在、利用可能なモデル [Azureのドキュメント]

モデル	説明
○シリーズ モデル	高度な問題解決、増強された集中力と能力を備えた推論モデル。
GPT-4o、GPT-4o mini GPT-4 Turbo	最新の最も能力の高い Azure OpenAI モデルであり、テキストと画像の両方を入力として受け入れることができるマルチモーダルバージョンを備えています。
よく使われるのは GPT-4o	01系は利用登録が必要 声出力" の会話のやり取り、またはオーディオ生成をサポートする
GPT-4	GPT-3.5 を基に改善され、自然言語とコードを理解し、生成できるモデルのセット。
GPT-3.5	GPT-3 を基に改善され、自然言語とコードを理解し、生成できるモデルのセット。
埋め込み	テキストを数値ベクトル形式に変換して、テキストの類似性を促進できるモデルのセット。
DALL-E	自然言語からオリジナルの画像を生成できるモデルのシリーズ。
Whisper	音声を文字起こしして音声テキスト変換を翻訳できる一連のモデル。
テキスト読み上げ (プレ ビュー)	テキストを音声に合成できるプレビューステップの一連のモデル。

仮
ご

gpt-4o - Azure OpenAI Service | Product Availability by Region | Azure OpenAI Service - Pricing | Azure OpenAI Service モデル - Azu | Azure OpenAI Service モデル - Azu | +

https://ai.azure.com/resource/models/gpt-4o/version/2024-11-20/registry/azure-openai?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/p...

Azure AI Foundry | Azure OpenAI Service / gpt-4o-test-1 / モデル / gpt-4o

すべてのリソース 5 シンボル gpt-4o-test-1 (japaneast, S0) 8

gpt-4o

▷ デプロイ ▷ 微調整

詳細 既存のデプロイ コードサンプル ライセンス

gpt-4o offers a shift in how AI models interact with multimodal inputs. By seamlessly combining text, images, and audio, gpt-4o provides a richer, more engaging user experience.

Matching the intelligence of gpt-4 turbo, it is remarkably more efficient, delivering text at twice the speed and at half the cost. Additionally, GPT-4o exhibits the highest vision performance and excels in non-English languages compared to previous OpenAI models.

gpt-4o is engineered for speed and efficiency. Its advanced ability to handle complex queries with minimal resources can translate into cost savings and performance.

The introduction of gpt-4o opens numerous possibilities for businesses in various sectors.

詳細表示

モデルバージョン

リージョンの可用性に関する詳細情報 [?] Japan East

モデルID	可用性	ライフサイクル	最大要求	廃止予定
2024-05-13	グローバル標準, プロビジョニング	一般提供	Input: 128000	2025年5月20日 (火)

詳細表示

データ、メディア、言語

Property	説明
サポートされるデータ型	入力
	出力

クイック ファクト OpenAI

gpt-4o chat-completion

最後のトレーニング

リージョンによってはすべてのモデルが選べるわけではない

モデルのヘルプ

関連リソース

Azure OpenAI Service モデル

Azure OpenAI Service モデルの非推奨と廃止

gpt-4o - Azure OpenAI Service | Product Availability by Region | Azure OpenAI Service - Pricing | Azure OpenAI Service モデル - Azu | Azure OpenAI Service モデル - Azu | +

https://ai.azure.com/resource/models/gpt-4o/version/2024-11-20/registry/azure-openai?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/p...

Azure AI Foundry | Azure OpenAI Service / gpt-4o-test-1 / モデル / gpt-4o すべてのリソース 5 シンボル gpt-4o-test-1 (japaneast, S0) 8

gpt-4o

▶ デプロイ □ 微調整

詳細 既存のデプロイ コード

gpt-4o offers a shift in how AI models process text, images, and audio, gpt-4o provides

Matching the intelligence of gpt- speed and at half the cost. Additionally, gpt-4o can translate between non-English languages compared to other AI models.

gpt-4o is engineered for speed and efficiency, so minimal resources can translate into faster responses.

The introduction of gpt-4o opens up new possibilities for AI applications.

デプロイ名* gpt-4o-test

デプロイの種類 グローバル標準

Global Standard: 最も高い制限が適用された API 呼び出しあたりの支払。グローバル展開の種類 の詳細をご覧ください。

データはリソースの Azure 地域の外部でグローバルに処理される可能性がありますが、データストレージは AI リソースの Azure 地域に残ります。データ所在地 の詳細をご覧ください。

デプロイの詳細

モデルバージョン 2024-11-20

AI リソース / 空白 10223-m58are1c-eastus2

リソースの場所 East US 2

1 分あたりのトークン数 (TPM)

処理はEast US 2で行われる

Japan Eastにはリソースがない

選択したリソースにデプロイする 取消し

モデルのヘルプ

関連リソース

Azure OpenAI Service モデル

Azure OpenAI Service モデルの非推奨と廃止

リージョンによるモデルの対応

- [Azure OpenAI Service モデル - Azure OpenAI | Microsoft Learn](#)

Global-Standard モデルの提供状況

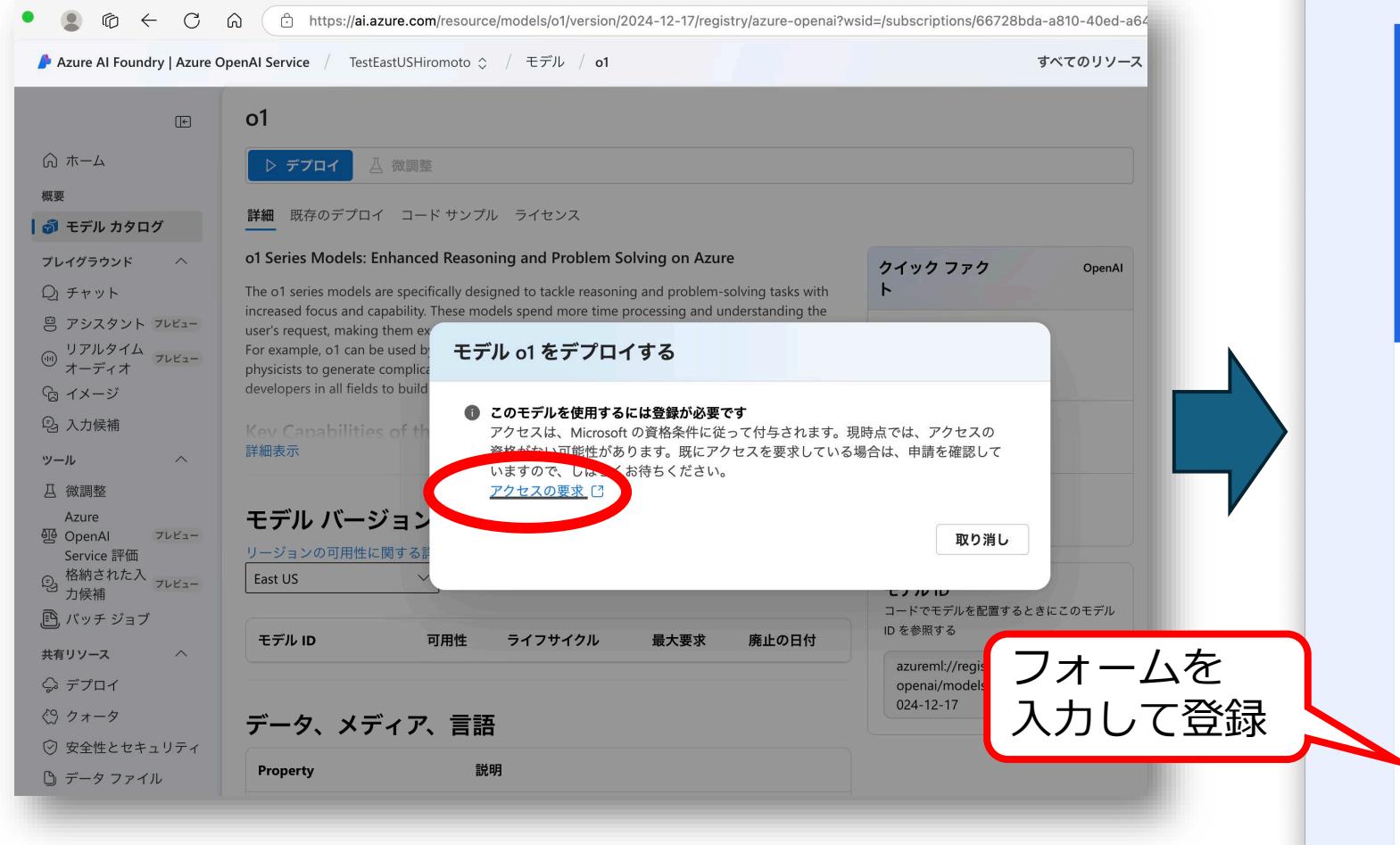
[] テーブルを展開する

リージョン	o1-preview、 2024-09-12	o1-mini、 2024-09-12	gpt-4o、 2024年5月13日	gpt-4o、 2024年8月6日	gpt-4o、 2024年11-20	gpt-4o-mini、 2024-07-18	gpt-4o-realtime-preview、 2024-10-01	gpt-4o-realtime-preview、 2024-12-17
australiaeast	-	-	✓	✓	-	✓	-	-
brazilsouth	-	-	✓	✓	-	✓	-	-
canadaeast	-	-	✓	✓	-	✓	-	-
eastus	✓	✓	✓	✓	✓	✓	-	-
eastus2	✓	✓	✓	✓	✓	✓	-	-
francecentral	-	-	✓	✓	-	✓	-	-
germanywestcentral	-	-	✓	✓	-	✓	-	-
japaneast	-	-	✓	✓	-	✓	-	-
koreacentral	-	-	✓	✓	-	✓	-	-

Japan Eastには2024-11-20の
gpt-4oはない
→ East US 2が自動選択される

2025年2月現在、GPT-o1系はアクセス要求が必要

- 「デプロイ」をクリックすると「このモデルを使用するには登録が必要です」と出る→「アクセスの要求」をクリック

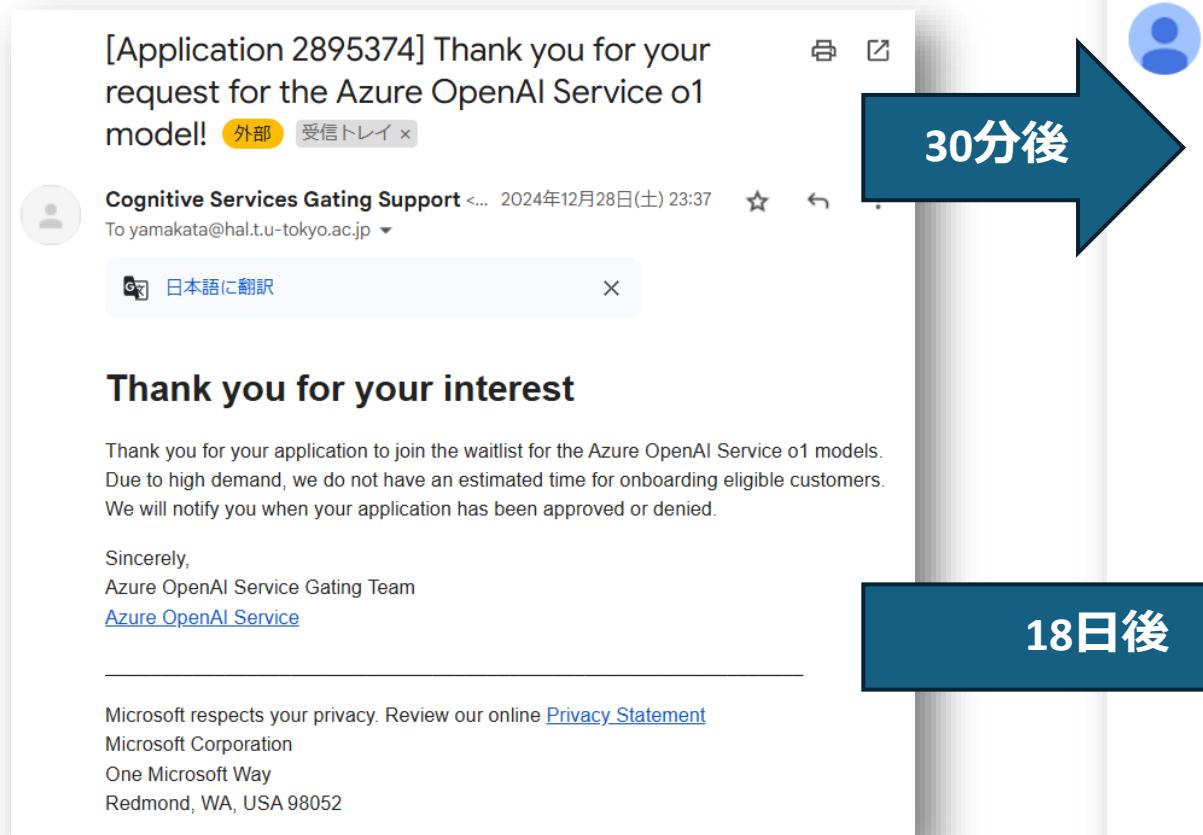


The screenshot shows the 'Request Access to the Azure OpenAI Service' form. The title is 'Request Access to the Azure OpenAI Service o* family of models'. The form includes instructions: 'Please read all instructions carefully and complete form as instructed'. It notes that 'The Azure OpenAI Service o* family of models require registration. Due to high demand, customers may remain on the waitlist until space is available.' A red speech bubble with the Japanese text 'フォームを入力して登録' (Input the form and register) points to the 'Your First Name' field at the bottom of the form.

