

Azure AI Search ドキュメント

従来または生成検索シナリオでのベクターおよびテキスト コンテンツの大規模な情報取得。

Azure AI Search について

☰ 概念

[Azure AI Search とは?](#)

[AI エンリッチメント](#)

[セマンティックの再ランク付け](#)

⚡ クイックスタート

[Search Service の作成](#)

[検索インデックスを作成する](#)

ベクトルのサポート

☰ 概念

[Azure AI Search のベクター](#)

[組み込みベクター化 \(プレビュー\)](#)

[組み込みスカラー量子化 \(プレビュー\)](#)

[取得拡張生成 \(RAG\)](#)

⚡ クイックスタート

[ベクトル インデックスを作成する](#)

[データに基づいたグラフ](#)

[ベクトルインデックスにクエリを実行する](#)

📁 SAMPLE

ベクトルのサンプル

Azure AI Studio

攻略ガイド

[AI Studio でベクター インデックスを作成する](#)

[質問と回答のコパイロットをビルドする](#)

データのインデックス作成

概念

[検索インデックスとは](#)

[データのインポート](#)

[インデクサーの概要](#)

攻略ガイド

[Azure Blob Storage のインデックス](#)

[Azure SQL Database のインデックス](#)

[Azure Cosmos DB のインデックス](#)

[任意のデータのインデックス作成](#)

アプリの開発

チュートリアル

[Web アプリに検索を追加する](#)

攻略ガイド

[.NET での開発](#)

[REST での開発](#)

リファレンス

[Azure REST API リファレンス](#)

[Azure SDK for .NET](#)

[Azure SDK for Python](#)

[Azure SDK for Java](#)

[Azure SDK for JavaScript](#)

クエリ データ

概念

[クエリの種類と構成](#)

[簡単なクエリを作成する](#)

[詳細なクエリを作成する](#)

リファレンス

[単純な構文 \(既定\)](#)

[OData 言語のリファレンス](#)

[ドキュメントを検索する \(REST\)](#)

Azure AI Search とは

[アーティクル] • 2024/05/03

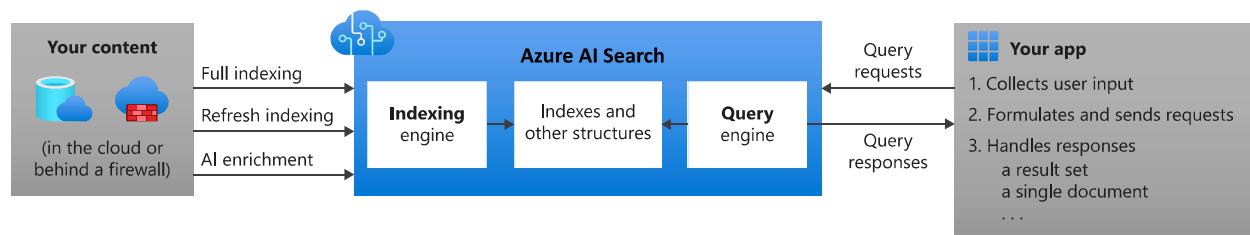
Azure AI Search (旧称 "Azure Cognitive Search") は、従来の検索アプリケーションおよび生成型 AI 検索アプリケーションでのユーザー所有コンテンツに対する安全な情報取得を大規模に提供します。

情報取得は、テキストとベクトルを表示するすべてのアプリの基礎となります。一般的なシナリオには、カタログまたはドキュメントの検索、データ探索、一般的になりつつある独自のグラウンディング データを活用したチャットスタイルのアプリがあります。検索サービスを作成するときは、次の機能を使用します。

- ベクトル検索用の検索エンジンおよび検索インデックスに対するフルテキスト検索とハイブリッド検索
- 統合型データ チャンキングとベクター化 (レビュー)、テキストの字句解析、コンテンツの抽出と変換のためのオプションの AI エンリッチメントによる豊富なインデックス作成
- ベクター クエリ、テキスト検索、ハイブリッド クエリ、あいまい検索、オートコンプリート、地理空間検索などに対応した豊富なクエリ構文
- Azure のスケール、セキュリティ、リーチ
- データレイヤー、機械学習レイヤー、Azure AI サービス、Azure OpenAI での Azure 統合

Search Service の作成

アーキテクチャ上、検索サービスは、インデックスが設定されていないデータを含む外部データストアと、検索インデックスにクエリ要求を送信して応答を処理するクライアントアプリとの間に配置されます。



クライアントアプリでは、検索エクスペリエンスは Azure AI Search の API を使用して定義され、関連性のチューニング、セマンティック ランク付け、オートコンプリート、同意語の一致、あいまい一致、パターン マッチング、フィルター、並べ替えを含めることができます。

Azure プラットフォーム全体で、Azure AI Search は、他の Azure サービスと統合できます。その形式は、Azure データソースからのデータのインジェストと取得を自動化

するインデクサー、および Azure AI サービスで提供されている消費型 AI (画像や自然言語処理など) や Azure Machine Learning 内で作成したり Azure Functions 内にラップしたりするカスタム AI を組み込んだスキルセットです。

検索サービスの内部

検索サービス自体は、"インデックスの作成" と "クエリの実行" の 2 つが主なワークフローとなります。

- **インデックス作成**は、取り込みのプロセスです。コンテンツを検索サービスに読み込んで検索可能にします。内部では、受信テキストはトークンに処理され、反転インデックスに保存され、受信ベクトルはベクトルインデックスに保存されます。Azure AI Search でインデックスを付けられるドキュメント形式は JSON です。作成した JSON ドキュメントをアップロードすることや、インデクサーを使用してデータを取得して JSON にシリアル化することができます。

[スキルセット](#)による [Applied AI](#) により、画像モデルと言語モデルを使用したインデックス作成が拡張されます。ソースドキュメントに画像や大きな非構造化テキストがある場合は、OCR の実行、画像の説明、構造の推測、テキストの翻訳などを行うスキルをアタッチできます。 [データチャンкиングとベクトル化](#)を実行するスキルをアタッチすることもできます。

- **クエリの実行**は、検索可能なテキストがインデックスに取り込まれると、クライアントアプリが検索サービスにクエリ要求を送信して応答を処理するときに発生することがあります。すべてのクエリは、ユーザーが制御する検索インデックスを介して実行されます。

