



Guide to create and call Azure Search Service

Version 1.0

Evelyn Tan

Last updated 23 April 2021

Brief Overview

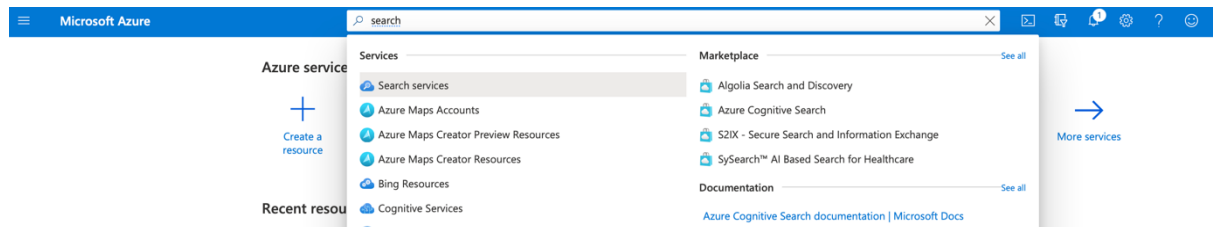
In this guide, you would need to have an Azure account created in order for you to follow through the steps below. If you do not have an Azure account created, please create one before you continue. After this guide, you would be able to understand how to create a azure search service resource, import data into azure search service and create a config file to call azure search service api using python code. The code to call azure search api can be found in the notebook ***predict-emailservice-xgboost-part1-github.ipynb***.

Content Page

Brief Overview.....	1
Step 1: Click on search service resource.....	2
Step 2: Create basics of search service resource	2
Step 3: Create scale of search service resource.....	3
Step 4: Create networking of search service resource	4
Step 5: Create tags of search service resource	4
Step 6: Import Data	5
Step 7: Connect to your data.....	6
Step 8: Customize cognitive skills	8
Step 9: Customize target index	9
Step 10: Create an indexer.....	10
Step 11: Search Explorer	10
Step 12: Search Explorer UI.....	11
Step 13: Creating Azure Search Service Config file	12

Step 1: Click on search service resource

On the top bar of the home page of portal.azure.com, type in search and click on the search service that appears at the top of this image.



Step 2: Create basics of search service resource

Once clicking the search service icon, you would be prompted to this page. Click on new to create the search service. You would need to create a resource group if you have not done so. Resource group is like a workspace where it stores all the services that is created in azure. You can also specify the location where your search service would be located at. Different location has different pricing. In this guide, I would be using Southeast Asia and all the basic tiers. Once done, click Next: Scale at the bottom of the page.

A screenshot of the "New Search Service" page in the Azure portal. The page is divided into two main sections: "Search services" on the left and "New Search Service" on the right. In the "Search services" section, the "New" button is circled in blue. In the "New Search Service" section, there are tabs for "Basics", "Scale", "Networking", "Tags", and "Review + create". The "Basics" tab is selected. Under "Project Details", there are dropdown menus for "Subscription" (set to "Microsoft Azure Enterprise") and "Resource Group" (with a "Create new" link below it). Under "Instance Details", there is a "Service name" field (with a "Required field" tooltip) and a "Location" dropdown (set to "West US 2"). There is also a "Pricing tier" dropdown. The "Standard" pricing tier is selected, with details: "25 GB/Partition*, max 12 replicas, max 12 partitions, max 36 search units" and a "Change Pricing Tier" link. At the bottom of the page, there are three buttons: "Review + create", "Previous", and "Next: Scale". The "Next: Scale" button is circled in blue.

Step 3: Create scale of search service resource

Once clicking Next: Scale, you can choose the number of replicas and partitions that you would want for your search service. In this guide, I would select 3 replicas and 1 partition. After configuring what you need, just click Next: Networking.

[Home](#) > [Search services](#) >

Search services

[+ New](#) [Manage view](#) [...](#)

Filter for any field...

