

Contents

What is Azure Cognitive Search?	2
Import Data.....	4

Azure Cognitive Search

What is Azure Cognitive Search?

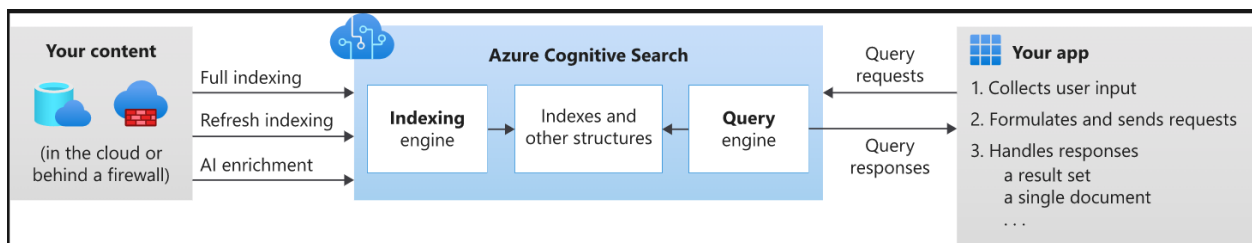
Azure Cognitive Search ([formerly known as "Azure Search"](#)) is a cloud search service that gives developers infrastructure, APIs, and tools for building a rich search experience over private, heterogeneous content in web, mobile, and enterprise applications.

Search is foundational to any app that surfaces text content to users, with common scenarios including catalog or document search, online retail, or knowledge mining for data science.

When you create a search service, you'll work with the following capabilities:

- A search engine for full text search with storage for user-owned content in a search index
- Rich indexing, with [text analysis](#) and [optional AI enrichment](#) for advanced content extraction and transformation
- Rich query capabilities, including simple syntax, full Lucene syntax, and typeahead search
- Programmability through REST APIs and client libraries in Azure SDKs for .NET, Python, Java, and JavaScript
- Azure integration at the data layer, machine learning layer, and AI (Cognitive Services)

Architecturally, a search service sits between the external data stores that contain your un-indexed data, and your client app that sends query requests to a search index and handles the response.



Across the Azure platform, Cognitive Search can integrate with other Azure services in the form of *indexers* that automate data ingestion/retrieval from Azure data sources, and *skillsets* that incorporate consumable AI from Cognitive Services, such as image and natural language processing, or custom AI that you create in Azure Machine Learning or wrap inside Azure Functions.

Key terms used in the hackathon:

Indexer	A configuration object specifying a data source, target index, an optional skillset, optional schedule, and optional configuration settings for error handling and base-64 encoding.
Data Source	Persists connection information to source data, including credentials. A data source object is used exclusively with indexers.
Index	Physical data structure used for full text search and other queries.

Import Data

The **Import data wizard** in the Azure portal creates multiple objects used for indexing and AI enrichment on a search service.

In the [Azure portal](#), open the search service page from the dashboard or [find your service](#) in the service list. In the service Overview page at the top, click **Import data**.



When you click on Import data, it will start the wizard. The wizard is organized into four main steps:

1. Connect to a supported Azure data source.

- Select Azure CosmosDB from the datasource drop down menu.

A screenshot of the 'Import data' wizard in the Azure portal, specifically the 'Connect to your data' step. The wizard has a title bar 'Import data' with a close button. Below the title bar are four tabs: 'Connect to your data' (active), 'Add cognitive skills (Optional)', 'Customize target index', and 'Create an indexer'. A descriptive text states: 'Create and load a search index using data from an external data source. Azure Cognitive Search crawls the data structure you provide, extracts searchable content, optionally enriches it with cognitive skills, and loads it into an index.' Below this is a 'Data Source' section with a dropdown menu currently set to 'Existing data source'. At the bottom, there are three columns with headers: 'Name' (with up/down arrows), 'Type' (with up/down arrows), and 'Table/Collection' (with up/down arrows). The table area currently shows 'No results.'

- In the **data source** page, the source must be **Cosmos DB**, with the following specifications:
- **Name** is the name of the data source object. Once created, you can choose it for other workloads.
- **Cosmos DB account** should be in one of the following formats:
 - The primary or secondary connection string from Cosmos DB with the following format: AccountEndpoint=https://<Cosmos DB account name>.documents.azure.com;AccountKey=<Cosmos DB auth key>;.
 - Select Choose an existing connection and it will show you the CosmosDb instance that is created in the resource group
- **Database** is an existing database from the account. - fsihack
- **Collection** is a container of documents. Documents must exist in order for import to succeed. - claims
- **Query** can be blank if you want all documents, otherwise you can input a query that selects a document subset. **Query** is only available for the SQL API.

