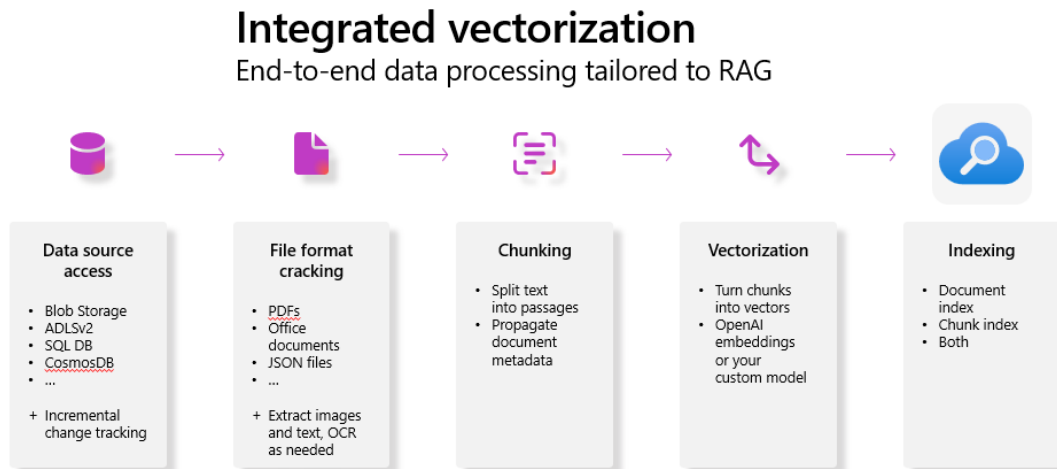


Azure AI Search - Create a Hybrid Search Solution (including vector search)

In this lab, you will create a hybrid search solution using Azure AI Search integrated vectorization solution. You can do this task using Azure AI Studio or Azure Portal. We are going to do it using Azure AI Studio.



We will use sample data (uploaded to Azure storage account) in this step. We will also use Azure OpenAI embedding model to turn document chunks into vectors.

Therefore, this step will require Azure AI Studio to connect to your Azure Storage account (to get pdf documents) and Azure OpenAI Service (for embedding model)

1. Prerequisites:


- a. Azure OpenAI service with embedding model deployed.
- b. Azure storage account
- c. Azure AI Studio with AI Hub created.

2. Upload sample data (PDF documents) to Azure storage container (using Azure portal).

- a. Create a container within the Azure storage account you created before. To do this
 - i. Go to the Azure storage account (in Azure portal)
 - ii. Click on Containers
 - iii. Click on the “+ Container” to create a new container.
 - iv. Give the container a name (Ex: ragdatacontainer) and click Create.
- b. Upload sample data within this container. To do this:
 - i. Get all PDF documents from the GitHub repository. Sample data available on GitHub at <https://github.com/amulchapla/GenAI-LLMapps-lifecycle-workshop/tree/main/AzureAISearch-VectorSearch>
 - ii. Go to the Container you created and click Upload.
 - iii. **VERY IMPORTANT**: Enter “Upload to folder” name as “hrbenefitsdata” (as shown below)
 - iv. Select all PDF documents to upload and click “Upload” button.

Upload blob





6 file(s) selected: Benefit_Options.pdf, employee_handbook.pdf, Nort...

Drag and drop files here or [Browse for files](#)

☐ Overwrite if files already exist

Advanced

Blob type ⓘ

Block blob

☒ Upload .vhd files as page blobs (recommended)

Block size ⓘ

4 MiB

Access tier ⓘ

Hot (Inferred)

Upload to folder

hrbenefitsdata

Blob index tags ⓘ

Key

Value

3. Add your data (in Azure AI Studio)

In this step, we will bring our own data so you can chat with your own data using the GPT models.

- Go to Azure AI Studio. You can login to AI Studio at <https://ai.azure.com/> OR go to your Azure OpenAI service overview page and click on “Build in Azure AI Studio” button.
- Once in Azure AI Studio, click on the Project.
- Click on Playground (on the left hand side)
- Go to “Add your data” tab
- Click on “Add your data” and click on “Add connection” button. Here you are going to connect AI Studio to your Azure Storage account.
- Enter information related to your storage account (see screenshot below). You will need to enter Account Key for your Azure storage account which you can get from Azure Portal.