しばらくすると承認される

私の場合…

- O1-previewとo1-miniはすぐ承認
- O1は15日後に承認



30分後



18日後





← gpt-4o-test

詳細 メトリック リスクと安全性

プレイグラウンドで開く

編集 削除

配置情報

名前	gpt-4o-test	プロビジョニングの状態	成功
展開の種類	グローバル標準	作成日	2025-01-30T08:19:22.6269308Z
作成者	1923306880@utac.u-tokyo.ac.jp	変更日	Jan 30, 2025 5:19 PM
更新者	1923306880@utac.u-tokyo.ac.jp	バージョン更新ポリシー	新しい既定のバージョンが使用可能になったら
レート制限 (1分あたりのトークン数)	10,000	レート制限 (1分あたりの要求数)	60
モデル名	gpt-4o	モデル バージョン	2024-11-20
ライフサイクルの状態	GenerallyAvailable	作成日	Dec 3, 2024 9:00 AM
更新日	Dec 3, 2024 9:00 AM	モデルの廃止日	Aug 20, 2025 9:00 AM

エンドポイント

ターゲット URI

[https://19233-m58are1c-eastus2.cognitiveservices.azure.com/...](https://19233-m58are1c-eastus2.cognitiveservices.azure.com/)

キー

.....

デプロイのヘルプ

Azure AI Foundry でのデプロイとは?

Azure AI Foundry では、大規模言語モデル(LLM)、フロー、Web アプリのデプロイがサポートされています。LLM またはフローをデプロイすると、それを Web サイトやアプリケーション、またはその他の運用環境で使用できるようになります。

[詳細情報](#)

関連リソース

監視と安全性

コンテンツ フィルター

DefaultV2

アプリケーション開発に役立つリンク

[コードサンプルリポジトリ](#)[チュートリアル](#)

3. チャットを行う

←

gpt-4o-test

詳細 メトリック リスクと安全性

プレイグラウンドで開く

編集 削除

ホーム

概要

モデル カタログ

プレイグラウンド ^

チャット

アシスタント プレビュー

リアルタイム プレビュー
オーディオ

イメージ

入力候補

ツール ^

微調整

Azure
OpenAI Service 評価

格納された入力候補 プレビュー

パッチ ジョブ

共有リソース ^

デプロイ

クォータ

安全性とセキュリティ

データ ファイル

ベクトルス トア プレビュー

配置情報

名前	gpt-4o-test	プロビジョニングの状態	成功
展開の種類	グローバル標準	作成日	2025-01-30T08:19:22.6269308Z
作成者	1923306880@utac.u-tokyo.ac.jp	変更日	Jan 30, 2025 5:19 PM
更新者	1923306880@utac.u-tokyo.ac.jp	バージョン更新ポリシー	新しい既定のバージョンが使用可能になったら
レート制限 (1分あたりのトークン数)	10,000	レート制限 (1分あたりの要求数)	60
モデル名	gpt-4o	モデル バージョン	2024-11-20
ライフサイクルの状態	GenerallyAvailable	作成日	Dec 3, 2024 9:00 AM
更新日	Dec 3, 2024 9:00 AM	モデルの廃止日	Aug 20, 2025 9:00 AM

エンドポイント

ターゲット URI

https://19233-m58are1c-eastus2.cognitiveservices.azure.com/...

キー

.....

監視と安全性

コンテンツ フィルター

DefaultV2

アプリケーション開発に役立つリンク

コードサンプル リポジトリ

チュートリアル

デプロイのヘルプ

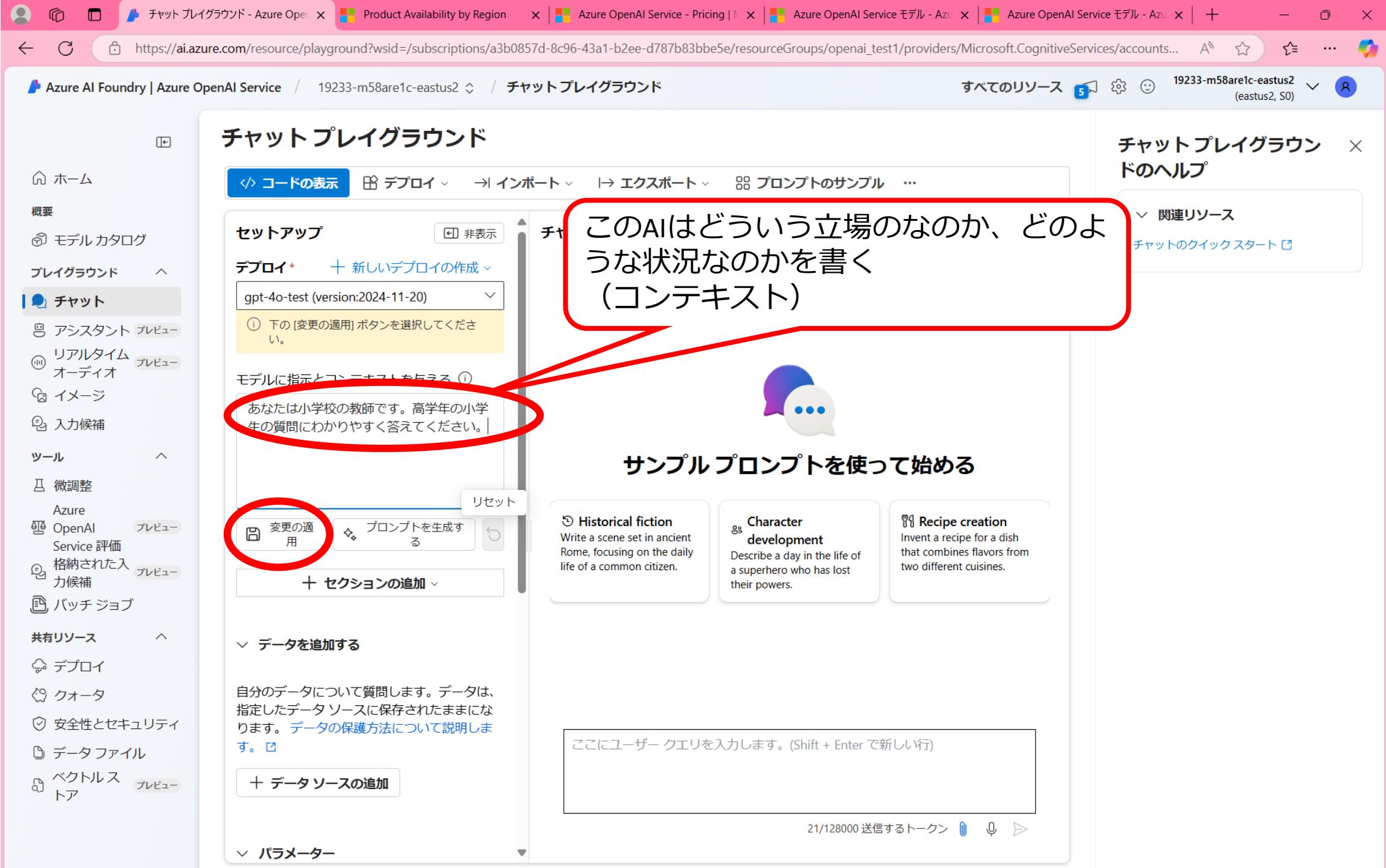
Azure AI Foundry でのデプロイ
とは?

Azure AI Foundry では、大規模言語モデル(LLM)、フロー、Web アプリのデプロイがサポートされています。LLM またはフローをデプロイすると、それを Web サイトやアプリケーション、またはその他の運用環境で使用できるようになります。

詳細情報

> 関連リソース





チャット プレイグラウンド - Azure OpenAI Service | Product Availability by Region | Azure OpenAI Service - Pricing | Azure OpenAI Service モデル - Azu | Azure OpenAI Service モデル - Azu | +

← ↻ 🔍 https://ai.azure.com/resource/playground?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/accounts... A ⭐ 🔍 ... 🎨

Azure AI Foundry | Azure OpenAI Service / 19233-m58are1c-eastus2 / チャット プレイグラウンド すべてのリソース 5 🚧 😄 19233-m58are1c-eastus2 (eastus2, S0) 🔍

チャット プレイグラウンド

コードの表示 デプロイ インポート エクスポート プロンプトのサンプル ...

セットアップ

デプロイ * + 新しいデプロイの作成
gpt-4o-test (version:2024-11-20)
① 下の [変更の適用] ボタンを選択してください。

モデルに指示とコンテキストを与える ①
あなたは小学校の教師です。高学年的小学生の質問にわかりやすく答えてください

変更の適用 プロンプトを生成する

+ セクションの追加

チャットの履歴

システム メッセージを更新しますか?

システム メッセージを更新すると、新しいチャットセッションが開始されます。以前のメッセージは新しい API 要求には含まれません。

今後、このメッセージを表示しない

続行 **キャンセル**

使って始める

Recipe creation
Invent a recipe for a dish that combines flavors from two different cuisines.

ここにユーザー クエリを入力します。(Shift + Enter で新しい行)

39/128000 送信するトークン 🔍 🔍 ▶

ホーム 概要 モデル カタログ ブレイグラウンド チャット アシスタント プレビュー リアルタイム オーディオ プレビュー イメージ 入力 候補 ツール 微調整 Azure OpenAI Service 評価 格納された 入力 候補 パッチ ジョブ 共有 リソース デプロイ クォータ 安全性とセキュリティ データ ファイル ベクトル ストア プレビュー 共有 リソース デプロイ クォータ 安全性とセキュリティ データ ファイル ベクトル ストア プレビュー パラメーター

チャット プレイグラウンド のヘルプ

関連 リソース チャット の クイック スタート

仮 https://ai.azure.com/resource/playground?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/accounts... A すべてのリソース 19233-m58are1c-eastus2 (eastus2, S0) チャットプレイグラウンド

チャットプレイグラウンド

コードの表示 デプロイ インポート エクスポート プロンプトのサンプル ...

セットアップ

デプロイ * + 新しいデプロイの作成
gpt-4o-test (version:2024-11-20)

モデルに指示とコンテキストを与える あなたは小学校の教師です。高学年的小学生の質問にわかりやすく答えてください。

変更の適用 プロンプトを生成する + セクションの追加

データを追加する

自分のデータについて質問します。データは、指定したデータソースに保存されたままになります。データの保護方法について説明します。 + データソースの追加

パラメーター

過去のメッセージを含む

チャットの履歴

応答形式 テキスト

サンプルプロンプトを使って始める

Historical fiction
Write a scene set in ancient Rome, focusing on the daily life of a common citizen.

Character development
Describe a day in the life of a superhero who has lost their powers.

Recipe creation
Invent a recipe for a dish that combines flavors from two different cuisines.

大規模言語モデルの仕組みについて教えてください。

67/128000 送信するトークン ▶

チャットプレイグラウンドのヘルプ

関連リソース チャットのクイックスタート

チャットプレイグラウンド

/ コードの表示 デプロイ インポート エクスポート プロンプトのサンプル ...