[セマンティックランク付け](#)は、クエリ実行の拡張機能です。検索結果の処理に言語理解が加わり、最も意味の関連が高い結果が上位に昇格されます。

Azure AI Search を使用する理由

Azure AI Search は、次のアプリケーションシナリオに適しています。

- 従来のフルテキスト検索や次世代型のベクトル類似性検索に使用できます。キーワード検索と類似性検索の強みを活用した情報取得で生成 AI アプリを支援します。両方のモダリティを使用して、最も関連性の高い結果を取得します。
- 異種コンテンツを、ベクトルとテキストで構成されるユーザーが定義および設定した検索インデックスに統合します。検索可能なコンテンツの所有権とコントロールを保持します。

- 生成 AI および RAG アプリ用にデータ チャンクとベクトル化を統合する。
- ドキュメント レベルで細分性の高いアクセス制御² を適用します。
- インデックス作成とクエリワークロードを専用の検索サービスにオフロードします。
- 関連性のチューニング、ファセットナビゲーション、フィルター (地理空間検索)、同意語マッピング、オートコンプリートなど、検索に関連した機能を容易に実装。
- Azure Blob Storage や Azure Cosmos DB に格納されている未区分の大きなテキスト ファイル、画像ファイル、またはアプリケーション ファイルを検索可能なチャンクに変換。これは、Azure AI から外部処理を追加するコグニティブ スキルを使ってインデックスを作成するときに行われます。
- 言語分析またはカスタム テキスト分析を追加。英語以外のコンテンツがある場合、Azure AI Search では Lucene アナライザーと Microsoft の自然言語プロセッサの両方がサポートされます。また、生コンテンツの特殊な処理 (分音記号のフィルター処理や、文字列内のパターンの認識と保持など) を実行するようにアナライザーを構成することもできます。

特定の機能の詳細については、「[Azure AI Search の機能](#)」を参照してください。

ファースト ステップ

機能は、Azure portal、シンプルな REST API、または Azure SDK for .NET などの Azure SDK を使用して公開されます。Azure portal は、インデックスとスキルセットのプロトタイプ作成やクエリを実行するためのツールによるサービスの管理とコンテンツの管理をサポートしています。

次の 4 つのステップにより、主要な検索機能をエンド ツー エンドで試すことができます。

1. [レベル](#)とリージョンを決定します。サブスクリプションごとに 1 つだけ無料 Search サービスが許可されます。すべてのクイックスタートは、Free レベルで完了できます。さらに容量と機能が必要な場合は、[有料のレベル](#)² が必要です。
2. [Azure portal ポータルで Search サービスを作成](#)します。
3. [\[データのインポート\] ウィザードから開始します](#)。組み込みのサンプルまたはサポートされているデータ ソースを選ぶと、インデックスの作成、ロード、クエリを数分で実行できます。

4. ポータル クライアントを使って作成した検索インデックスのクエリを実行して、[Search エクスプローラー](#)で終了します。

また、検索インデックスの作成、読み込み、クエリをアトミックな手順で行うこともできます。

1. ポータル、[REST API](#)、[.NET SDK](#)、または他の SDK を使用して[検索インデックスを作成](#)します。検索可能なコンテンツの構造は、インデックススキーマによって定義されます。
2. "プッシュ" モデルを使用して任意のソースから JSON ドキュメントをプッシュして[コンテンツをアップロード](#)します。または、ソースデータが[サポートされる型](#)である場合は "プル" モデル (インデクサー) を使用します。
3. ポータルの[検索エクスプローラー](#)、[REST API](#)、[.NET SDK](#)、または他の SDK を使用して、[インデックスを照会](#)します。

💡 ヒント

複雑なソリューションやカスタムソリューションについては、Azure AI Search テクノロジにおける深い専門知識を持つ[パートナーにお問い合わせください](#)。

検索オプションの比較

お客様から、Azure AI Search は他の検索関連のソリューションと比較してどうなのかについてよくお問い合わせいただきます。主な相違点を次の表に示します。

[+] テーブルを展開する

比較対象	主な相違点
Microsoft Search	Microsoft Search は、SharePoint 内のコンテンツに対してクエリを実行する必要がある、Microsoft 365 の認証済みユーザーを対象としています。 Azure AI Search は、Azure および任意の JSON データセット全体からコンテンツをプルします。
Bing	Bing API では、Bing.com のインデックスに対してクエリを実行して一致する語句を検索します。 Azure AI Search では、ユーザーのコンテンツが入力されたインデックスを検索します。データインジェストとスキーマを制御するのはユーザーです。
データベース検索	Azure SQL には、 フルテキスト検索 と ベクトル検索 があります。 Azure Cosmos DB にも、 テキスト検索 と ベクトル検索 があります。 Azure AI Search は、関連性のチューニングなどの機能、または異種ソースからのコンテンツが必要な場合に魅力的な選択肢になります。リソース使用率も、別途考慮する必要があります。インデックス作成とクエリは、計算負荷が高くなります。DBMS から検索をオフ

比較対象	主な相違点
	ロードすることにより、トランザクション処理用のシステム リソースが保持されます。
専用の検索ソリューション	全機能を持つ専用の検索を使用することにした場合、最終的なカテゴリ比較は検索テクノロジ間で行われます。情報取得とコンテンツ ナビゲーションの両方を主に検索に依存するアプリを使用する場合、Azure AI Search は、Azure 上のコンテンツに対するベクター、キーワード、ハイブリッド ワークロードの点でクラウド プロバイダーの中でも最強です。

以下のような大きな強みがあります。

- ベクトルおよび非ベクトル (テキスト) のインデックス作成とクエリのサポート。ベクトル類似性検索を使用すると、検索語句が完全に一致しない場合でも、検索 クエリと意味的に類似した情報を見つけることができます。キーワード検索とベクトル検索の長所を最大限に生かすには、ハイブリッド検索を使用します。
- セマンティック ランク付けとスコアリング プロファイルによるランク付けと関連性のチューニング。クエリ構文では、用語ブーストとフィールドの優先順位付けがサポートされます。
- インデックス層での Azure データ統合 (クローラー)。
- コンテンツのテキストとベクターを検索可能にする変換のための Azure AI 統合。
- 信頼された接続用の Microsoft Entra セキュリティと、インターネットなしのシナリオでのプライベート接続用の Azure Private Link。
- [完全な検索エクスペリエンス](#): 56 言語での言語分析およびカスタム テキスト分析。ファセット、オートコンプリート クエリ、候補の結果、シノニム。
- Azure のスケール、信頼性、グローバルなリーチ。