Add a connection to external assets

Select your service

Service *
 Azure Blob Storage

Account selection method *
☐ Manually enter account information

Subscription ID *
 MCAPS-Hybrid-REQ-43647-2022-amchapla (ef55aae2-f9b7-40df-94ca-ee242aa44595)

Storage account *
 stazureaiamu861189642980 (azuremlworkspacerg)

Blob container *
 ragdatacontainer

Authentication method * ⓘ
 Credential based

Authentication type * ⓘ
 Account key

Account key * ⓘ

Connection name * ⓘ
 ragdatastoragecontainer

Access ⓘ
 Project

Create connection Cancel

g. Select Azure AI Search service

Add your data

- Source data
- Index storage
- Search settings
- Index settings
- Review and finish

Index storage
 Select where to store your new index

Index storage *
 Azure AI Search

Select Azure AI Search service * ⓘ
 gptkbtthxdj6b6uiqtk

Back Next Create Cancel

h. Configure Azure OpenAI service connection for Embedding model. Embedding model is used to create embedding of your documents in this step.

Add your data

- ✓ Source data
- ✓ Index storage
- 3 Search settings
- ✓ Index settings
- 5 Review and finish

Configure search settings
Adding vector search supports: Hybrid (vector + keyword search), Hybrid + Semantic (most accurate search results for generative AI applications), Vector, Semantic and Keyword retrieval. Hybrid will be set as default and can be changed at inference time in the playground. Not adding vector search supports: Keyword and Semantic retrieval. Keyword will be set as default and can be changed at inference time in the playground. Adding vector search requires an Azure OpenAI embedding model. [Learn more](#)

Vector settings
☒ Add vector search to this search resource

Select an embedding model [?] ⓘ
Default_AzureOpenAI

☒ I acknowledge that Azure OpenAI embedding model, **text-embedding-ada-002 (Version 2)**, will be deployed if not already. It will incur usage to my account. [View pricing](#)

Back Next Create Cancel

- i. Enter index name and other settings (you can leave them default)

Add your data

- ✓ Source data
- ✓ Index storage
- ✓ Search settings
- 4 Index settings
- 5 Review and finish

Index settings
Configure your index

Index name [?] ⓘ
ragdataaistudioindex

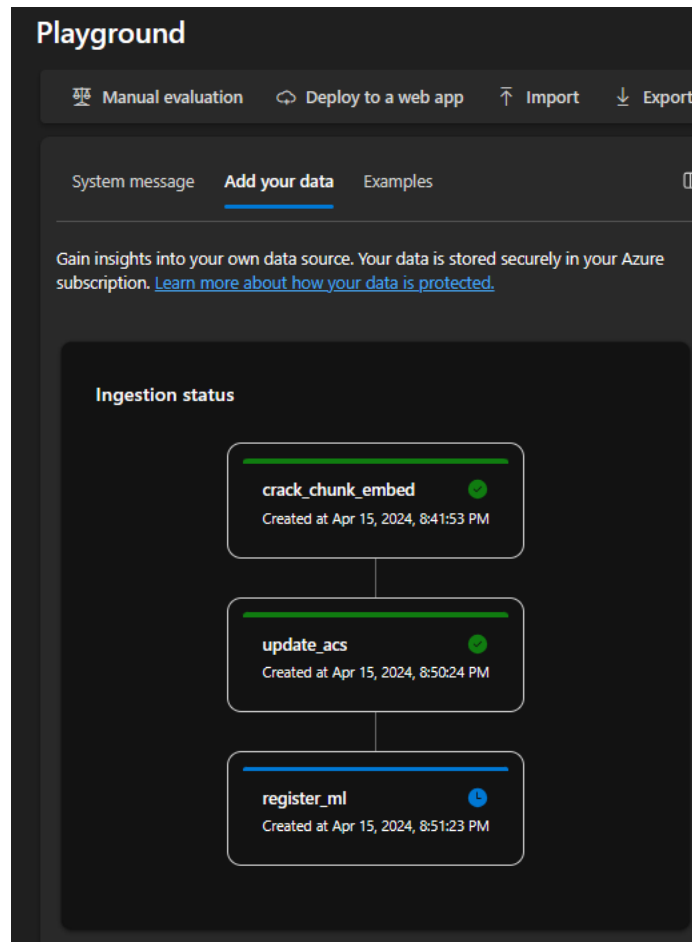
Schedule updates [?] ⓘ
One time indexing (no scheduled updates) ▾