Name [↑↓](#)

New Search Service

[Basics](#) [Scale](#) [Networking](#) [Tags](#) [Review + create](#)

SEARCH UNITS [ⓘ](#) EST. COST PER MON... [ⓘ](#)
1 / 36 **\$224.69**

Replicas

Replicas distribute workloads across the service. We guarantee 99.9% availability for read operations with 2 replicas, and for read and write operations with 3 or more replicas. [Learn more about the SLA for Azure Cognitive Search](#)

1

✖

99.9% availability for read operations (choose 2 or more replicas)

✖

99.9% availability for read/write operations (choose 3 or more replicas)

Partitions

Partitions allow for scale as well as faster data ingestion by spanning your index over multiple units. The allowed values of partitions are 4, 6, and 12.

1

25 GB Storage

< Page 1 of 1 >

[Review + create](#) [Previous](#) [Next: Networking](#)

Step 4: Create networking of search service resource

Once clicking Next: Networking, you can choose public or private for the endpoint. If you do not wish to allow public to access your endpoint for api calls, you should select private. In this guide, I would be selecting public. After configuring what you need, just click Next: Tags.

The screenshot shows the 'New Search Service' configuration page in the Azure portal. The left sidebar displays 'Search services' with a '+ New' button and a 'Manage view' dropdown. The main content area has tabs for 'Basics', 'Scale', 'Networking', 'Tags', and 'Review + create'. The 'Networking' tab is active, showing 'Endpoint network connectivity' with a description: 'Public endpoints can be reached by any machine on the Internet. Use a private endpoint to limit access to your service only to authorized VPNs via a private link. [Learn more](#)'. Below this, there is a section 'Endpoint connectivity (data)' with two radio buttons: 'Public' (selected) and 'Private'. At the bottom, there are navigation buttons: 'Review + create' (highlighted in blue), 'Previous', and 'Next: Tags'.

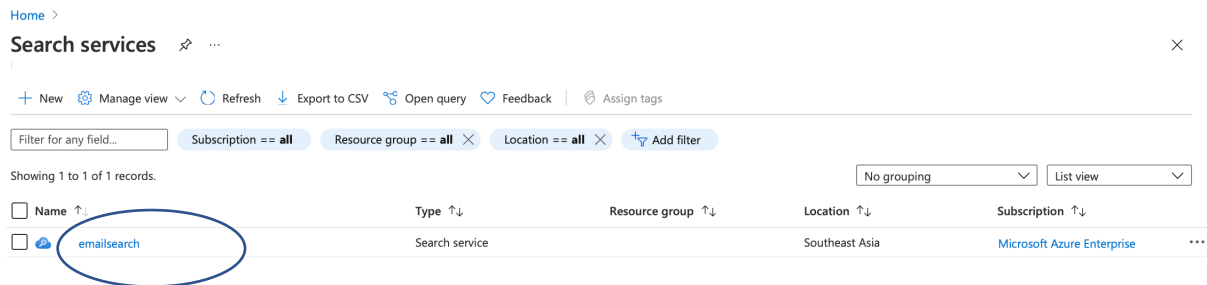
Step 5: Create tags of search service resource

Once clicking Next: Tags, you can configure the tags to associate with your resource. In this guide, I will be skipping this step as I would not need tags. After configuring what you need, just click Next: Review + create.