Azure AI Search の新機能

[アーティクル] • 2024/05/03

Azure Cognitive Search は Azure AI Search になりました。 Azure AI Search の機能、ドキュメント、サンプルの最新情報について説明します。

① 注意

プレビュー機能はここで発表されていますが、[プレビュー機能の一覧](#)も維持管理されているため、1か所で参照できます。

2024 年 4 月

[+] テーブルを展開する

Item	型	説明
情報漏えいに対処するセキュリティ更新プログラム	API	GET 応答では接続文字列もキーも返されなくなりました。 GET スキルセット、GET インデックス、および GET インデクサーに適用されます。この変更は、AI Search と統合された Azure 資産を未承認のアクセスから保護するのに役立ちます。
Basic レベルと Standard レベルでのストレージの拡張	機能	Basic では、最大 3 つのパーティションと 3 つのレプリカがサポートされるようになりました。Basic レベルおよび Standard (S1, S2, S3) レベルでは、パーティションごとの課金レートは同じで、パーティションあたりのストレージが大幅に増えます。追加される容量はリージョンの可用性の対象となり、2024 年 4 月 3 日以降に作成された新しい検索サービスに適用されます。現在、インプレース アップグレードはありません。そのため、新しい検索サービスを作成して追加のストレージを取得してください。
ベクターのクオータの増加	機能	ベクター クオータも同様に、選択したリージョンで 2024 年 4 月 3 日以降に作成された新しいサービスで増加します。
組み込みのベクター量子化、narrow ベクターデータ型、および新しい stored プロパティ (プレビュー)	機能	このレビューでは、3 つの機能強化により、より大きなベクターワークロードのサポートが、より低いコストで追加されます。1 番目としては、スカラー量子化により、メモリ内とディスク上のベクター インデックスサイズが圧縮されます。2 番目としては、使用できるベクター フィールドについて、narrow データ型を割り当てることができます。3 番目として、より柔軟なベクター フィールド ストレージ オプションを追加しました。

Item	Type	Description
2024-03-01-preview Search REST API	API	新しいデータ型、ベクター圧縮プロパティ、およびストレージオプションのための Search REST API の新しいレビュー バージョン。
2024-03-01-preview Management REST API	API	コントロールプレーン操作用の新しいレビュー バージョンの Management REST API。
2023-07-01-preview 非推奨のお知らせ	API	2024 年 4 月 8 日 非推奨のお知らせ。2024 年 7 月 8 日に提供終了。これは、ベクトル検索サポートを提供した最初の REST API でした。新しい API バージョンでは、ベクトル構成が異なります。できるだけ早く 新しいバージョンに移行 することをお勧めします。

2024 年 2 月

[+] [テーブルを開く](#)

Item	Type	Description
ディメンションの新しい制限	機能	ベクトルフィールドのディメンションの上限が <code>2048</code> から引き上げられて <code>3072</code> になりました。

2023 年のお知らせ

[+] [テーブルを開く](#)

Month	Type	お知らせ
11 月	機能	ベクトル検索、一般提供 。カスタマー マネージド キー (CMK) に対する以前の制限は現在、解除されています。 事前フィルター処理 と、網羅的な K 近傍法も一般提供になりました。
11 月	機能	セマンティック ランク付け、一般提供
11 月	機能	垂直統合 (レビュー) は、インデックス作成中のデータのチャンクへの変換とテキストからベクトルへの変換を追加し、クエリ実行時のテキストからベクトルへの変換も追加します。
11 月	機能	[データのインポートとベクトル化] ウイザード (レビュー) は、データ チャンクとベクトル化を自動化します。対象は、 2023-10-01-Preview REST API です。

Month	Type	お知らせ
11月	機能	インデックスのプロジェクト（レビュー） は、エンリッチメントパイプラインのコンテンツが複数のインデックスをターゲットにできる1対多のインデックスパターンに使用されるセカンダリインデックスの形状を定義します。
11月	API	2023-11-01 Search REST API は、ベクトル検索およびセマンティックランク付け用の安定バージョンのSearch REST APIです。一般公開された機能への移行手順については、 REST API のアップグレードに関するページ を参照してください。
11月	API	2023-11-01 Management REST API では、セマンティックランク付けを有効または無効にするAPIが追加されています。
11月	スキル	Azure OpenAI埋め込みスキル（レビュー） は、Azure OpenAIリソースにデプロイされている埋め込みモデルに接続して、スキルセットの実行中に埋め込みを生成します。
11月	スキル	テキスト分割スキル（レビュー） は、ネイティブデータのチャunk分割をサポートするために、 2023-10-01-Preview で更新されています。
11月	ビデオ	ベクトル検索とセマンティックランク付けによってGPTプロンプトを向上させる方法 は、ハイブリッド検索によって、有用なAI応答を生成するための最適な基礎データ入手して、概念とキーワードの両方で検索できるようにする方法について説明します。
11月	サンプル	生成AIアプリケーションでのロールベースのアクセス制御 は、Microsoft Entra IDとMicrosoft Graph APIを使用して、インデックス内のチャunk分割されたコンテンツに対する詳細なユーザーアクセス許可をロールアウトする方法について説明します。
10月	サンプル	"データとチャット"ソリューションアクセラレータ 。 Azure AI Searchをトリガーとして使用するエンドツーエンドのRAGパターン。インデックス作成、データチャunk、オーケストレーションを提供します。
10月	機能	網羅的Kニアレストネイバー(KNN) は、ベクトル空間での類似性検索のためのスコアリングアルゴリズムです。2023-10-01-Preview REST APIでのみ使用できます。
10月	機能	ベクトル検索でのプリフィルター は、クエリの実行前にフィルター条件を評価し、検索する必要があるコンテンツの量を減らします。2023-10-01-Preview REST APIでのみ使用できます。要件に応じて、 <code>vectorFilterMode</code> (既定値) または <code>preFilter</code> に設定できるクエリに対して新しい <code>postFilter</code> プロパティを使います。
10月	API	2023-10-01-Preview Search REST API は、ベクトルフィールドまたはベクトルクエリの定義の破壊的変更です。
8月	機能	拡張セマンティックランク付け 。アップグレードされたモデルが、セマンティック再ランク付けのためにロールアウトされ、より多くのリージョンが利用