2. Create an index schema, inferred by sampling source data.

- Optionally, add AI enrichments to extract or generate content and structure. Inputs for creating a knowledge store are collected in this step. For our hackathon we will “Skipto : Cutsomize target index”
- Change the index name to “claimsindex”. The index that we will create will be used in the front-end UI application. If we want to make the field available to filter on, select the appropriate configuration. Select Next to create the indexer

Import data ... ×

Field name	Type	Retrievable	Filterable	Sortable	Facetable	Searchable	Analyzer	Suggester
Address	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
CountryRegion	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
DateOfBirth	Edm.DateTi...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
DateOfExpiration	Edm.DateTi...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
DocumentNumber	Edm.Stri...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
DocumentType	Edm.String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
FileName	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
FirstName	Edm.String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
LastName	Edm.String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
MachineReadableZone	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Nationality	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Region	Edm.String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
Sex	Edm.String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
claimId	Edm.String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luce...	<input type="checkbox"/>
formtype	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
id	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
rid	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

Previous: Add cognitive skills (Optional) Next: Create an indexer

- Run the wizard to create objects, load data, set a schedule and other configuration options. Change it to run hourly and click submit.

Import data ...

* Connect to your data Add cognitive skills (Optional) * Customize target index **Create an indexer**

Indexer

Name * ✓

Schedule ⓘ Once Hourly Daily Custom

Description

Advanced options

- Upon successful completion of the indexer it will process the document in cosmosdb and make it available as searchable index document in Azure search.

[+ Add index](#)
[Import data](#)
[Search explorer](#)
[Refresh](#)
[Delete](#)
[Move](#)

Get 99.9% availability guaranteed with 3 replicas or more. →

^ Essentials JSON View

Resource group (Move) : FSIUC3-478734 Url : https://od478734azs.search.windows.net
 Status : Running Pricing tier : Basic
 Location : East US Replicas : 1 (No SLA)
 Subscription (Move) : Shared Sub - 156 Partitions : 1
 Subscription ID : a96b8f1a-f683-4f26-918c-f083e93873fc Search units : 1
 Tags (Edit) : DeploymentId : 478734 LaunchId : 18597 LaunchType : ON_DEMAND_LAB TemplateId : 3933 TenantId : 277

Would you like to try our new semantic search capabilities? Sign up for the preview.

Get Started Usage Monitoring Indexes Indexers Data sources Skillsets Debug sessions

[+ New Indexer](#)

Status ↑↓	Name ↑↓	Last run ↑↓	Docs succeeded	Errors/Warnings
Success	claimsindexer	a few seconds ago	7/7	0/0

7. To view the documents, click on Index tab and select the index name

[+ Add index](#)
[Import data](#)
[Search explorer](#)
[Refresh](#)
[Delete](#)
[Move](#)

Get 99.9% availability guaranteed with 3 replicas or more. →

^ Essentials JSON View

Resource group (Move) : FSIUC3-478734 Url : https://od478734azs.search.windows.net
 Status : Running Pricing tier : Basic
 Location : East US Replicas : 1 (No SLA)
 Subscription (Move) : Shared Sub - 156 Partitions : 1
 Subscription ID : a96b8f1a-f683-4f26-918c-f083e93873fc Search units : 1
 Tags (Edit) : DeploymentId : 478734 LaunchId : 18597 LaunchType : ON_DEMAND_LAB TemplateId : 3933 TenantId : 277

Would you like to try our new semantic search capabilities? Sign up for the preview.

Get Started Usage Monitoring Indexes Indexers Data sources Skillsets Debug sessions

[+ New Index](#)

Name ↑↓	Document Count ↑↓	Storage Size ↑↓
claimsindex	0	0 B

8. Click on the search and it will show you indexed documents

claimsindex ...
index

Save Discard Refresh Create Demo App Delete

Query string ⓘ

API version

Search

https://od478734azs.search.windows.net/indexes/claimsindex/docs?api-version=2020-06-30-Preview&search=*

```
1 {
2   "@odata.context": "https://od478734zs.search.windows.net/indexes('claimsindex')/$metadata#docs(*)",
3   "value": [
4     {
5       "@search.score": 1,
6       "Address": null,
7       "CountryRegion": null,
8       "DateOfBirth": null,
9       "DateOfExpiration": null,
10      "DocumentNumber": null,
11      "DocumentType": null,
12      "FileName": "Georgia_Insurance.jpg",
13      "FirstName": null,
14      "LastName": null,
15      "MachineReadableZone": null,
16      "Nationality": null,
17      "Region": null,
18      "Sex": null,
19      "claimId": "Georgia",
20      "formtype": "Insurance",
21      "rid": "LXNON8FMcjMy4mNHQUFBQUFBQT090"
22    }
23  ]
24 }
```