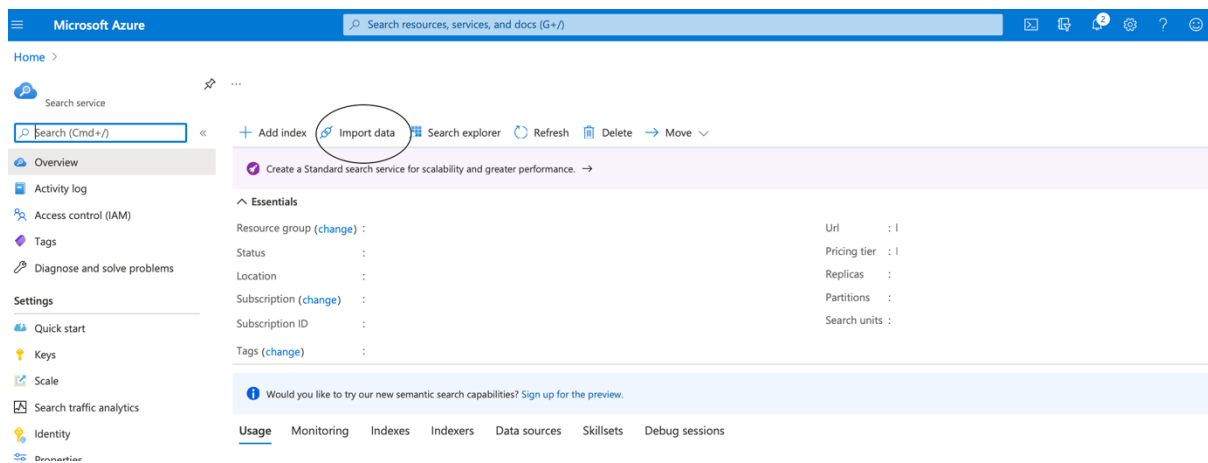
The screenshot shows the 'New Search Service' configuration page in the Azure portal, specifically the 'Tags' tab. The left sidebar is identical to the previous step. The main content area has tabs for 'Basics', 'Scale', 'Networking', 'Tags', and 'Review + create'. The 'Tags' tab is active, showing a description: 'Tags are name/value pairs that enable you to categorize resources and view consolidated billing by applying the same tag to multiple resources and resource groups. [Learn more](#)'. Below this, a note states: 'Note that if you create tags and then change resource settings on other tabs, your tags will be automatically updated.' There is a checkbox labeled 'Key' and a text input field for 'Value'. At the bottom, there are navigation buttons: 'Review + create' (highlighted in blue), 'Previous', and 'Next: Review + create'.

Step 6: Import Data

After reviewing and creating, it will take a few minutes to create the resource. Once the service is created, you can click on the name of the service.



and you will be prompted to this page. Click on Import Data after creating a search service



Step 7: Connect to your data

You will be prompted to choose your data location. Select the location where your data source is stored at. For my guide, I would be choosing Data Lake Gen 2 and select all metadata for data to extract. As my data type is in csv, you will need to select delimited text under parsing mode. The data source can be a folder of files or file or table. Azure Search Service will create a dataset object which duplicates this dataset and will be reading this object when you query the service. Next select choose an existing connection and fill up the container name and blob folder name.


The steps to choose existing connection and creating a container and a folder in the storage account to upload data is shown below.

Import data ...

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

Create and load a search index using data from an existing Azure data source in your current subscription. Azure Cognitive Search

Data Source Azure Data Lake Storage Gen2

 Support for Azure Data Lake Storage Gen2 is in preview and not intended for production use.

Data source name * demodata

Data to extract ⓘ All metadata

Parsing mode Delimited text

First Line Contains Header ⓘ ☒

Delimiter Character ⓘ ,

Connection string * DefaultEndpointsProtocol=https;AccountName=csidema...

[Choose an existing connection](#)