セットアップ

非表示

デプロイ * + 新しいデプロイの作成

gpt-4o-test (version:2024-11-20)

モデルに指示とコンテキストを与える ⓘ

あなたは小学校の教師です。高学年的小学生の質問にわかりやすく答えてください。

変更の適用 プロンプトを生成する

+ セクションの追加

> データを追加する

> パラメーター

チャットの履歴



|

応答形式 テキスト

大規模言語モデルの仕組みについて教えてください。

大規模言語モデル（例えばChatGPTなど）は、たくさんの文章を学習して、言葉と言葉のつながりやパターンを理解するように作られたAIの一種です。簡単に言うと、たくさんの本や記事、ウェブページなどを読んで、どんな言葉がどんなふうにつながるのかを覚えている「とても大きな辞書みたいなもの」と考えるとわかりやすいかもしれません。

仕組みを簡単に説明すると

1. 学習する（トレーニング）

AIは人間が書いたたくさんの文章を見て、「次にどんな言葉が来そうか」を予測する練習をします。例えば、「今日は空がとても〇〇です」という文章を見たとき、「〇〇」の部分には「青い」や「きれい」といった言葉が来そうだ、と予測するんです。

2. パターンを覚える

AIは言葉と言葉の関係や、文の流れをたくさん覚えます。例えば、「犬」という言葉が出てきたら「かわいい」「散歩」「吠える」などの関連する言葉を覚えたりします。

3. 質問に答える

質問されたり指示を出されたりすると、覚えたパターンを使って「一番自然な答え」を考えます。ただし、AIは実際に考えているわけではなく、あくまで「学んだパターンに基づいて次の言葉を予測している」だけです。

ここにユーザー クエリを入力します。(Shift + Enter で新しい行)

838/128000 送信するトークン

チャットプレイグラウンドのヘルプ

関連リソース

チャットのクイックスタート

チャット プレイグラウンド - Azure OpenAI Service | Product Availability by Region | Azure OpenAI Service - Pricing | Azure OpenAI Service モデル - Azu | Azure OpenAI Service モデル - Azu | +

https://ai.azure.com/resource/playground?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/accounts...

Azure AI Foundry | Azure OpenAI Service / 19233-m58are1c-eastus2 / チャット プレイグラウンド

すべてのリソース 5 19233-m58are1c-eastus2 (eastus2, S0)

チャット プレイグラウンド

コードの表示 デプロイ インポート エクスポート プロンプトのサンプル ...

セットアップ

デプロイ * + 新しいデプロイの作成
gpt-4o-test (version:2024-11-20)

モデルに指示とコンテキストを与える

あなたはAzureの専門家です。初めてAzureを使うユーザーに、わかりやすくAzureの使い方を教えてください。

変更の適用 プロンプトを生成する

+ セクションの追加

データを追加する

自分のデータについて質問します。データは、指定したデータソースに保存されたままになります。データの保護方法について説明します。

+ データソースの追加

> パラメーター

チャットの履歴

応答形式 テキスト

Azureで仮想マシンを起動して13BのLlama-3モデルをLoRA fine-tuningしたいと思います。リソースは何を選んだらいいですか？

Azureで13BのLlama-3モデルをLoRA (Low-Rank Adaptation) を使ってファインチューニングする場合、以下の手順とリソース選定のポイントを考慮してください。

1. 仮想マシンの選定

Llama-3のような大規模な言語モデルを扱う場合、GPU搭載VM (VM) が必要です。LoRAは従来のファインチューニングでも13BモデルはGPUメモリや計算リソースを大量に消費します。

途中までしか生成されない場合は、パラメータを調整

推奨される仮想マシンタイプ

Azureでは、以下のGPU搭載VMシリーズが適しています。

1. NDシリーズ (NVIDIA A100搭載)
例: ND96asr_v4

ここにユーザー クエリを入力します。(Shift + Enter で新しい行)

2104/128000 送信するトークン

チャット プレイグラウンドのヘルプ

関連リソース チャットのクイックスタート



ホーム

概要

モデル カタログ

プレイグラウンド ▾

チャット

アシスタント プレビュー

リアルタイム
オーディオ プレビュー

イメージ

入力候補

ツール ▾

微調整

Azure
OpenAI Service 評価

格納された入力候補 プレビュー

パッチ ジョブ

共有リソース ▾

デプロイ

クォータ

安全性とセキュリティ

データ ファイル

ベクトルス トア プレビュー

チャット プレイグラウンド

コードの表示

デプロイ

インポート

エクスポート

プロンプトのサンプル

...

自分のデータについて質問します。データは、指定したデータソースに保存されたままになります。[データの保護方法について説明します。](#)

+ データソースの追加

パラメーター

過去のメッセージを含む ⓘ

10

最大応答 ⓘ

2400

温度 ⓘ

0.7

上位 P ⓘ

0.95

シーケンスの停止 ⓘ

シーケンスの停止

頻度のペナルティ ⓘ

0

プレゼンスのペナルティ ⓘ

0

チャットの履歴



|

応答形式

テキスト



Azureで仮想マシンを起動して13BのLlama-3モデルをLoRA fine-tuningしたいと思います。リソースは何を選んだらいいですか？

Azureで13BのLlama-3モデルをLoRA (Low-Rank Adaptation) を使ってファインチューニングする場合、以下の手順とリソース選定のポイントを考慮してください。

1. 仮想マシンの選定

Llama-3のような大規模な言語モデルを扱う場合、GPUを搭載した仮想マシン（VM）が必要です。LoRAは従来のファインチューニングより軽量ですが、それでも13BモデルはGPUメモリや計算リソースを大量に使用します。

推奨される仮想マシンタイプ

Azureでは、以下のGPU搭載VMシリーズが適しています：

1. NDシリーズ (NVIDIA A100搭載)

- 例: ND96asr_v4

ここにユーザー クエリを入力します。(Shift + Enter で新しい行)

2104/128000 送信するトークン



チャット プレイグラウンド のヘルプ

関連リソース

チャットのクイックスタート

4. 外部のプログラムから 呼び出す

Google Colaboratoryのノートブックから、APIを介してOpenAI Service のモデルにプロンプトを送り、結果を受け取る

学校 Web アプリの作成 gpt-4o-test - Az App Service ブラ Web App Service microsoft web チャット プレイグラフ Azure AI Foundry Azure AI Foundry +

https://ai.azure.com/resource/deployments/%2Fsubscriptions%2Fa3b0857d-8c96-43a1-b2ee-d787b83bbe5e%2FresourceGroups%2Fopenai_test1%2Fproviders%2FMicrosoft.CognitiveServices%2FEndpoints%2Fgpt-4o-test

Azure AI Foundry | Azure OpenAI Service / 19233-m58are1c-eastus2 / デプロイ / gpt-4o-test すべてのリソース 19233-m58are1c-eastus2 (eastus2, S0)

← gpt-4o-test

ホーム 概要 モデル カタログ プレイグラウンド チャット アシスタント プレビュー リアルタイム オーディオ プレビュー イメージ 入力候補 ツール 微調整 Azure OpenAI Service 評価 格納された入力 候補 パッチ ジョブ デプロイ 安全性とセキュリティ データ ファイル ベクトルス トア

詳細 メトリック リスクと安全性 プレイグラウンドで開く 編集 削除

配置情報

名前	gpt-4o-test	プロビジョニングの状態	成功
展開の種類	グローバル標準	作成日	2025-01-30T08:19:22.6269308Z
作成者	1923306880@utac.u-tokyo.ac.jp	変更日	Jan 30, 2025 5:19 PM
更新者	1923306880@utac.u-tokyo.ac.jp	バージョン更新ポリシー	新しい既定のバージョンが使用可能になったら
レート制限 (1分あたりのトークン数)	10,000	レート制限 (1分あたりの要求数)	60
モデル名	gpt-4o	モデル バージョン	2024-11-20
ライフサイクルの状態	作成日		

エンドポイント

ターゲット URI
https://19233-m58are1c-eastus2.cognitiveservices.azure.com/...
キー
.....

監視と安全性

コンテンツ フィルター DefaultV2

アプリケーション開発に役立つリンク

コードサンプル リポジトリ チュートリアル

でコピーしてエンドポイントにセット

をクリックしてapiキーを表示させたうえで をクリックしてコピー

注意：エンドポイントの情報が知られてしまうと、誰でもあなたのリソースを使えてしまう（支払い義務はあなたに発生）ので、絶対に公開しないこと！

Google Colaboratoryサンプルコード

- ECCSアカウントにログインして[こちら](#)をColabで開いてください。
- 保存したい場合は「ドライブにコピー」をクリックして、ご自分のGoogle ドライブにコピーしてください。
- 「2. パラメータの設定」にある以下の部分を書き換えてください。

```
[ ] import os
```

```
# 以下の[YOUR TARGET URL]をエンドポイントのターゲットURLに書き換える  
os.environ["AZURE_OPENAI_ENDPOINT"] = "[YOUR TARGET URL]"
```

```
# 以下の[YOUR KEY]をエンドポイントのキーに書き換える  
os.environ["AZURE_OPENAI_KEY"] = "[YOUR KEY]"
```

ターゲットURLに書き換え

キーに書き換え

```
▶ from openai import AzureOpenAI
```

```
client = AzureOpenAI(  
    azure_endpoint = os.getenv("AZURE_OPENAI_ENDPOINT"),  
    api_key = os.getenv("AZURE_OPENAI_KEY"),  
    api_version="2024-08-01-preview",  
)
```

```
response = client.chat.completions.create(  
    model=os.getenv("AZURE_OPENAI_MODEL"),  
    messages=[
```

```
        {"role": "system", "content": "あなたはAzureの専門家です。初めてAzureを使うユーザに、わかりやすくAzu  
        {"role": "user", "content": "Azure OpenAI Serviceでデプロイしたgpt-4oのモデルにGoogle Colab上で動かす"},  
    ],
```

```
    max_tokens=800,  
    temperature=0.7,  
    top_p=0.95,  
    frequency_penalty=0,  
    presence_penalty=0,  
    stop=None,  
    stream=False  
)
```

コンテキストとユーザクエリを
設定。

```
▶ print(response.choices[0].message.content)
```

➡ Azure OpenAI ServiceでデプロイしたGPT-4 (gpt-4o) モデルにGoogle Colabからアクセスするには、以下の手順を実行します。

1. **Azure OpenAI Serviceの準備**

モデルが返してきたメッセージを書き出すコード。

まず、Azure上でOpenAI Serviceを設定し、GPT-4のモデルをデプロイします。

手順

1. **Azureポータルにログイン**:

- [Azureポータル] (<https://portal.azure.com>) にアクセスし、アカウントでログインします。

2. **Azure OpenAIリソースを作成**:

- Azureポータルの検索バーで「OpenAI」を検索し、「Azure OpenAI」を選択します。
- 「+ 作成」をクリックしてリソースを作成します。
- サービス名やリージョンを選択し、リソースをデプロイします。

3. **モデルをデプロイ**:

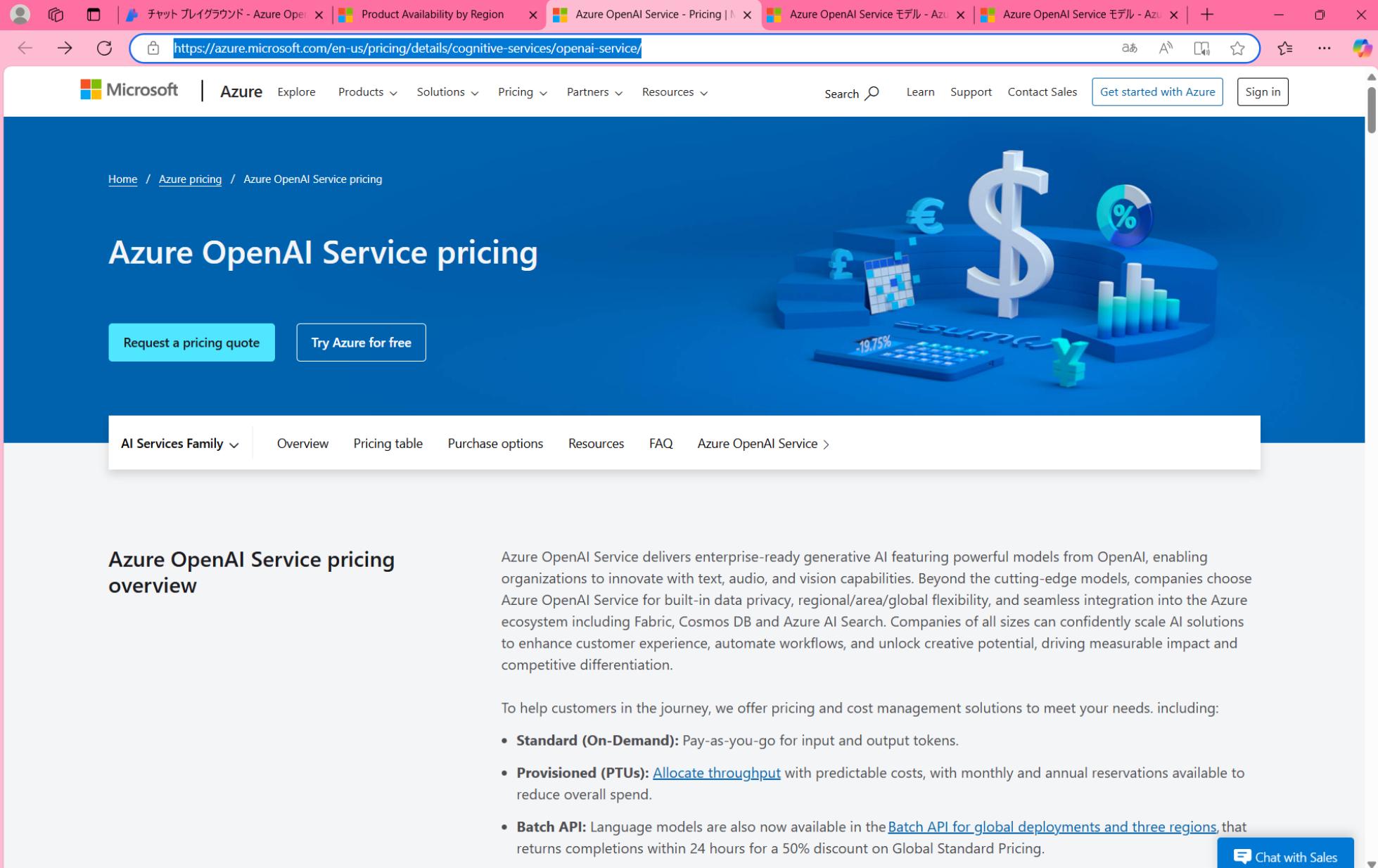
- 作成したAzure OpenAIリソースに移動し、「モデルデプロイ」を選択。
- 使用したいモデル（例: `gpt-4` または `gpt-4o`）を選択してデプロイします。
- デプロイ名（例: `gpt-4o-model`）を指定してデプロイを完了させます。

5. その他の話題

5. 1 OpenAI Serviceの価格

料金の見積もり

Azure OpenAI Serviceの価格 ([こちらを参照](#))



The screenshot shows the Azure OpenAI Service pricing page. At the top, there are four tabs: 'Product Availability by Region', 'Azure OpenAI Service - Pricing', 'Azure OpenAI Service モデル - Azu', and 'Azure OpenAI Service モデル - Azo'. Below the tabs, the URL is https://azure.microsoft.com/en-us/pricing/details/cognitive-services/openai-service/. The Microsoft navigation bar includes 'Microsoft' logo, 'Azure', 'Explore', 'Products', 'Solutions', 'Pricing', 'Partners', 'Resources', 'Search', 'Learn', 'Support', 'Contact Sales', 'Get started with Azure', and 'Sign in'. The main content area has a blue background featuring a large dollar sign and various currency symbols (€, £, ¥) with charts and graphs. The title 'Azure OpenAI Service pricing' is centered above two buttons: 'Request a pricing quote' and 'Try Azure for free'. Below the title, a navigation bar includes 'AI Services Family', 'Overview', 'Pricing table', 'Purchase options', 'Resources', 'FAQ', and 'Azure OpenAI Service >'. The main content section starts with 'Azure OpenAI Service pricing overview' and a detailed description of the service's capabilities and integration with the Azure ecosystem. It then lists three pricing models: Standard (On-Demand), Provisioned (PTUs), and Batch API.

Azure OpenAI Service pricing overview

Azure OpenAI Service delivers enterprise-ready generative AI featuring powerful models from OpenAI, enabling organizations to innovate with text, audio, and vision capabilities. Beyond the cutting-edge models, companies choose Azure OpenAI Service for built-in data privacy, regional/area/global flexibility, and seamless integration into the Azure ecosystem including Fabric, Cosmos DB and Azure AI Search. Companies of all sizes can confidently scale AI solutions to enhance customer experience, automate workflows, and unlock creative potential, driving measurable impact and competitive differentiation.

To help customers in the journey, we offer pricing and cost management solutions to meet your needs. including:

- **Standard (On-Demand):** Pay-as-you-go for input and output tokens.
- **Provisioned (PTUs):** [Allocate throughput](#) with predictable costs, with monthly and annual reservations available to reduce overall spend.
- **Batch API:** Language models are also now available in the [Batch API for global deployments and three regions](#), that returns completions within 24 hours for a 50% discount on Global Standard Pricing.

[Chat with Sales](#)

OpenAI GPTの料金

GPT-4o

GPT-4o is the most advanced multimodal model that's faster and cheaper than GPT-4 Turbo with stronger vision capabilities. The model has 128K context and an October 2023 knowledge cutoff.

Model	Pricing (1M Tokens)	Pricing with Batch API (1M Tokens)
GPT-4o-2024-1120 Global	Input: ¥393.88751 Cached Input: ¥196.9438 Output: ¥1,575.5501	Input: ¥196.94376 Output: ¥787.7751

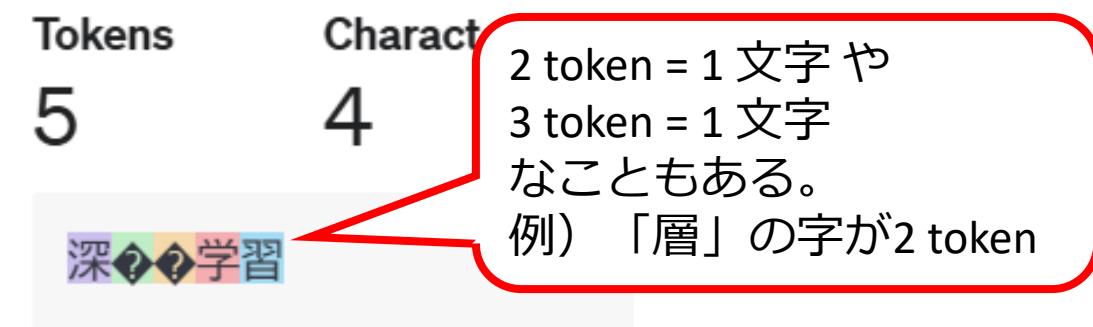
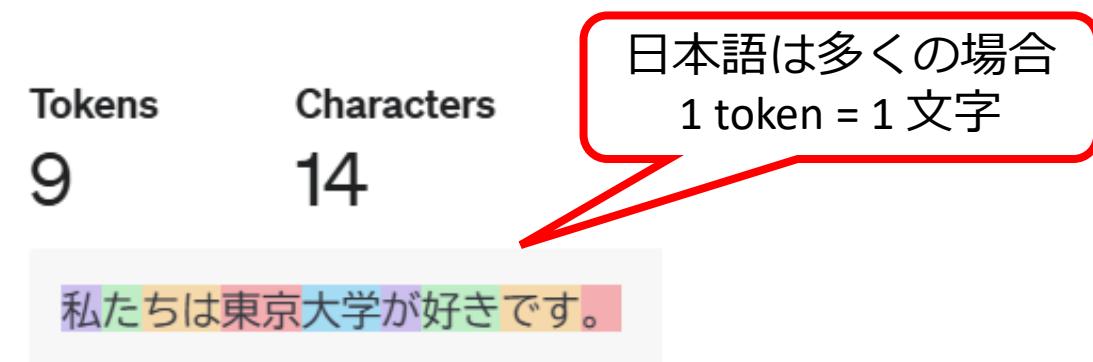
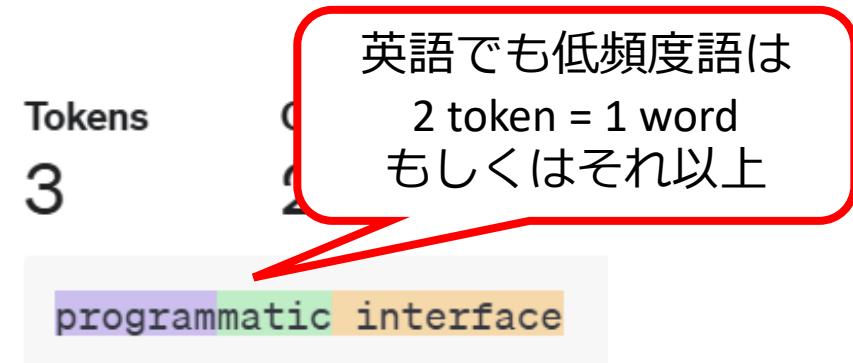
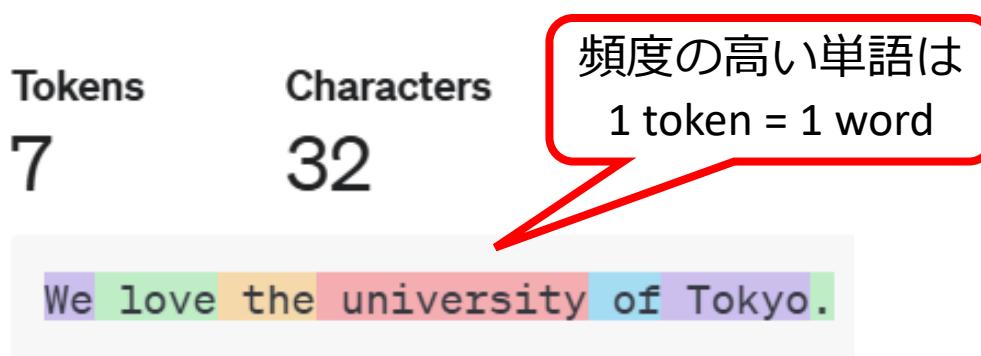
o1

o1 is the new reasoning model series for complex tasks. The model has 200K context and an October 2023 knowledge cutoff.

Model	Pricing (1M Tokens)
o1 1217 Global	Input: ¥2,363.3251 Cached Input: ¥1,181.6626 Output: ¥9,453.300060

Azure OpenAI Serviceの価格

- ・入力文（Input）と出力文（Output）のtoken数によって決まる
- ・OpenAIにおけるtoken数の調べ方：[Web版tokenizer](#)
- ・英語は100 tokens ≈ 75 words だが、低頻度語は対応するtoken数が増える
- ・日本語は全言語から見れば低頻度→英語よりtoken数が圧倒的に多い



学校 gpt-4o-test GitHub - A IImjp2 - M Pricing - L Pricing - M Azure Disk azure port サポートリク vCPU のク IIm-jp/IIm +

https://ai.azure.com/resource/deployments/%2Fsubscriptions%2Fa3b0857d-8c96-43a1-b2e0-707f22115625

Azure AI Foundry | Azure OpenAI Service / 19233-m58are1c-eastus2 / デプロイ / gpt-4o-test 19233-m58are1c-eastus2 (eastus2, S0) ヘルプ

「デプロイ>[名前]>メトリック」で
トークン数を確認できる。

ホーム
概要
モデルカタログ
プレイグラウンド
チャット
アシスタント プレビュー
リアルタイム オーディオ プレビュー
イメージ
入力候補
ツール
微調整
Azure OpenAI Service 評価
格納された入力候補 プレビュー
パッチジョブ
デプロイ
データファイル
ベクトルストア プレビュー

← gpt-4o-test

詳細 メトリック リスクと安全性

Azure Monitor で表示 01/24/2025 - 01/31/2025 最終日 7日 1か月 ...