Month	Type	お知らせ
		できるように拡張されています。一意のトークンの最大数が2倍になり128から256に増えました。
7月	サンプル	ベクター デモ (Azure SDK for JavaScript) 。埋め込みの生成、インデックスの作成と読み込み、およびいくつかのベクタークエリの実行のために、Node.js と @azure/search-documents 12.0.0-beta.2 ライブラリを使用します。
7月	サンプル	ベクター デモ (Azure SDK for .NET) 。埋め込みの生成、インデックスの作成と読み込み、およびいくつかのベクタークエリの実行のために、Node.js と Azure.Search.Documents 11.5.0-beta.3 ライブラリを使用します。このサンプル は 、Azure SDK チームから試すこともできます。
7月	サンプル	ベクター デモ (Azure SDK for Python) は、azure.search.documents の最新ベータリリースを使用して、埋め込みの生成やインデックスの作成と読み込みに加え、複数のベクトルクエリの実行を行います。その他のベクトル検索デモについては、 azure-search-vector-samples/demo-python リポジトリを参照してください。
6月	機能	ベクトル検索のパブリック プレビュー。
6月	機能	セマンティック検索の利用可能性 、Basic レベルで使用できます。
6月	API	2023-07-01-Preview Search REST API 。ベクトル検索のサポート。
5月	機能	Azure RBAC (ロールベースのアクセス制御、一般提供)。
5月	API	Azure ロールを使用するように検索を構成するためのサポートを備えた 2022-09-01 Management REST API 。検索サービスの認証オプションをサポートするよう、Azure PowerShell の Az.Search モジュールと Azure CLI の Az search モジュールが更新されました。 Terraform プロバイダー を使用して認証オプションを構成することもできます(詳細については、こちらの Terraform のクリックスタート を参照してください)。
4月	サンプル	事業継続とディザスター リカバリーのための Azure AI Search のマルチリージョン展開 。エンドポイントが失敗した場合にコンテンツの同期と要求のリダイレクトを行うためのオプションを含む、Azure AI Search のマルチリージョンソリューションを完全に構成するデプロイ スクリプト。
3月	サンプル	Azure OpenAI と Azure AI Search を使用した ChatGPT と Enterprise データ (GitHub) 。Azure AI Search を OpenAI の大規模な言語モデルと組み合わせるための Python コードとテンプレート。背景については、こちらの Tech Community のブログ記事「 Revolutionize your Enterprise Data with ChatGPT 」(ChatGPT を使用してエンタープライズデータに革命を起こす)を参照してください。
重要なポイント:		
Azure AI Search を使用して、検索可能なコンテンツを統合してインデックスを作成します。		

Month	Type	お知らせ
		インデックスにクエリを実行して最初の検索結果を取得します。
		これらの結果からプロンプトをアセンブルし、Azure OpenAI の gpt-35-turbo (レビュー) モデルに送信します。
		クロス ドキュメントの回答を返し、顧客向けアプリで引用文献と透明性を提供して、ユーザーが応答を評価できるようにします。

前年度のお知らせ

- [2022 年のお知らせ](#)
- [2021 年のお知らせ](#)
- [2020 年のお知らせ](#)
- [2019 年のお知らせ](#)

サービスのブランド変更

このサービスには長年にわたって複数の名前が付けられてきました。ここでは、それらを逆時系列順に示します。

- **Azure AI 検索** (2023 年 11 月) Azure AI サービスと顧客の期待に合わせて名前が変更されました。
- **Azure Cognitive Search** (2019 年 10 月) サービス操作でのコグニティブ スキルと AI 処理の使用の拡大 (ただし省略可能) を反映するために名前が変更されました。
- **Azure Search** (2015 年 3 月) 最初の名前です。

サービスの更新情報

Azure AI Search に関する[サービス更新のお知らせ](#)は、Azure の Web サイトで参照できます。

機能名の変更

セマンティック検索は、既存の結果セットの L2 ランク付けを提供する機能をより適切に説明するために、2023 年 11 月に[セマンティック ランク付け](#)に名前が変更されました。

Azure AI Search の機能

[アーティクル] • 2024/04/04

Azure AI Search は情報取得機能を備え、オプションの AI 統合を利用してさらなるテキストと構造コンテンツを抽出します。

次の表は、カテゴリ別の機能をまとめたものです。Azure AI Search と他の検索テクノロジの比較については、「[検索オプションの比較](#)」を参照してください。

Azure のあらゆるパブリック クラウド、プライベート クラウド、ソブリン クラウドに機能パリティがありますが、一部のリージョンでは特定の機能がサポートされていません。 詳細については、「[リージョン別の利用可能な製品](#)」を参照してください。

① 注意

プレビュー機能をお探しですか? 「[プレビュー機能リスト](#)」を参照してください。

インデックス作成機能

[:] テーブルを展開する

カテゴリ	機能
データ ソース	検索インデックスは、JSON ドキュメントとして送信されれば、すべてのソースからのテキストを受け入れることができます。 インデクサー は、プライマリ データ ストア内の検索可能なコンテンツを抽出するために、サポートされているデータ ソースからのデータ インポートを自動化する機能です。インデクサーによって JSON のシリアル化は自動的に処理されます。また、ほとんどの場合、何らかの形の変更と削除の検出がサポートされます。 Azure SQL Database、Azure Cosmos DB、Azure BLOB ストレージを含む さまざまなデータ ソース に接続できます。
階層データ構造と入れ子になったデータ構造	複合型 とコレクションでは、検索インデックス内に事実上すべての種類の JSON 構造をモデル化できます。一对多と多対多のカーディナリティは、コレクション、複合型、および複合型のコレクションを通じてネイティブに表現できます。
Linguistic Analysis	アナライザーは、インデックス作成および検索操作中のテキスト処理に使用するコンポーネントです。既定では、汎用の標準 Lucene アナライザーを使用するか、言語アナライザー、ユーザーが構成したカスタム アナライザー、または必要な形式でトークンを生成する別の定義済みアナライザーで既定をオーバーライドすることができます。