Virtual machine [?] ⓘ
☒ Auto select ☐ Select from recommended options ☐ Select from all options

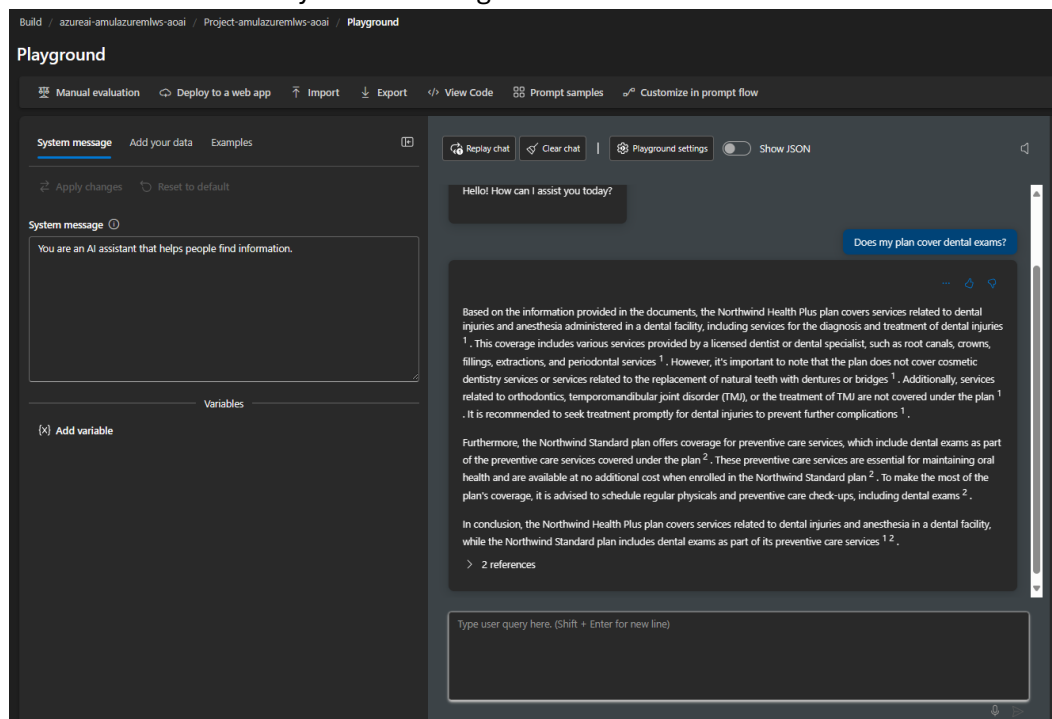
Selecting a virtual machine will incur additional costs.

Back Next Create Cancel

- j. Review information and click Create.
- k. It will take a few minutes to complete the process (shown below). It will update the status as shown below.



- l. You can now chat with your data using the GPT model.



4. [OPTIONAL] Create Azure AI Search service and create vector store

- Login to Azure Portal
- Create a new Azure AI Search service (if one doesn't exist already)
- Go to Azure AI Search service and click "Import and vectorize data". Follow the data import wizard.
 - Set up your data connection:** Connect to an external Azure Blob storage containing PDFs and other unstructured data files. You can leave Blob folder empty.