7 要求の合計数
1.35万 トーカンの合計数
4918 プロンプト トーカン数
8539 完了トーカン数

トーカンの使用状況

5772
5291
4810
4329
3848
3367
2886
2405
1924
1443
962
481
0

24. Jan 25. Jan 26. Jan 27. Jan 28. Jan 29. Jan 30. Jan 31. Jan 1. Feb

● トーカンの合計数 ■ プロンプト トーカン総数 □ 完了トーカン数

クリップボード (24/24)
アイテムは収集されません。クリップボードの容量以上のサイズです。

5.2 1分当たりのトークン制限

Azure OpenAI Serviceにはクオータ制限がある

gpt-4o-test-1 - Microsoft Azure | モデル クオータ - Azure OpenAI Service +

https://ai.azure.com/resource/quota?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/accounts/19233-m58are1c-eastus2

Azure AI Foundry | Azure OpenAI Service / 19233-m58are1c-eastus2 / モデル クオータ

すべてのリソース 5 19233-m58are1c-eastus2 (eastus2, S0) ヘルプ

クオータの使用状況の監視および追跡

サブスクリプションとリージョン別にクオータを表示し、デプロイ全体の使用状況を追跡します。デプロイを作成するにはクオータが必要であり、トラフィックのニーズに応じて柔軟にサイズを変更できます。

サブスクリプション ①
1923306880-YamakataLab

Azure OpenAI Service Standard + パッチ Azure OpenAI Service がプロビジョニングされました

クオータの要求 最新の情報に更新 ビューのリセット グループ化 クオータの種類、モデルとリージョン

デプロイ ↑	モデル名	バージョン	クオータ...	リージョン	リソース	クオ...
> GlobalStandard						

検索 フィルター すべてのクオータを表示する 列

可用性が最も低いクオータ

● 使用可能なクオータ
● 割り当てられたクオータ

OpenAI Service モデルのクオータはサブスクリプションレベルで管理され、リージョンおよびモデルの種類ごとに指定されます。同じリージョンで、1つまたは複数の Azure AI リソースを実行している同じ種類のすべてのデプロイは、1つのクオータ制限を共有します。

リージョンとクオータの種類全体におけるモデルの可用性

ホーム 概要 モデル カタログ ブレイグランド チャット アシスタント プレビュー リアルタイム オーディオ プレビュー イメージ 入力候補 ツール 微調整 Azure OpenAI Service 評価 格納された入 力候補 パッチ ジョブ 共有リソース テプロトコル クオータ 安全性とセキュリティ データ ファイル ベクトルス トア プレビュー

ホーム

概要

モデル カタログ

ブレイクラウンド ^

チャット

アシスタント プレビュー

リアルタイム オーディオ プレビュー

イメージ

入力候補

ツール ^

微調整

Azure OpenAI Service 評価

格納された入力候補 プレビュー

パッチ ジョブ

共有リソース ^

デプロイ

クオータ

安全性とセキュリティ

データ ファイル

ベクトルス トア プレビュー

クオータの使用状況の監視および追跡

サブスクリプションとリージョン別にクオータを表示し、デプロイ全体の使用状況を追跡します。デプロイを作成するにはクオータが必要であり、トラフィックのニーズに応じて柔軟にサイズを変更できます。

サブスクリプション ⓘ

1923306880-YamakataLab

Azure OpenAI Service Standard + パッチ

Azure OpenAI Service がプロビジョニング中

クオータの要求

最新の情報に更新

ビューのリセット

検索

デプロイ ↑

モデル名

バージョン

ク

GlobalStandard

gpt-4o

GlobalStandard

East US 2

gpt-4o

GlobalStandard

20K

gpt-4o

2024-11-20

GlobalStandard

East US 2

aifoundrytest0...

10KTPM

gpt-4o-test

2024-11-20

GlobalStandard

East US 2

19233-m58are1...

10KTPM

o1

GlobalStandard

East US 2

o1

GlobalStandard

10K の 5M TPM

- Kは1000、TPMはTokens-per-Minuteで、1分当たりのトークン数
- 初期状態では1分に10000トークンの制限がかかっている
- モデル、リージョンによってクオータは異なる（参照）

クオータの種類、モデルとリージョン

クオータ

クオータ

使用可能なクオータ

割り当てられたクオータ

OpenAI Service モデルのクオータはサブスクリプションレベルで管理され、リージョンおよびモデルごとに指定されます。同じリージョンで、1つ複数の Azure AI リソースを実行している同じ種別のデプロイは、1つのクオータ制限を共有します。

リージョンとクオータの種類全体におけるルールの可用性

gpt-4o-test-1 - Microsoft Azure モデルクオータ - Azure OpenAI Service +

https://ai.azure.com/resource/quota?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/accounts/19233-m58are1c-eastus2

Azure AI Foundry | Azure OpenAI Service / 19233-m58are1c-eastus2 / モデルクオータ すべてのリソース 5 19233-m58are1c-eastus2 (eastus2, S0) 詳細情報

クオータの使用状況の監視および追跡

サブスクリプションとリージョン別にクオータを表示し、デプロイ全体の使用状況を追跡します。デプロイを作成するにはクオータが必要であり、トラフィックのニーズに応じて柔軟にサイズを変更できます。

サブスクリプション 1923306880-YamakataLab

Azure OpenAI Service Standard + パッチ Azure OpenAI Service がプロビジョニングされました

このデプロイに割り当てられたクオータの編集

容量 Tokens Per Minute (thousands) - gpt-4o - GlobalStandard
1120K/29990K

確認 キャンセル

スライドバーで希望する
サイズまで引き上げ

グループ化 クオータの種類、モデルとリージョン

可用性が最も低いクオータ

OpenAI Service モデルのクオータはサブスクリプションレベルで管理され、リージョンおよびモデルごとに指定されます。同じリージョンで、1つ以上の複数の Azure AI リソースを実行している場合、すべてのデプロイは、1つのクオータ制限を共有します。

リージョンとクオータの種類全体における
ルの可用性

リージョン	モデル	クオータの種類	可用性
East US 2	gpt-4o	GlobalStandard	20K の 30M TPM
East US 2	gpt-4o-test	GlobalStandard	10KTPM
East US 2	o1	GlobalStandard	10K の 5M TPM

gpt-4o-test-1 - Microsoft Azure モデル クォータ - Azure OpenAI Service + https://ai.azure.com/resource/quota?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/openai_test1/providers/Microsoft.CognitiveServices/accounts/19233-m58are1c-eastus2... A ⚡ ⚡ ...

Azure AI Foundry | Azure OpenAI Service / 19233-m58are1c-eastus2 / モデル クォータ すべてのリソース 5 🚧 🚧 19233-m58are1c-eastus2 (eastus2, S0) ⚡ ⚡ ⚡

クォータの使用状況の監視および追跡

サブスクリプションとリージョン別にクォータを表示し、デプロイ全体の使用状況を追跡します。デプロイを作成するにはクォータが必要であり、トラフィックのニーズに応じて柔軟にサイズを変更できます。

詳細情報

サブスクリプション ① 1923306880-YamakataLab

Azure OpenAI Service Standard + パッチ Azure OpenAI Service がプロビジョニングされました

クォータの要求 最新の情報に更新 ビュー

検索 検索

デプロイ ↑ モデル名

- GlobalStandard
- gpt-4o
- East US 2 gpt-4o GlobalStandard 1.13M の 30M TPM
- gpt-4o 2024-11-20 GlobalStandard 10KTPM 増えた！
- gpt-4o-test 2024-11-20 GlobalStandard East US 2 19233-m58are1c-eastus2 1.12MTPM
- o1 GlobalStandard

グループ化 クォータの種類、モデルとリージョン

可用性が最も低いクォータ

クォータの合計 10M 0 クォータ (モデル + リージョン) クォータの種類

● 使用可能なクォータ ● 割り当てられたクォータ

OpenAI Service モデルのクォータはサブスクリプションレベルで管理され、リージョンおよびモデルごとに指定されます。同じリージョンで、複数の Azure AI リソースを実行している同じデプロイは、1 つのクォータ制限を共有します。

リージョンとクォータの種類全体における可用性

初期状態では30M TPMが上限
もっと増やしたいときは
クォータの引き上げを要求
(承認されるまで数日かかる！)

増えた！

クオータの 引き上げを要求

- クオータ増加リクエスト
フォーム
- 氏名やメールアドレス、組織名、要求タイプ・サイズ
- なぜクオータ引き上げをしたいのかを説明
(Justification)
- 人力で判定している?
承認まで数日かかる
(早めにやっておく)

The screenshot shows a web browser window with a pink header bar. The address bar contains the URL <https://customervoice.microsoft.com/Pages/ResponsePage.aspx?id=v4j5cvGGr0GRqy180B...>. The main content area has a black header with the Microsoft Azure logo. Below it, the title "Azure OpenAI Service: Request for Quota Increase" is displayed. A descriptive text explains that the form is used to submit requests for quota increases in quota for the standard deployment type. It states that requests are being accepted and will be filled in the order received, with priority given to customers who generate traffic that consumes the existing quota allocation. It also notes that requests may be denied if the condition is not met. At the bottom left, there is a note in red text: "* 必須". The main body of the form contains instructions: "Read the instructions carefully and answer each question completely before submitting the request". It also provides a note: "Use this form to request an increase due to your forecasted usage for Azure OpenAI Service. Microsoft will use the information you provide to assess your usage volume and patterns, allowing us to allocate the necessary GPU capacity to support your work. We will make every effort to accommodate your request; however, allocation is based on our current capacity and future deployments, and is subject to availability." A "Please Note:" section at the bottom states: "Please Note: For Microsoft personnel, do not fill out this form. Find more information [here](#)." The form includes a field labeled "1. First Name *".

5.3 Azure OpenAI Service のデータ、 プライバシー、セキュリティ

データはAzureやOpenAIに使われるのか？

Data, privacy, and security for Azure OpenAI Service - Azure AI services | Microsoft Learn

プロンプト（入力）と完成品（出力）、埋め込み、トレーニング データ：

- ・他のお客様はご利用いただけません。
- ・OpenAIには利用できません。
- ・OpenAIモデルの改善には使用されません。
- ・は Azure OpenAI Service 基盤モデルのトレーニング、再トレーニング、または改善には使用されません。
- ・お客様の許可または指示なしに、Microsoft または第三者の製品またはサービスを改善するために使用することはありません。
- ・微調整された Azure OpenAI モデルは、お客様専用に使用できます。

Azure OpenAI Service は、Microsoft が Azure サービスとして運営しています。Microsoft は Microsoft の Azure 環境で OpenAI モデルをホストしており、本サービスは OpenAI が運営するサービス (ChatGPT や OpenAI API など) と対話しません。

目的

- 学習済みモデルを利用する2つの選択肢
 - OpenAIのモデルをすぐに使いたい → Azure OpenAI Service
 - OpenAIを含め様々なモデルを使いたい → Azure AI Foundry
- Azure AI Foundryのロードマップ
 1. Hubを作成してからProjectを作成
 2. モデルカタログからモデルを選択（たとえば”deepseek-R1”を選ぶ）
 3. チャットプレイグラウンドからチャットを実行

DeepSeekを使う

OpenAI以外のモデルを使いたいならAI Foundryを利用

2025年2月現在のモデル一覧

最新版は[こちら](#)を参照

会社	モデル	タイプ
AI21 Labs	AI21-Jamba-1.5-Mini, AI21-Jamba-1.5-Large	chat-completion
Azure OpenAI	o3-mini, o1, o1-preview, o1-mini, gpt-4o, gpt-4o-mini	chat-completion
Azure OpenAI	gpt-4o-realtime-preview	real-time
Azure OpenAI	text-embedding-3-large, text-embedding-3-small	embeddings
Cohere	Cohere-embed-v3-english, Cohere-embed-v3-multilingual	embeddings
Cohere	Cohere-command-r-plus-08-2024, Cohere-command-r-08-2024, Cohere-command-r-plus, Cohere-command-r	chat-completion
Core42	jais-30b-chat	chat-completion
DeepSeek	DeepSeek-R1	chat-completion
Meta	Llama-3.3-70B-Instruct, Llama-3.2-11B-Vision-Instruct, Llama-3.2-90B-Vision-Instruct, Meta-Llama-3.1-405B-Instruct, Meta-Llama-3.1-8B-Instruct, Meta-Llama-3.1-70B-Instruct, Meta-Llama-3-8B-Instruct, Meta-Llama-3-70B-Instruct	chat-completion
Microsoft	Phi-3-mini-128k-instruct, Phi-3-mini-4k-instruct, Phi-3-small-8k-instruct, Phi-3-medium-128k-instruct, Phi-3-medium-4k-instruct, Phi-3.5-vision-instruct, Phi-3.5-MoE-instruct, Phi-3-small-128k-instruct, Phi-3.5-mini-instruct, Phi-4	chat-completion
Mistral AI	Mistral-3B, Mistral-large, Mistral-small, Mistral-Nemo, Mistral-large-2407, Mistral-Large-2411, Codestral-2501	chat-completion
NTT Data	Tsuzumi-7b	chat-completion