カテゴリ	機能
	<p>Lucene または Microsoft の言語アナライザーは、動詞の時制や名詞の性、不規則な複数形の名詞（例：'mouse' と 'mice'）、二重複合語、（スペースを使用しない言語の）改行などをインテリジェントに処理するために使用されます。</p> <p>カスタム字句アナライザーは、発音照合や正規表現などの複雑なクエリ形式で使用されます。</p>

ベクトル検索とハイブリッド検索

 テーブルを展開する

カテゴリ	機能
ベクター インデックス	検索インデックス内に ベクトルフィールド を追加することで、 ベクトル検索 のシナリオをサポートできるようになります。ベクトルフィールドは、同じ検索ドキュメント内の非ベクトルフィールドと共に存できます。
ベクトル クエリ	单一 および複数のベクトル クエリを作成します。
ベクトル検索のアルゴリズム	Hierarchical Navigable Small World (HNSW) または 完全な K 近傍法 (KNN) を使用して、検索インデックス内の同様のベクトルを検索します。
ベクトル フィルター	情報を取得する際に クエリの実行前または実行後にフィルターを適用 して、精度を高めます。
ハイブリッド情報の取得	1 つの ハイブリッド クエリ要求 にある概念とキーワードを検索します。 ハイブリッド検索 にはベクトル検索とテキスト検索が統合されており、またオプションのセマンティック ランク付けと関連性のチューニングを行うことで最適な結果が得られるようになっています。
データ チャンキングとベクトル化の統合 (プレビュー)	Text Split スキル によるネイティブのデータ チャンキングと、 ベクタライザー と AzureOpenAIEmbeddingModel スキル によるネイティブのベクトル化。 垂直統合 (プレビュー) を使用すると、ソース ファイルからクエリに繋がるエンドツーエンドのインデックス作成パイプラインを利用できます。
統合されたベクター圧縮と量子化	組み込みのスカラー量子化 を使って、メモリ内とディスク上のベクター インデックス サイズを圧縮します。また、不要なベクターの格納を省略したり、narrow データ型をベクター フィールドに割り当ててストレージ要件を減らすこともできます。
データのインポートとベク	Azure portal で利用できる 新しいウィザード を使用すると、データ チャンキングとベクトル化を含む完全なインデックス作成パイプラインを作成できます。

カテゴリ	機能
トル化 (プレビュー)	ウィザードでは、すべてのオブジェクトと構成設定を作成できます。

AI エンリッチメントとナレッジ マイニング

[+] テーブルを展開する

カテゴリ	機能
インデックス作成中の AI 処理	AI エンリッチメント とは、本来であれば全文検索用にインデックスを作成できないコンテンツからテキストや情報を抽出する、インデクサー パイプラインの埋め込み画像や自然言語処理です。AI 処理は、後でインデクサーにアタッチするスキルセットに、スキルを追加して組み合わせることによって実現されます。AI には、テキスト翻訳や光学式文字認識 (OCR) などの Microsoft の組み込みスキルと、お客様が提供するカスタム スキルがあります。
検索以外のシナリオでの分析と使用のためのエンリッチされたコンテンツの格納	ナレッジストア は、ナレッジマイニングやデータ サイエンス処理など、検索以外のシナリオを想定した、充実したコンテンツの永続的なストレージです。ナレッジストアは、スキルセットで定義されますが、Azure Storage ではオブジェクトまたは表形式の行セットとして作成されます。
キャッシュされたエンリッチメント	増分エンリッチメント (プレビュー) とは、スキルセットの実行中に再利用できるキャッシュされたエンリッチメントを指します。キャッシュは、処理コストが高い OCR や画像分析を含むスキルセットで特に重要です。

クエリとユーザー エクスペリエンス

[+] テーブルを展開する

カテゴリ	機能
自由形式のテキスト検索	フルテキスト検索 は、大部分の検索ベース アプリの主な用途です。クエリは、サポートされている構文を使用して作成できます。
単純なクエリ構文	単純なクエリ構文 では、論理演算子、語句検索演算子、後置演算子、優先順位演算子を使用できます。
完全な Lucene クエリ構文	完全な Lucene クエリ構文 には、あいまい検索、近接検索、用語ブースト、正規表現の拡張機能を含む、簡単な構文でのすべての操作が含まれています。

カテゴリ	機能
関連性	簡単なスコアリング は Azure AI Search の主な利点です。スコアリングプロファイルを使用して、ドキュメント自体の値の関数として、関連性をモデル化できます。たとえば、新しい製品や割り引き製品を検索結果の上位に表示することが望ましい場合があります。あるいは、追跡記録し、個別に保存しておいた顧客の検索傾向に基づいてパーソナライズされたスコアリングのタグを利用し、スコアリングプロファイルを作成できます。
	セマンティック ランカー は、クエリとのセマンティックな関連性に基づいて結果を再ランク付けする Premium 機能です。コンテンツとシナリオによっては、ほぼ最小限の構成または労力で検索の関連性を大幅に向上させることができます。
地理空間検索	地理空間関数 は、地理座標をフィルター処理して照合します。距離で一致させることも、多角形に含めて一致させることもできます。
フィルターと ファセット	1つのクエリ パラメーターで ファセットナビゲーション を有効にできます。Azure AI Search は、ファセットナビゲーション構造を返します。これをカテゴリ一覧の背後のコードとして使用すると、自律フィルター処理(たとえば、価格帯やブランド別のカタログ品目のフィルター処理)を実現できます。
	フィルター を使用することで、ファセットナビゲーションをアプリケーションの UI に組み込み、クエリ形成を拡張し、ユーザーまたは開発者が指定した条件に基づいてフィルター処理することができます。フィルターを作成するには、OData 構文を使用します。
ユーザースペース	検索バーでの先行入力クエリで、 オートコンプリート を有効にすることができます。 検索候補 も検索バーでの部分テキスト入力に使用できますが、結果はクエリ用語ではなくインデックス内の実際のドキュメントです。
スコアリング	シノニム は、ユーザーが代替用語を提供する必要がなく、クエリのスコープを暗黙的に拡張する同等の用語を関連付けます。
	ヒットの強調表示 は、検索結果内の一一致するキーワードにテキスト書式を適用します。強調表示されるスニペットを返すフィールドを選択できます。
	並べ替え は、インデックススキーマを介して、複数のフィールドで利用でき、クエリ時に1つの検索パラメーターで切り替えることができます。
	検索結果の ページング とスロットは、Azure AI Search によって検索結果に適用される微調整された制御によって簡単に行うことができます。

