

サポート用Webページ プリザンターを利用したRAGの構築

プリザンターを利用したRAGの構築で設定するAzureサービスの内容は以下を参照してください。

1. リソースグループの主な設定

Azureのポータルにログインし新しいリソースグループを1つ作成します。

・表I:リソースグループの主な設定

内容	設定値
サブスクリプション	使用するサブスクリプション
リソース グループ名	リソースグループに付ける名前
リージョン	(Asia Pacific) Japan East

[!CAUTION]

チャットシステムの構築ではAzureのサービスを利用します。本記事の執筆時点では、Azureの新規利用時に30日間利用可能な無料クレジットが提供されるサービスがあり、筆者の検証では無料クレジットの範囲内で構築することができました。ただし、リソースの使い方や課金体系の仕様変更によって課金が発生する可能性がありますので、最新のAzureの仕様をよくご確認の上で構築してください。

<https://azure.microsoft.com/ja-jp/pricing/purchase-options/azure-account>

2. Azure OpenAIの主な設定

リソースグループにAzure OpenAIのリソースを1つ作成します。

・表II:Azure OpenAIの主な設定

内容	設定値
リージョン	Japan East
名前	Azure OpenAIのリソースに付ける名前 (例:open-ai-0227)
価格レベル:	Standard S0
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可
Tags	(設定不要)

[!CAUTION]

契約内容によりますがAzure OpenAIのリソース作成上限が1つという場合があります。1つも作成できない場合は、Azure AI servicesの削除されたリソースの管理にAzure OpenAIが残っていないかを確認してください。削除されたリソースから消去することで新たにAzure OpenAIを作成できます。

3. モデルのデプロイ

Azure OpenAIのリソースに共有リソースとしてモデルを2種類デプロイします。共有リソースのデプロイ手順は手順を参照してください。

・手順

[①作成したAzure OpenAIの管理画面に移動]→[②Go to Azure AI Foundry portalボタン]→[③Azure AI Foundryのポータルに移動]→[④共有リソース]→[⑤デプロイ]→[⑥+モデルのデプロイ]→[⑦基本モデルをデプロイする]

・表III:gpt-4o-miniのデプロイ設定

内容	設定値
モデル	gpt-4o-mini
デプロイ名	gpt-4o-mini
デプロイの種類	グローバル標準
モデルバージョン	2024-07-18
AI リソース	Azure OpenAIの作成で作成したリソース
容量	250K - 1分あたりのトークン数
リソースの場所	Japan East
備考	Japan East (クォータなし)が表示される場合は注意事項を要確認

・表IV:text-embedding-ada-002のデプロイ設定

内容	設定値
モデル	text-embedding-ada-002
デプロイ名	text-embedding-ada-002
デプロイの種類	Standard
モデルバージョン	2(既定)
AI リソース	Azure OpenAIの作成で作成したリソース
容量	120K - 1分あたりのトークン数
リソースの場所	Japan East

[!CAUTION]
契約内容によってgpt-4o-miniのリソースの場所にJapan East (クォータなし)と表示される可能性があります。この場合は**手順**を実施してクォータを要求してください。リソースの場所にJapan Eastを選択できるようになるまで1〜2日程度かかります。また、リソースの場所を変えてしまうとAzure OpenAIのリソースと紐づかなくなりますので、リソースの場所はAzure OpenAIと揃えてください。

・手順

[①Azure AI Foundryのポータルに移動]→[②共有リソース]→[③クォータ]→[④クォータの要求]

[!CAUTION]
クォータの要求の入力フォームは全文英語で記載されていますが、日本語で内容を入力しても問題ありません。内容を正確に記述するように心がけてください。

4. Azure AI Searchの主な設定

リソースグループにAzure AI Searchのリソースを1つ作成します。

・表V:Azure AI Searchの主な設定

内容	設定値
サービス名	Azure AI Searchのリソースに付ける名前 (例:ramen-search-0227)
リージョン	(Asia Pacific) Japan East
価格レベル:	Free
スケーリング	(設定不要)
タグ	(設定不要)

5. インデックスの作成

Azure AI Searchに格納するインデックスを定義します。インデックスを作成する手順は手順を参照してください。

・手順

[①作成したSearch Serviceの管理画面に移動]→[②検索管理]→[③インデックス]→[④+インデックスの追加]→[インデックスの追加]

インデックス名を入力してください。(例:ramen-20250301)

次に、インデックスフィールドを8つ追加します。デフォルトで表示されているidフィールドは削除してください。インデックスフィールドは「+フィールドの追加」ボタンで追加します。表VIを参照してフィールドの内容を入力してください。

・表VI:追加するインデックスフィールド

フィールド名	種類	属性の構成
ID	Edm String	キー、取得可能、並べ替え可能
StoreName	Edm String	取得可能
Reviews	Edm String	取得可能
Location	Edm String	取得可能
Style	Edm String	取得可能
RecommendedMenu	Edm String	取得可能
Keyword	Edm String	取得可能

次にTextVectorのフィールドを追加しますが、少し手順が他のフィールドと異なります。まず表VIIの通りにフィールドの内容を入力します。

・表VII:追加するインデックスフィールド(TextVector)

フィールド名	種類	属性の構成	ディメンション
TextVector	Collection(Edm Single)	ストレージに含める、検索可能	1536

TextVectorにはベクター検索プロファイルを構成する「アルゴリズム」と「ベクトル化」を設定する必要があります。

初期状態でベクター検索プロファイルがありませんのメッセージの下に「作成」ボタンをクリックすると、画面の状態が変わります。アルゴリズムの構成がありませんとベクトル化なしの2つのメッセージの各配下に「作成」ボタンが表示されていることを確認してください。

アルゴリズムとベクトル化は各「作成」ボタンをクリックして表VIIIおよび表IXの通りに設定してください。

・ 表VIII:アルゴリズム(ベクトルアルゴリズム)

内容	設定値
アルゴリズム名	vector-config-xxxxxx(自動付与される)
Kind	hnsw
双方向リンク数(m)	4
efConstruction	400
efSearch	500
類似性メトリック	cosine

・ 表IX:ベクトル化

内容	設定値
名前	vectorizer-xxxxxx(自動付与される)
Kind	Azure OpenAI
サブスクリプション	使用するサブスクリプション
Azure OpenAI Service	Azure OpenAIの作成で作成したリソース
モデル デプロイ	text-embedding-ada-002
認証の種類	APIキー

ここまで設定できたらインデックスの作成の「作成」ボタンをクリックします。