1. HubとProjectを作成

学校 ホーム - Microsoft チャット プレイグラウンド azure foundry Azure の料金概要 azure price - 検索 Azure Pricing 料金計算ツール Product Availability

https://portal.azure.com/#home

Microsoft Azure

AI Foundry Copilot 9 8 7 6 5 4 3 2 1 ? ヘルプ

1923306880@ut.ac.jp THE UNIVERSITY OF TOKYO

AI Foundry サービス (13) サービス

Azure AI Foundry AI Search Community Training Azure OpenAI

Azure AI Foundry Marketplace もっと見る

Azure AI Foundry アーキテクチャ - Azure AI Foundry Azure AI Foundry SDK を使い始める方法 - Azure AI Foundry Azure AI Foundry ポータルでモデル カタログについて確認する - Azure AI Foundry Azure AI サービスのドキュメント

Microsoft Entra ID サービス プリンシパル

AI-Foundry-test0 ai-foundry-project0 Microsoft Entra ID で検索を続行してください

すべて表示 Azure AI services フィードバックの送信

リソース

最近 お気に入り

名前

- AI-Foundry-test0
- AI-Foundry-hub0
- 1923306880-3791_ai
- rg-1923306880-1245_ai
- deepseek0
- openai_test1
- 19233-m58are1c-eastus2
- 1923306880-0528
- deepseek00567883523
- 1923306880-YamakataLab
- gpt-4o-test-1
- ai-19233068803791ai761221281178

最終表示日

- 10 時間前
- 10 時間前
- 11 時間前
- 22 時間前
- 3 日前
- 4 日前

Azure AI Foundry を選択

リソースの作成 リソース グループ

クォータ Azure Machine... その他のサービス

移動

学校 Azure AI Foundry - Microsoft Azure チャット プレイグラウンド - Azure AI Foundry azure foundry sku - 検索 Models available in Azure AI mod +

https://portal.azure.com/#browse/Microsoft.MachineLearningServices%2Faistudio

Microsoft Azure リソース、サービス、ドキュメントの検索 (G+/)

Copilot 7 セットアップ ヘルプ フィードバック THE UNIVERSITY OF TOKYO 1923306880@ut.ac.jp

ホーム > Azure AI Foundry The University of Tokyo

+ Create ビューの管理 CSV

Project Collaborate, organize, and track work to build AI apps.

Hub Grouping container for projects. Provides security, connectivity, and compute management.

1923306880-5791_ai ai-foundry-project AI-Foundry-test0 aifoundrytest00903255598

リソース グループ ↑↓ の値と等しいすべて の値と等しいすべて リソース グループ 次の値と等しいすべて 場所 次の値と等しいすべて フィルターの追加

グループ化なし リストビュー

リソース グループ ↑↓	種類 ↑↓	場所 ↑↓	サブスクリプション ↑↓
openai_test1	Azure AI services	East US 2	1923306880-YamakataLab
rg-1923306880-1245_ai	Azure AI hub	East US 2	1923306880-YamakataLab
AI-Foundry-hub0	Azure AI project	East US 2	1923306880-YamakataLab
AI-Foundry-hub0	Azure AI hub	East US 2	1923306880-YamakataLab
AI-Foundry-hub0	Azure AI services	East US 2	1923306880-YamakataLab

まずHubを作る
(Projectから始めても、既存のHubがなければ作ることになる)

<前へ ページ 1 / 1 次へ> フィードバックの送信

学校 Azure AI ハブ - Microsoft Azure | azure foundry sku - 検索 | Models available in Azure AI mod | +

https://portal.azure.com/#view/Microsoft_Azure_MLTeamAccounts/CreateAIStudioResourceBlade/_provisioningContext~/%7B"initialValues"%3A%7B"subscriptionIds"%3A%5B"a3b0... 6 A? ? Copilot

Microsoft Azure リソース、サービス、ドキュメントの検索 (G+/)

1923306880@utac.u... THE UNIVERSITY OF TOKYO

ホーム > Azure AI Foundry >

Azure AI ハブ

Azure AI ハブリソースの作成

基本 ストレージ ネットワーク 暗号化 ID タグ 確認および作成

組織

デプロイされたリソースとコストを管理するサブスクリプションを選択してください。フォルダーなどのリソース グループを使用してすべてのリソースを整理し管理します。AI ハブは、チームがプロジェクト作業、モデル エンドポイント、コンピューティング、(データ) 接続、セキュリティ設定を共有し、使用状況を管理するためのコラボレーション環境です。

サブスクリプション * ① 1923306880-YamakataLab

リソース グループ * ① (新規) AI-Foundry-hub0
新規作成

リージョン * ① East US 2

リソースの詳細

名前 * ① AI-Foundry-test0

フレンドリ名 ① AI Foundry test0

既定のプロジェクトリソース グループ ①

Same as hub resource group

Azure AI サービスのベース モデル

OpenAI を含む AI サービスに接続する * ① (新規) aifoundrytest00903255598
新規作成

確認および作成

次へ: ストレージ

インタフェースは Azure OpenAI Service と概ね同じ

学校 Azure AI ハブ - Microsoft Azure | azure foundry sku - 検索 | Models available in Azure AI mod | +

https://portal.azure.com/#view/Microsoft_Azure_MLTeamAccounts/CreateAIStudioResourceBlade/_provisioningContext~/%7B"initialValues"%3A%7B"subscriptionIds"%3A%5B"a3b0... 6 A? ? Copilot

Microsoft Azure リソース、サービス、ドキュメントの検索 (G+/)

Copilot 6 A? ? Copilot 1923306880@utac.u... THE UNIVERSITY OF TOKYO

ホーム > Azure AI Foundry >

Azure AI ハブ

Azure AI ハブリソースの作成

✓ 検証に成功しました

基本 ストレージ ネットワーク 暗号化 ID タグ 確認および作成

基本

サブスクリプション	1923306880-YamakataLab
リソース グループ	(新規) AI-Foundry-hub0
名前	AI-Foundry-test0

リソース

リージョン	East US 2
AI サービス	(新規) aifoundrytest00903255598
ストレージ アカウント	(新規) aifoundrytest02974364556
キー コンテナー	(新規) aifoundrytest02322244864
Application Insights	なし
コンテナー レジストリ	なし

ネットワーク

接続方法	すべてのネットワークからのパブリック アクセスを有効にする
ネットワークの分離	パブリック

暗号化

暗号化の種類	Microsoft マネージド キー
--------	--------------------

ID

ID の種類: フォールバックハンドル

作成 < 前へ 次へ > Automation のテンプレートをダウンロードする

The 'Create' button at the bottom left is circled in red.

学校 Microsoft.MachineLearningService x azure foundry sku - 検索 x | Models available in Azure AI mod x +

https://portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/~/overview/id/%2Fsubscriptions%2Fa3b0857d-8c96-43a1-b2ee-d787b83bbe5e%2FresourceGroups%2FAI-... 7 A? ? ⚡ ... 🎨

Microsoft Azure リソース、サービス、ドキュメントの検索 (G+/) Copilot 1 THE UNIVERSITY OF TOKYO 1923306880@utac.u...

ホーム > Microsoft.MachineLearningServices | 概要 ×

デプロイ

検索 削除 キャンセル 再デプロイ ダウンロード 最新の情報に更新

概要

入力 出力 テンプレート

... デプロイが進行中です

デプロイ名 : Microsoft.MachineLearningServices
サブスクリプション : 1923306880-YamakataLab
リソース グループ : AI-Foundry-hub0

開始日時 : 2025/2/8 23:22:50
関連付け ID : 3df3a6f4-caef-4654-86c7-36b50a2049c2

デプロイの詳細

リソース	種類	状態	操作の詳細
aifoundrytest02974364556	ストレージ アカウント	Accepted	操作の詳細
aifoundrytest02322244864	キー コンテナー	OK	操作の詳細

Microsoft Defender for Cloud
アプリおよびインフラストラクチャをセキュリティで保護します
次に移動: Microsoft Defender for Cloud >

無料の Microsoft チュートリアル
今すぐ学習を開始する >

エキスパートとの共同作業
Azure のエキスパートは、Azure での資産の管理を支援し、サポートの最前線となるサービス プロバイダー パートナーです。
Azure Expert の検索 >

学校 Microsoft.MachineLearningService x azure foundry sku - 検索 x | Models available in Azure AI mod x +

https://portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/~/overview/id/%2Fsubscriptions%2Fa3b0857d-8c96-43a1-b2ee-d787b83bbe5e%2FresourceGroups%2FAI-... 7 A THE UNIVERSITY OF TOKYO

Microsoft Azure リソース、サービス、ドキュメントの検索 (G+/)

Microsoft.MachineLearningServices | 概要

デプロイ

検索 削除 キャンセル 再デプロイ ダウンロード 最新の情報に更新

概要

✓ デプロイが完了しました

デプロイ名 : Microsoft.MachineLearningServices
サブスクリプション : 1923306880-YamakataLab
リソース グループ : AI-Foundry-hub0

開始日時 : 2025/2/8 23:22:50
関連付け ID : 3df3a6f4-cacf-4654-86c7-36b50a2049c2

デプロイの詳細

次の手順

リソースに移動

Cost Management
予算内に収め、予期しない料金が請求されないように通知を受け取ります。
コストのアラートを設定 >

Microsoft Defender for Cloud
アプリおよびインフラストラクチャをセキュリティで保護します
次に移動: Microsoft Defender for Cloud >

無料の Microsoft チュートリアル
今すぐ学習を開始する >

エキスパートとの共同作業
Azure のエキスパートは、Azure での資産の管理を支援し、サポートの最前線となるサービス プロバイダー パートナーです。
Azure Expert の検索 >

学校 AI-Foundry-test0 - Microsoft Azure サービス azure foundry sku - 検索 Models available in Azure AI mod +

https://portal.azure.com/#@univtokyo.onmicrosoft.com/resource/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourcegroups/AI-Foundry-hub0/providers/Microsoft....

Microsoft Azure リソース、サービス、ドキュメントの検索 (G+/) Copilot 7 🚧 🚧 🚧 🚧 1923306880@utac.u... THE UNIVERSITY OF TOKYO

ホーム > Microsoft.MachineLearningServices | 概要 >

AI-Foundry-test0

Azure AI hub

検索 Create project Download config.json Delete

概要

要点

JSON ビュー

アクティビティ ログ	リソース グループ : AI-Foundry-hub0	Project resource group (... : AI-Foundry-hub0
アクセス制御 (IAM)	場所 : East US 2	Storage : aifoundrytest02974364556
タグ	サブスクリプション : 1923306880-YamakataLab	Container Registry (edit) : ...
問題の診断と解決	サブスクリプション ID : a3b0857d-8c96-43a1-b2ee-d787b83bbe5e	Application Insights (edit) : ...
イベント	Key Vault : aifoundrytest02322244864	Provisioning State : Succeeded

Govern the environment for your team in AI Foundry

Your Azure AI hub provides enterprise-grade security, and a collaborative environment to build AI solutions. Centrally audit usage and cost, and set up connections to your company resources that all projects can use. [learn more about the Azure AI Foundry](#)

Launch Azure AI Foundry

学校 | AI-Foundry-test0 - Microsoft Azure | ハブの概要 - Azure AI Foundry | azure foundry sku - 検索 | Models available in Azure AI mod | +

https://ai.azure.com/managementCenter/hub/overview?tid=f07baf4f-2b70-47d7-9a02-0875caf94c84&wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourcegroups/AI... A ⚡ ⚡ ⚡ ⚡ ⚡

Azure AI Foundry | 管理センター / ハブの概要

AI Foundry test0

+ 新しいプロジェクト (最新の情報に更新 ...)