セキュリティ機能

[+] テーブルを展開する

カテゴリ	機能
データの暗号化	<p>Microsoftによって管理される保存中の暗号化は内部ストレージ レイヤーに組み込まれており、取り消しはできません。</p> <p>ユーザーが Azure Key Vault で作成および管理するユーザー管理の暗号化キーは、インデックスとシノニム マップの補助的な暗号化に使用できます。 インデックス付きコンテンツを完全に二重に暗号化する CMK での暗号化は、2020 年 8 月 1 日以降に作成されたサービスでは、一時ディスク上のデータにも拡張されています。</p>
エンドポイント保護	<p>受信ファイアウォールの IP ルールのサポートを使用すると、検索サービスが要求を受け入れる IP 範囲を設定できます。</p> <p>Azure Private Link を使用してプライベート エンドポイントを作成し、すべての要求が仮想ネットワークを経由するようにします。</p>
受信アクセス	<p>Azure ロールベースのアクセス制御では、検索コンテンツと操作へのアクセスを制御するため、Microsoft Entra ID のユーザーとグループにロールが割り当てられます。 Azure テナントがない場合、キーベースの認証を使用することもできます。</p>
アウトバウンド セキュリティ (インデクサー)	<p>プライベート エンドポイントを使用したデータ アクセスにより、インデクサーは Azure Private Link を介して保護されている Azure リソースに接続できます。</p> <p>信頼された ID を使用したデータ アクセスは、外部データ ソースへの接続文字列でユーザー名とパスワードを省略できることを意味します。 植物検索サービスが以前に信頼されたサービスとして登録されていた場合、インデクサーがデータ ソースに接続するときに、リソースは接続を許可します。</p>

ポータル機能

[+] テーブルを展開する

カテゴリ	機能
プロトタイプ ピングと検査のためのツール	<p>インデックスの追加は、ポータルのインデックス デザイナーであり、属性付きのフィールドと他のいくつかの設定で構成される基本スキーマを作成するために使用できます。 インデックスを保存したら、SDK または REST API を使用して情報を移入し、データを提供できます。</p> <p>データのインポート ウィザードでは、インデックス、インデクサー、スキルセット、およびデータ ソースの定義が作成されます。 データが Azure に存在する場合、このウィザードを使用すると、特に概念実証の調査や探究で時間と労力を大幅に節約できます。</p>

カテゴリ	機能
	<p>検索エクスプローラーは、クエリをテストしたり、スコアリング プロファイルを調整したりするために使用します。</p> <p>デモ アプリの作成は、検索エクスペリエンスのテストに使用できる HTML ページを生成するために使用します。</p> <p>デバッグ セッションは、スキルセットを対話形式でデバッグできるビジュアルエディターです。依存関係、出力、変換が表示されます。</p>
監視と診断	<p>監視機能を有効にして、ポータルに常に表示される、ひとめで確認できるメトリックの先に進みます。1秒あたりのクエリ数、待ち時間、スロットルに関するメトリックが取得され、ポータルページで報告されます。追加の構成は必要ありません。</p>

プログラミング

[+] [テーブルを展開する](#)

カテゴリ	機能
REST	<p>サービス REST API は、データプレーン操作で使用します (インデックス作成、クエリ、および AI エンリッチメントに関連するすべての操作を含む)。また、このクライアントライブラリを使用して、システム情報と統計を取得することもできます。</p> <p>管理 REST API は、Azure Resource Manager によるサービスの作成とプロビジョニング用です。また、この API を使用して、キーと容量を管理することもできます。</p>
Azure SDK for .NET	<p>Azure.Search.Documents は、データプレーン操作で使用します (インデックス作成、クエリ、および AI エンリッチメントに関連するすべての操作を含む)。また、このクライアントライブラリを使用して、システム情報と統計を取得することもできます。</p> <p>Microsoft.Azure.Management.Search は、Azure Resource Manager によるサービスの作成とプロビジョニング用です。また、この API を使用して、キーと容量を管理することもできます。</p>
Azure SDK for Java	<p>com.azure.search.documents は、データプレーン操作で使用します (インデックス作成、クエリ、および AI エンリッチメントに関連するすべての操作を含む)。また、このクライアントライブラリを使用して、システム情報と統計を取得することもできます。</p> <p>com.microsoft.azure.management.search は、Azure Resource Manager によるサービスの作成とプロビジョニング用です。また、この API を使用して、キーと容量を管理することもできます。</p>

カテゴリ	機能
Azure SDK for Python	<p>azure-search-documents は、データプレーン操作で使用します(インデックス作成、クエリ、および AI エンリッチメントに関連するすべての操作を含む)。また、このクライアントライブラリを使用して、システム情報と統計を取得することもできます。</p> <p>azure-mgmt-search は、Azure Resource Manager によるサービスの作成とプロビジョニング用です。また、この API を使用して、キーと容量を管理することもできます。</p>
Azure SDK for JavaScript/TypeScript	<p>azure/search-documents は、データプレーン操作で使用します(インデックス作成、クエリ、および AI エンリッチメントに関連するすべての操作を含む)。また、このクライアントライブラリを使用して、システム情報と統計を取得することもできます。</p> <p>azure/arm-search は、Azure Resource Manager によるサービスの作成とプロビジョニング用です。また、この API を使用して、キーと容量を管理することもできます。</p>

関連項目

- [Azure AI Search の新機能](#)
- [Azure AI Search のプレビュー機能](#)

Azure AI Search についてよく寄せられる質問

よく寄せられる質問

Azure AI Search についてよく寄せられる質問と回答を確認してください。

全般

Azure AI Search とは？

Azure AI Search は、フルテキストおよびベクトル検索シナリオ向けに専用の検索エンジンと検索可能なコンテンツの永続的なストレージを提供します。また、より多くのテキストと構造を生コンテンツから抽出し、ベクトル検索のためにコンテンツをチャンク化およびベクトル化するための、省略可能な統合 AI も含まれています。