[Home](#) > [Azure AI services | AI Search](#) > [gptkb-thxdj6b6uiqtk](#) >

Import and vectorize data

gptkb-thxdj6b6uiqtk

● Connect to your data

○ Vectorize and enrich data

○ Review and create

Set up your data connection

Connect to an external Azure Blob storage containing PDFs and other unstructured data files. [Learn more](#)

Subscription *

Blob storage account *

Blob container * ⓘ

sourcedata-container

Blob folder ⓘ

your/folder/here

☐ Enable deletion tracking ⓘ

☐ Authenticate using managed identity. [Learn more](#)

- Vectorize your data with Azure OpenAI:** Connect to an Azure OpenAI service and select an embedding model for vector generation.

You need Azure OpenAI service with embedding model deployed for this step. If you don't have this then complete that task first.

[Home](#) > [Azure AI services | AI Search](#) > [gptkb-thxdj6b6uiqtk](#) >

Import and vectorize data

gptkb-thxdj6b6uiqtk

● Connect to your data

● Vectorize and enrich data

○ Review and create

Vectorize your data with Azure OpenAI

Connect to an Azure OpenAI service and select an embedding model for vector generation. [Learn more](#)

Subscription *

Azure OpenAI service * ⓘ

aoaiamuleastus2

[Create a new Azure OpenAI service](#)

Model deployment * ⓘ

text-embedding-ada-002

Authentication type ⓘ

☒ API key ☐ System assigned ☐ User assigned

☒ I acknowledge that connecting to an Azure OpenAI service will incur additional costs to my account. [View pricing](#)

Enrich your data with AI skills

If documents are image files, or contain embedded images, you can merge the text extracted from those images with the textual content of the documents. You will need to select an Azure AI service or create a new one in the same region as your search service to power your skills and enrich your data. [Learn more](#)

☐ Extract text from images

Advanced ranking and relevancy

Semantic ranker uses deep neural networks to provide relevant results and answers based on semantics, not just lexical analysis. [Learn more](#)

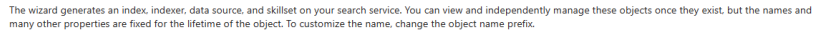
☒ Enable semantic ranker

Schedule indexing

Schedule

Once

- Update the search index name (optional):** The wizard generates an index, indexer, data source, and skillset on your search service. You can view and independently manage these objects once they exist, but the names and many other properties are fixed for the lifetime of the object. To customize the name, change the object name prefix.



vector-hrbenefits

Attached Azure OpenAI service	aoai:muleastus2
Deployment model	text-embedding-ada-002
Extracting text from images	Disabled
Semantic ranker	Enabled
Indexer run schedule	Once

+ Add index   Refresh  Delete

- | vector ✕ | | | |
|-----------------------------------|----------------|-------------------|--------------------|
| Name | Document count | Vector index size | Total storage size |
| vector-hrbenefits | 410 | 2.43 MB | 12.14 MB |

- vector-hrbenefits ...

Documents ⓘ	Total storage ⓘ	Vector index size ⓘ
410	12.14 MB	2.43 MB

 Query options View

Search

```

1 {
2   "odata.context": "https://gptkb-thxjd6b6uiqtgk.search.windows.net/indexes('vector-hrbenefits')/$metadata#dc
3   @search.answers": [
4     {
5       "key": "1227e008194b_aHr0cHm6Ly9zdGF6dXJlYW1hbXU4NjExODk2NDI5ODAuYmxvY15jb3JlLndpbmRvd3MubmV0L3NvdXJzZk
6       "text": "Performance reviews are a two-way dialogue between managers and employees. We encourage all
7       "highlights": "Performance reviews are<em>a two-way dialogue between managers and employees.</em> We
8       "score": 0.99560546875
9     }
10  ],
11  "@search.nextPageParameters": {
12    "search": "what is performance review",
13    "queryType": "semantic",
14    "semanticConfiguration": "vector-hrbenefits-semantic-configuration",
15    "captions": "extractive",
16    "answers": "extractive|count-3",
17    "queryLanguage": "en-US",
18    "skip": 50,
19    "vectorQueries": [
20      {
21        "kind": "text",
22        "k": 5,
23        "oversampling": null,
24        "fields": "vector",
25        "vector": [],

```