説明
ハブの説明をここに追加してください

ハブのプロパティ

名前	場所
AI-Foundry-test0	eastus2
サブスクリプション	リソース グループ
1923306880-YamakataLab	AI-Foundry-hub0

Azure Portal で管理する

API エンドポイントとキーを取得する

サブスクリプション クオータを表示する

ハブの削除

ユーザー 1

AI-Foundry-test0 Azure AI Administrator

+ 新しいユーザー

すべてのハブとプロジェクト

ハブ
AI-Foundry-test0

Overview

All hubs + projects

Quota

ハブ (AI-Foundry-test0) Overview

Users

Models + endpoints

Connected resources

Compute

最初のプロジェクトを作成する

プロジェクトを使用すると、AI 資産を整理し、さまざまなツール、データ接続などにまたがる "状態" (システムが以前のイベントや相互作用を記憶している状態) を保存できます。

Azure AI Foundry 内のプロジェクトの詳細を確認する

AI Foundry test0

説明

ハブの説明をここに追加してください

ハブのプロパティ

名前: AI-Foundry-test0 場所: eastus2

サブスクリプション: 1923306880-YamakataLab リソース グループ: AI-Foundry-hub0

Azure Portal で管理する

API エンドポイントとキーを取得する

サブスクリプション クオータを表示する

ハブの削除

ユーザー 1

AI-Foundry-test0 Azure AI Administrator

+ 新しいユーザー

すべて表示 →

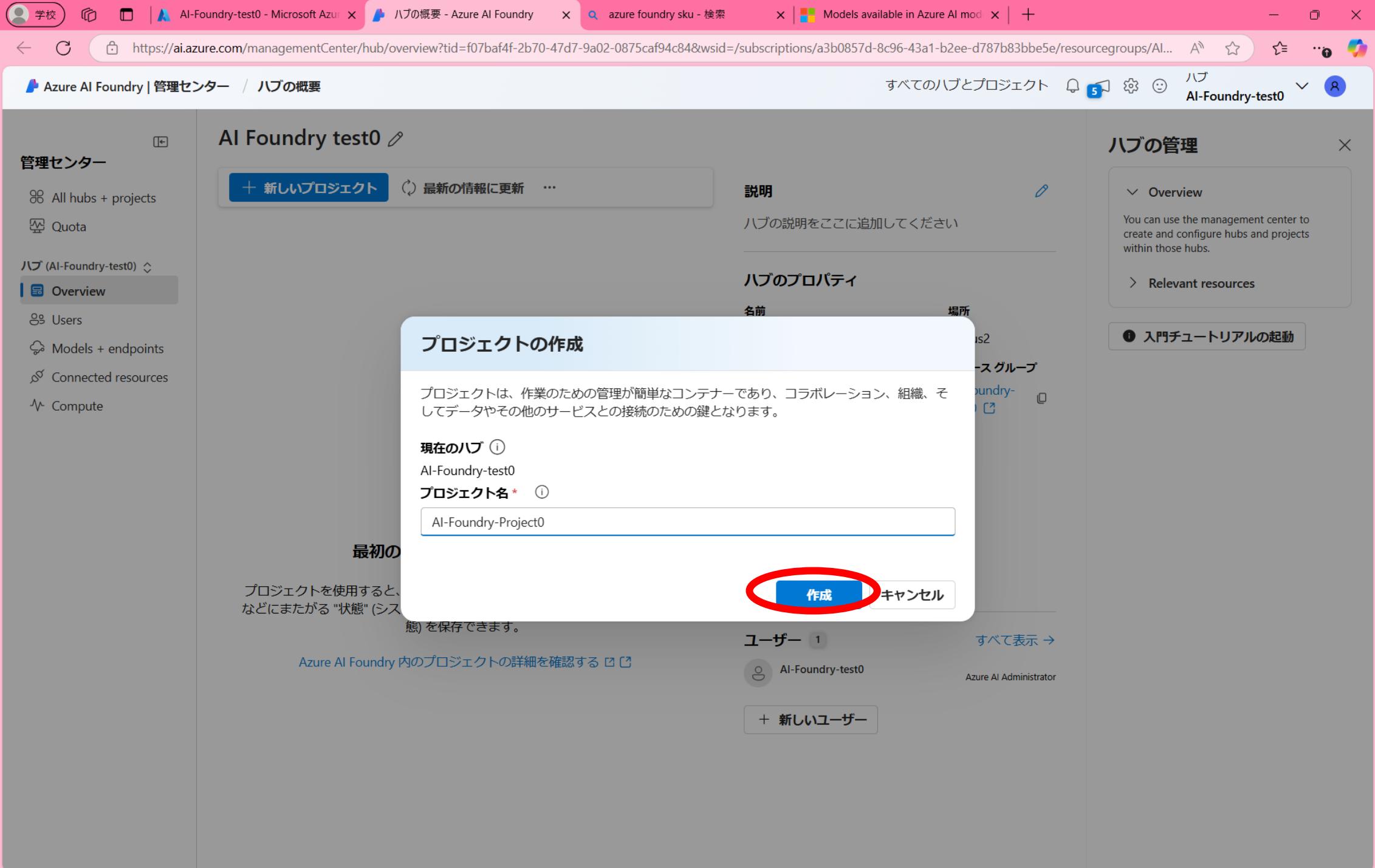
ハブの管理

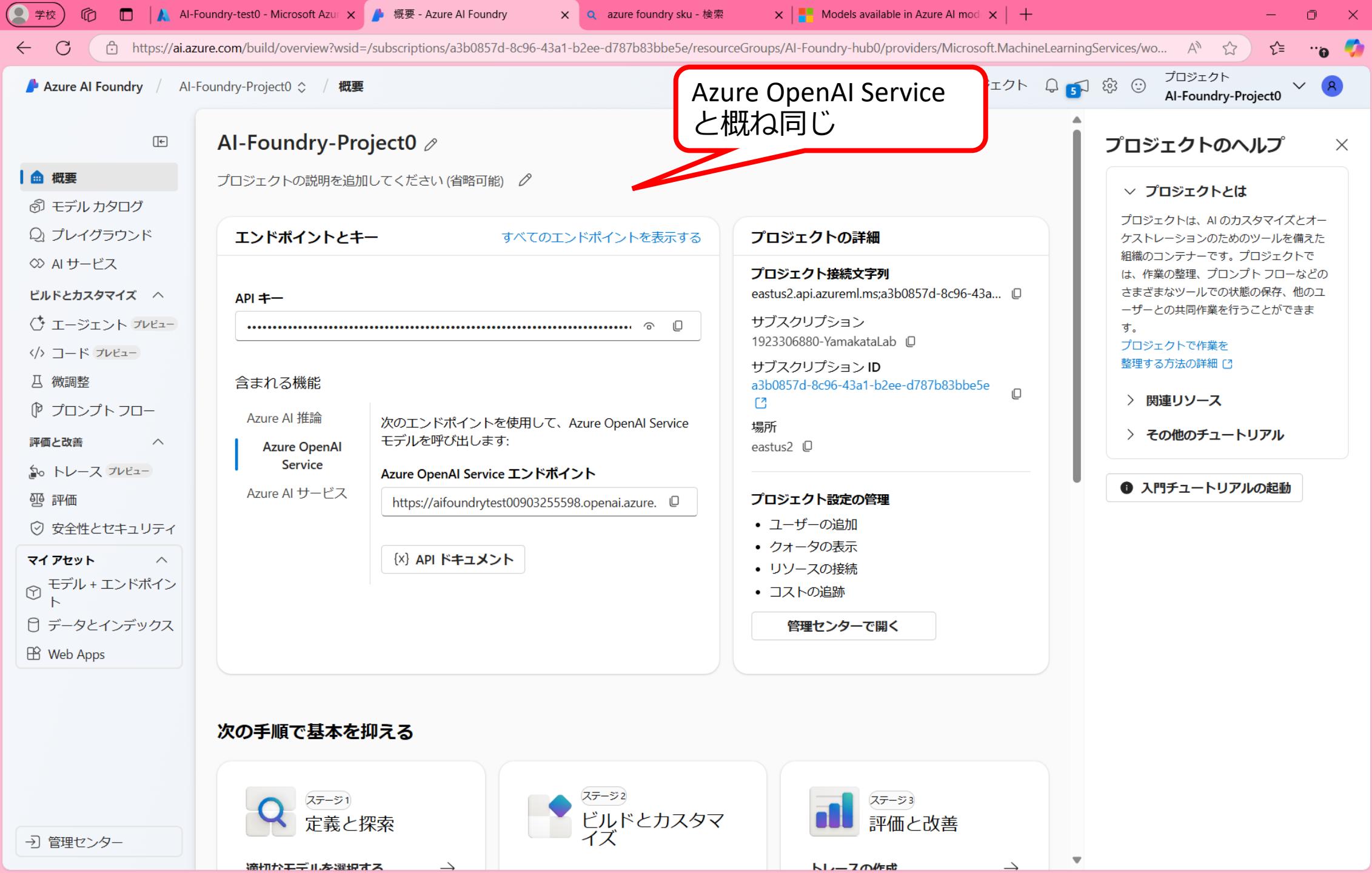
Overview

You can use the management center to create and configure hubs and projects within those hubs.

Relevant resources

入門チュートリアルの起動





2. モデルレカタログから
モデルを選択

学校 | AI-Foundry-test0 - Microsoft Azure | 概要 - Azure AI Foundry | azure foundry sku - 検索 | Models available in Azure AI mod | +

← ⌘ ⌘ https://ai.azure.com/build/overview?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/AI-Foundry-hub0/providers/Microsoft.MachineLearningServices/wo... A☆ ⌈ ⌋ ⌈ ⌋ ⌈ ⌋

Azure AI Foundry / AI-Foundry-Project0 / 概要 すべてのハブとプロジェクト 5 プロジェクト AI-Foundry-Project0 8

AI-Foundry-Project0

プロジェクトの説明を追加してください (省略可能)

モデルカタログ (選択) プレイグラウンド AI サービス ビルドとカスタマイズ エージェント プレビュー コード プレビュー 微調整 プロンプト フロー 評価と改善 トレース プレビュー 評価 安全性とセキュリティ マイアセット モデル + エンドポイント データとインデックス Web Apps

エンドポイントとキー

すべてのエンドポイントを表示する

API キー
.....

含まれる機能

Azure AI 推論 Azure OpenAI Service Azure AI サービス

次のエンドポイントを使用して、Azure OpenAI Service モデルを呼び出します:

Azure OpenAI Service エンドポイント
<https://aifoundrytest00903255598.openai.azure.com>

{x} API ドキュメント

プロジェクトの詳細

プロジェクト接続文字列
eastus2.api.azureml.ms;a3b0857d-8c96-43a1-b2ee-d787b83bbe5e

サブスクリプション
1923306880-YamakataLab

サブスクリプション ID
a3b0857d-8c96-43a1-b2ee-d787b83bbe5e

場所
eastus2

プロジェクト設定の管理

- ユーザーの追加
- クォータの表示
- リソースの接続
- コストの追跡

管理センターで開く

プロジェクトのヘルプ

▼ プロジェクトとは
プロジェクトは、AI のカスタマイズとオーケストレーションのためのツールを備えた組織のコンテナーです。プロジェクトでは、作業の整理、プロンプト フローなどのさまざまなツールでの状態の保存、他のユーザーとの共同作業を行うことができます。

▼ プロジェクトで作業を 整理する方法の詳細

▶ 関連リソース

▶ その他のチュートリアル

入门チュートリアルの起動

次の手順で基本を抑える



ステージ1
定義と探索



ステージ2
ビルトとカスタマイズ



ステージ3
評価と改善

→ 管理センター

→ 深めかエディルを選択オプス

トランクの作成 →

学校 | AI-Foundry-test0 - Microsoft Azure | モデル カタログ - Azure AI Foundry | azure foundry sku - 検索 | Models available in Azure AI mod | +

https://ai.azure.com/explore/models?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5/resourceGroups/AI-Foundry-hub0/providers/Microsoft.MachineLearningServices/w...

Azure AI Foundry / AI-Foundry-Project0 / モデル カタログ

すべてのハブとプロジェクト 5 プロジェクト AI-Foundry-Project0

カスタム AI ソリューションを構築するための適切なモデルを見つける

新着情報

4o Mini Audio is here!

Advanced audio at a fraction of the cost

モデルのチェックアウト

Check out models ブログを読む

2411/NeMo/3B で使用できるミストラル 微調整

カスタマイズにより、Mistral AI モデルのパフォーマンスを向上させます。

モデルのチェックアウト

Check out models ブログを読む

o3-mini はここにあります!

o3-mini には、高いパフォーマンスを必要とするシナリオに適したコスト効率の高い o1 機能が...

モデルのチェックアウト

Check out models ブログを読む

New model benchmarks available now in model catalog

Model benchmarks are integrated into model catalog for easier navigation. Compare benchmarks across models and datasets available in the industry to assess which one...