Azure AI Search を使用する方法

プライマリ ワークフローは、インデックスの作成、読み込み、クエリです。ポータルはほとんどのタスクに使用できますが、Azure AI Search はプログラムで使用し、クライアントコードから要求を処理することを目的としています。プログラムによるサポートは、REST API と .NET、Python、Java、JavaScript SDK for Azure のクライアントライブラリを通じて提供されます。

"Azure Search" と "Azure Cognitive Search" と "Azure AI Search" は同じ製品ですか？

Azure Search は、コア操作でのコグニティブ スキルと AI 処理の利用拡大 (任意) を反映するために、2019 年 10 月に Azure Cognitive Search に名前が変更されました。2023 年 10 月、Azure Cognitive Search は、Azure AI サービスに合わせて Azure AI Search に名前が変更されました。

サポートされている言語は何ですか？

トクン化に使われる既定のアナライザーは標準の Lucene であり、言語に依存しません。それ以外の場合、言語サポートは、受信 (インデックス作成) と送信 (クエリ) コンテンツに言語ルールを適用する[言語アナライザー](#)によって表現されます。スペル チェックなどの一部の機能は、言語のサブセットに制限されています。

検索をソリューションに統合する方法

クライアントコードでは、Azure SDK クライアント ライブラリまたは REST API を呼び出して検索インデックスに接続し、クエリを作成し、応答を処理する必要があります。インデックスをビルドして更新するコードを記述したり、インデクサーをプログラムやスクリプトで実行したりすることもできます。

さまざまな API で機能に類似性はありますか？

必ずしもそうではありません。REST API は常に、プレビュー API バージョンで新機能を実装する最初の機能です。Azure SDK のクライアント ライブラリは、一定期間内で新しい機能が選択されますが、独自のスケジュールでリリースされます。

REST API は最新のフィーチャーが最初に用意されますが、Azure SDK は、より多くのコーディングがサポートされているので、必要なフィーチャーが利用できない場合を除き、REST よりも推奨されます。

サービスを一時停止して課金を停止できますか？

検索サービスを一時停止することはできません。Azure AI Search では、サービスの作成時にコンピューティング リソースが割り当てられます。そのリソースを必要に応じて解放し、再請求することはできません。

サービスのアップグレード、ダウングレード、名前変更、または移動ができますか？

サービス レベル、名前、リージョンは、サービスの有効期間中は固定されます。

検索サービスを別のサブスクリプションまたはリソース グループに移行する場合、ダウンタイムは発生しますか？

[リソースを移動する前にチェックリストする](#) に従い、各ステップが完了していることを確認する限り、ダウンタイムは発生しません。

インデックス作成

Azure AI Search における "インデックス作成" にはどのような意味がありますか?

検索インデックスを設定するテキストのコンテンツとトークンを取り込み、解析し、格納することを指します。インデックス作成により、情報取得をサポートする逆インデックスやその他の物理的なデータ構造が作成されます。

スキーマにベクター フィールドが含まれている場合は、ベクトル インデックスが作成されます。

インデックスの移動、バックアップ、復元はできますか?

インデックスのポーティングはネイティブでサポートされません。検索インデックスは、ダウンストリーム データ構造と見なされ、運用データを収集する他のデータ ソースからのコンテンツを受け入れます。そのため、インデックスのバックアップと復元の組み込みサポートはありません。インデックスを削除した場合、またはインデックスを移動する場合は、ソース データからインデックスをリビルドすると予想されるためです。

ただし、検索サービス間でインデックスを移動する場合は、この [Azure AI Search .NET サンプル リポジトリ](#) にある `index-backup-restore` サンプル コードを試すことができます。[バックアップと復元の Python バージョン](#) もあります。

削除されたインデックスまたはサービスは復元できますか?

いいえ。Azure AI Search インデックスまたはサービスを削除した場合、復旧できません。検索サービスを削除すると、サービス内のすべてのインデックスが完全に削除されます。

SQL Database レプリカからインデックスを作成できますか?

Azure SQL Database に対して検索インデクサーを使用している場合、インデックスを最初から構築するときに、データ ソースとしてプライマリまたはセカンダリのレプリカの使用に制限はありません。ただし、(レコードの変更に基づく) 増分更新によるインデックスの更新には、プライマリ レプリカが必要です。この要件は SQL Database に由来するものです。SQL Database は変更追跡をプライマリ レプリカでのみ保証して

います。インデックス更新ワークロードにセカンダリ レプリカを使用しようとすると、すべてのデータが取得される保証はありません。

ベクトル

ベクター検索とは

ベクトル検索は、ベクトル表現を比較して最も似たドキュメントを検索する手法です。ベクトル表現の目的は、項目の本質的な特性を数値形式でキャプチャすることであるため、キーワードまたはタグに基づいて明示的な一致がない場合でも、ベクトル クエリで類似のコンテンツを識別できます。ユーザーが検索を実行すると、クエリがベクトル表現にまとめられ、ベクトル検索エンジンにより最も類似したドキュメントが識別されます。大規模なデータベースの効率を向上させるために、ベクトル検索では、多くの場合、クエリ ベクトルの近似ニアレストネイバーが提供されます。Azure AI Search のベクトル オファリングの詳細については、「[ベクトル検索の概要](#)」を参照してください。

Azure AI Search はベクトル検索をサポートしていますか？

Azure AI Search では、ベクトルインデックス作成と取得がサポートされています。プレビュー ライブラリとベータ ライブラリを使用すると、クエリ文字列とコンテンツをベクトル化できます。

Azure AI Search でのベクトル検索のしくみ

スタンドアロン ベクトル検索では、まず埋め込みモデルを使用して、埋め込み空間内のベクター表現にコンテンツを変換します。その後、ドキュメントペイロード内のかこれらのベクトルをインデックス作成用の検索インデックスに提供できます。検索要求を処理するには、インデックス作成と同じ DNN を使用して検索クエリをベクトル表現に変換し、ベクトル検索で最も類似したベクトルを検索し、対応するドキュメントを返します。

Azure AI Search では、テキストやその他の種類のコンテンツと共に、ドキュメント内のフィールドとしてベクトルデータにインデックスを付けることができます。ベクトルフィールドのデータ型は `Collection(Edm.Single)` です。

ベクトル クエリは、スタンドアロンで、または同じ検索要求の用語のクエリやフィルターなどの他のクエリの種類と組み合わせて発行することができます。

Azure AI Search でコンテンツまたはクエリをベクトル化できますか?

組み込みの統合ベクター化は現在、パブリック プレビュー段階になりました。

検索サービスでベクトル検索はサポートされていますか?

既存のサービスのほとんどではベクトル検索がサポートされています。ベクトル検索をサポートするパッケージや API を使用していて、インデックスの作成に失敗した場合、基になる検索サービスはベクトル検索をサポートしていないため、新しいサービスを作成する必要があります。これは、2019 年 1 月 1 日より前に作成されたサービスのごく一部で発生する可能性があります。

既存のインデックスにベクトル検索を追加できますか?

検索サービスでベクトル検索がサポートされている場合、既存と新規の両方のインデックスで、ベクトル フィールドに対応できます。

新しい検索サービスと既存の検索サービスの間で異なるベクトル インデックス サイズの制限が表示されるのはなぜですか?

Microsoft では新しい検索サービスの改善されたベクトル インデックス サイズの制限を世界全域にロールアウトしますが、特定のリージョンでは引き続きインフラストラクチャ容量を構築します。サポートされているリージョンで作成された新しい検索サービスでは、ベクトル インデックス サイズの制限の増加が見られます。残念ながら、既存のサービスを新しい制限に移行することはできません。

検索インデックスでベクトル検索を有効にする方法

インデックスでベクトル検索を有効にするには、次の手順に従います:

- フィールド コレクションに 1 つ以上のベクトル フィールドを追加します。
- ベクトル検索 フィールドで使用される構成を指定するインデックス スキーマに "vectorSearch" セクションを追加します。これには、HNSW など、使用される近

似ニアレストネイバー アルゴリズムのパラメーターが含まれます。

- [2023-11-01](#) または Azure SDK を使用して、インデックスの作成または更新、ドキュメントの読み込み、クエリの発行を行います。

クエリ

クエリの実行はどこで行われますか?

クエリは、検索サービス上でホストされている 1 つの検索インデックスに対して実行されます。複数のインデックスを結合して 2 つ以上のインデックス内のコンテンツを検索することはできませんが、[複数の検索サービスで同名のインデックスに対してクエリを実行する](#)ことはできます。

存在するとわかっている用語でも一致数が 0 件なのはなぜですか?

よくあるケースは、クエリの型ごとにサポートされる検索ビヘイビアと言語分析のレベルが異なることを把握していない場合です。主要なワークロードである全文検索には、用語を原形に分解する言語分析フェーズが含まれています。トークン化された用語は、より多くの数の変形と一致するため、このようなクエリ分析はより広い網を一致候補にかけます。

ただし、ワイルドカード、あいまいクエリ、正規表現クエリは、通常の用語または語句クエリのように分析されず、クエリが検索インデックス内の分析されたワードフォームと一致しない場合、リコールが不適切になる可能性があります。クエリの解析と分析について詳しくは、[クエリのアーキテクチャ](#)に関する記事をご覧ください。

ワイルドカード検索が遅い理由

ほとんどのワイルドカード検索クエリ(プレフィックス、あいまい、正規表現など)は、検索インデックス内の一一致する用語を使用して内部で再作成されます。この追加の処理により、待ち時間が増えます。さらに、`a*` などの広範な検索クエリでは、多くの用語で書き換えられる可能性が高いと、低速になる可能性があります。ワイルドカード検索のパフォーマンスを向上させるには、[カスタム アナライザー](#)の定義を検討してください。

複数のインデックスを対象に検索できますか?

いいえ。クエリは常に単一のインデックスが対象です。

検索スコアが一致ごとに 1.0 の定数である理由

検索スコアは、[一致する用語の統計プロパティ](#)に基づいて、フルテキスト検索クエリに対して生成され、結果セットで上位から下位に順序付けされます。フルテキスト検索ではないクエリの型(ワイルドカード、プレフィックス、正規表現)は、関連性スコアで順位付けされません。この動作は仕様です。定数のスコアによって、ランクには影響を与えずに、クエリ拡張によって見つかった一致を結果に反映することができます。

たとえば、ワイルドカード検索に「tour*」と入力すると、「tours」、「tourettes」、「tourmaline」との一致が生成されます。こうした結果の性質上、用語の相対的な重みを適切に推測することができません。そのため、ワイルドカード、プレフィックス、正規表現の種類のクエリでは、結果のスコア付けを行う際に、語句の出現頻度が無視されます。予期しない一致に対するバイアスを回避するために、部分的な入力に基づく検索結果には一定のスコアが与えられます。

セキュリティ

Azure AI Search では、顧客データはどこに保存されますか？

サービスがデプロイされている場所にデータが保存されます。Azure AI Search では、顧客データはデプロイ リージョンの外部には保存されません。

Azure AI Search では、処理のために他のサービスに顧客データを送信しますか？

はい。Azure AI サービスに基づいて組み込みのスキルを使用すると、インデクサーは内部ネットワーク経由で Azure AI サービスに要求を送信します。カスタム スキルを追加すると、インデクサーはパブリック ネットワーク経由でカスタム スキルで指定された URI にコンテンツを送信します。

ユーザー ID に基づいて検索結果へのアクセスを制御できますか？

そうとも限りません。通常、アプリケーションを実行する権限を持つユーザーには、すべての検索結果を表示する権限も与えられます。Azure AI Search では、行レベルまたはドキュメントレベルのアクセス許可は組み込みサポートされていませんが、回避策として [セキュリティ フィルター](#) を実装できます。

ユーザー ID に基づいて操作へのアクセスを制御できますか？

はい。コンテンツに対するデータプレーン操作への[ロールベースの承認](#)を使用できます。

検索サービスが IP ファイアウォールまたはプライベート エンドポイントの内側にある場合、Azure portal を使用して検索コンテンツを表示および管理できますか？

クライアントとポータルへのアクセスを許可するネットワーク例外を作成する場合は、ネットワークで保護された検索サービスで Azure portal を使用できます。 詳細については、[IP ファイアウォール経由での接続](#) または [プライベート エンドポイント経由の接続](#)に関する説明を参照してください。

次のステップ

こちらでご質問の回答が見つからない場合は、次のソースでさらに質問と回答を参照できます。

[Stack Overflow: Azure AI Search ↗](#)

[Azure AI Search でのフルテキスト検索のしくみ](#)

[Azure AI Search とは？](#)