Compare with benchmarks How model benchmarks are scored

コレクション 業界 機能 デプロイ オプション 推論タスク

タスクの微調整 ライセンス

モデルの比較

検索

モデル 1831

o3-mini チャットの完了

DeepSeek-R1 チャットの完了

Phi-4 チャットの完了

gpt-4o-mini-audio-preview オーディオの生成

gpt-4o-mini-realtime-preview オーディオの生成

o1 チャットの完了

クリップボード (24/24)
アイテムは収集されません。クリップボードの容量以上のサイズです。

概要

モデル カタログ

プレイグラウンド

AI サービス

ビルトとカスタマイズ

エージェント プレビュー

コード プレビュー

微調整

プロンプト フロー

評価と改善

トレース プレビュー

評価

安全性とセキュリティ

マイアセット

モデル + エンドポイント

データとインデックス

Web Apps

管理センター

モデルのヘルプ

How to use the model catalog

Search by name, filter, or browse to find the right model for your use case. Click to see model details and specifications from the publisher and to deploy a model.

Relevant resources

Microsoft 以外のモデルには異なる使用条件が適用されることに注意してください。デプロイする前にドキュメントをよくお読みください。

入门チュートリアルの起動



← DeepSeek-R1 プレビュー

▷ デプロイ[詳細](#) [既存のデプロイ](#) [コードサンプル](#) [ライセンス](#)*Learn more: [original model announcement]*

DeepSeek-R1 excels at reasoning tasks using a step-by-step training process, such as language, scientific reasoning, and coding tasks. It features 671B total parameters with 37B active parameters, and 128k context length.

DeepSeek-R1 builds on the progress of earlier reasoning-focused models that improved performance by extending Chain-of-Thought (CoT) reasoning. DeepSeek-R1 takes things further by combining reinforcement learning (RL) with fine-tuning on carefully chosen datasets. It evolved from an earlier version, DeepSeek-R1-Zero, which relied solely on RL and showed strong reasoning skills but had issues like hard-to-read outputs and [\[original model announcement\]](#).

[詳細表示](#)

データ、メディア、言語

Property	説明	
サポートされるデータ型	入力 text	出力 text
サポートされている言語	en, zh	

使用状況のガイダンス

Content Filtering

Prompts and completions are passed through a default configuration of Azure AI Content Safety classification models to detect and prevent the output of harmful content. Learn more about [Azure AI Content Safety](#). Configuration options for content filtering vary when you deploy a model for production in

[管理センター](#)

クイックファクト

DeepSeek-R1

[chat-completion](#)

最後のトレーニング
使用できません

モデル ID

コードでモデルを配置するときにこのモードルを参照する

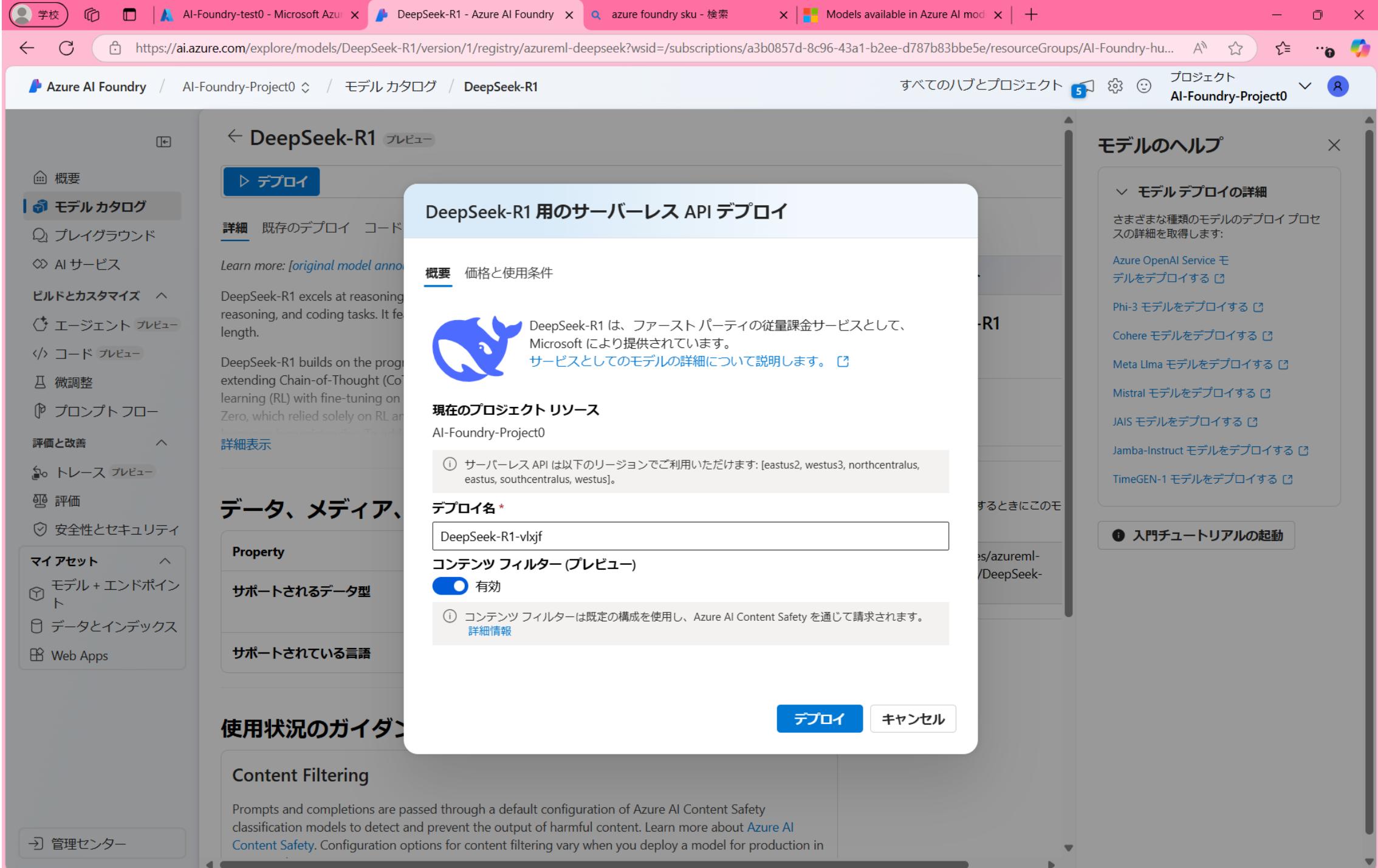
azureml://registries/azureml-deepseek/models/DeepSeek-R1/versions/1

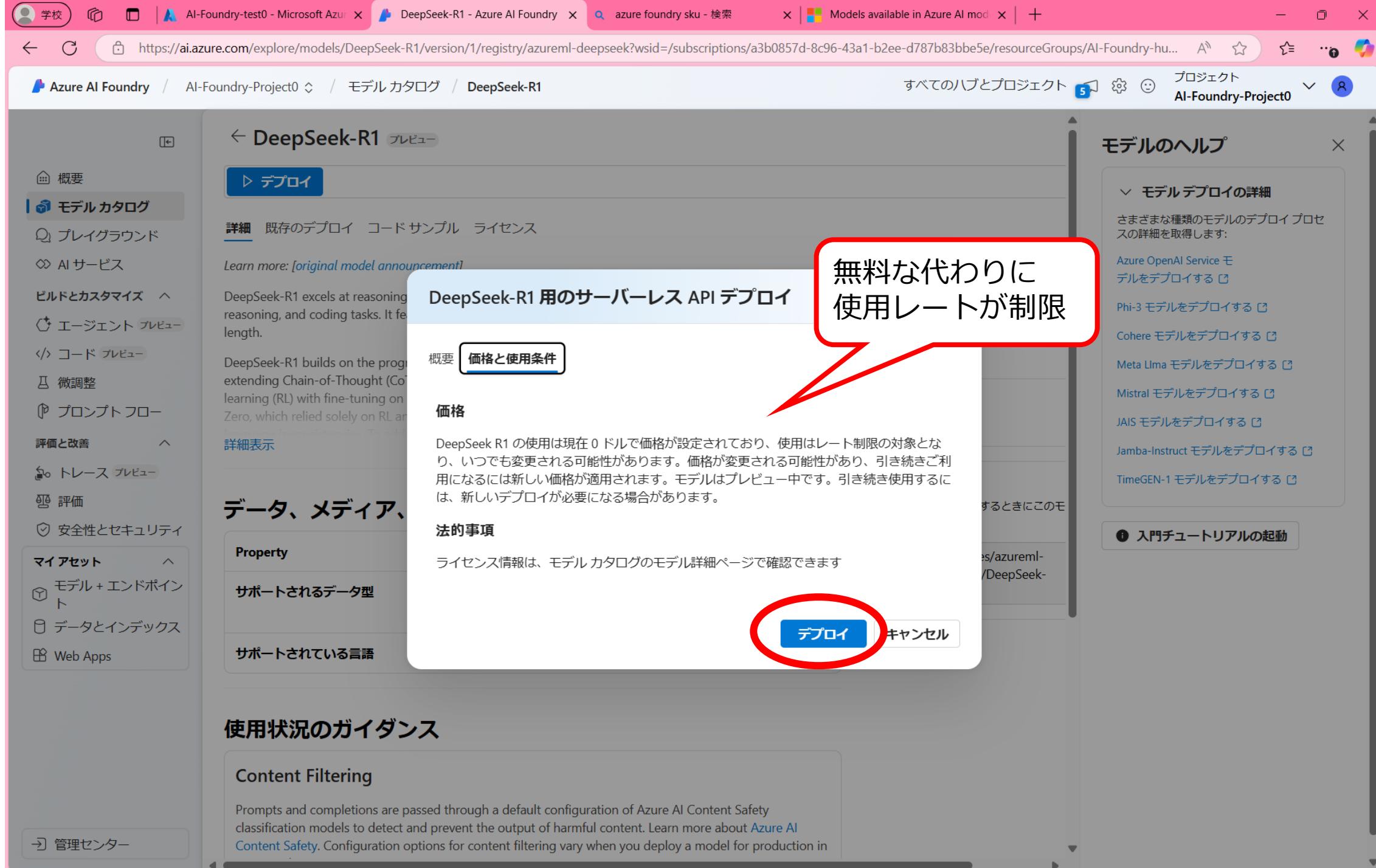
モデルのヘルプ

▼ モデルデプロイの詳細

さまざまな種類のモデルのデプロイプロセスの詳細を取得します:

[Azure OpenAI Service モデルをデプロイする](#)[Phi-3 モデルをデプロイする](#)[Cohere モデルをデプロイする](#)[Meta Llma モデルをデプロイする](#)[Mistral モデルをデプロイする](#)[JAIS モデルをデプロイする](#)[Jamba-Instruct モデルをデプロイする](#)[TimeGEN-1 モデルをデプロイする](#)[① 入門チュートリアルの起動](#)





学校 | AI-Foundry-test0 - Microsoft Azure | DeepSeek-R1-vlxjf - Azure AI Foundry | azure foundry sku - 検索 | Models available in Azure AI mod | +

https://ai.azure.com/build/deployments/serverless/DeepSeek-R1-vlxjf?wsid=/subscriptions/a3b0857d-8c96-43a1-b2ee-d787b83bbe5e/resourceGroups/AI-Foundry-hub0/providers/Micro... A ⭐ ⌂ ⌂ ⌂

Azure AI Foundry / AI-Foundry-Project0 / モデル + エンドポイント / DeepSeek-R1-vlxjf すべてのハブとプロジェクト 5 🚧 🛡️ 🎉 プロジェクト AI-Foundry-Project0 8

DeepSeek-R1-vlxjf

概要 実行する

プレイグラウンドで開く 最新の情報に更新 編集 削除

配置情報

名前: DeepSeek-R1-vlxjf
プロビジョニングの状態: 成功
最終更新日時: Feb 8, 2025 11:27 PM
作成者: 1923306880@utac.u-tokyo.ac.jp
作成日: Feb 8, 2025 11:27 PM

モデル: DeepSeek-R1

エンドポイント

ターゲット URI: <https://DeepSeek-R1-vlxjf.eastus2.models.ai.azure.com>
キー: [.....](#)

コンピューティングの種類: Consumption
Swagger URI: <https://DeepSeek-R1-vlxjf.eastus2.models.ai.azure.com/swagger.json>

API ルート

Azure AI model inference: Chat Completion
<https://DeepSeek-R1-vlxjf.eastus2.models.ai.azure.com/chat/completions>

deepseek-r1: Chat Completion
<https://DeepSeek-R1-vlxjf.eastus2.models.ai.azure.com/v1/chat/completions>

監視と安全性

Azure AI Content Safety: 有効

マイアセット マーク モデル + エンドポイント データとインデックス Web Apps

管理センター

3. チャットプレイグラウンド
からチャットを実行