☐ Authenticate using managed identity ⓘ

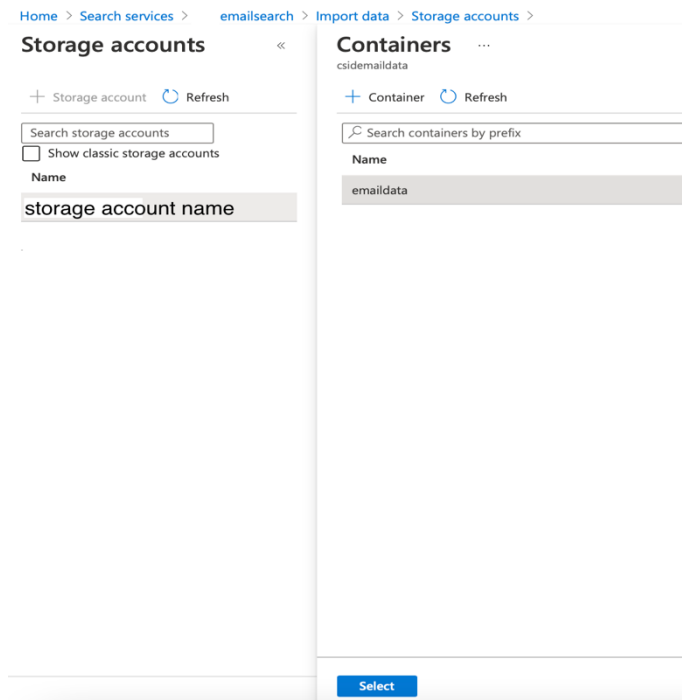
Container name * ⓘ emaildata

Blob folder ⓘ demo/

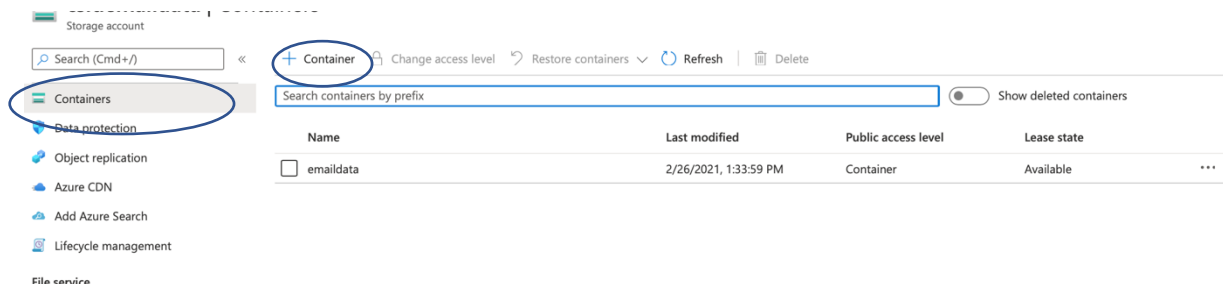
Description (optional)

Next: Add cognitive skills (Optional)

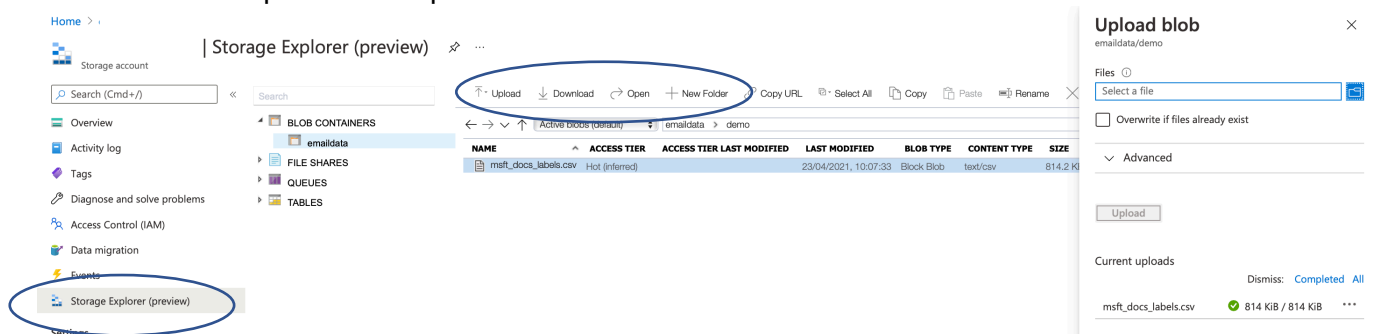
As I choose data lake gen 2 as my data source location, I would need to select my storage account and its following container. If you do not have a storage account created, create a storage account. The images below show the steps to create a container in the storage account and how to create folder and upload files in the container.



After creating a storage account, click on the name. It will prompt u to this page. click on Containers on the left menu-bar and + container icon to create a container in the storage account.



Next, click on storage explorer (preview) and click on Blob containers and the container name you just created. Create a folder and upload the data into the folder that you created in the container. In my guide, my container name is *emaildata* and the folder in this container is *demo/*. So, I would fill this information into search service *container name & blob folder* shown in the first picture in step 7 above.



Step 8: Customize cognitive skills

After connecting to your data, I will be adding a cognitive skill *extract key phrases* under add enrichment section. The other sections are untouched. This step can be skip if you do not wish to add any cognitive skills into the service.

Import data ...

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

⚠ Enrich and extract structure from your documents through cognitive skills using the same AI algorithms that power Cognitive Services. Select the document cracking options and the cognitive skills you want to apply to your document scenarios other than search. [Learn more](#)

▽ Attach Cognitive Services

△ Add enrichments

Run cognitive skills over a source data field to create additional searchable fields. [Learn about additional skills and extensibility here.](#)

Skillset name * ⓘ

azureblob-skillset

☐ Enable OCR and merge all text into **merged_content** field ⓘ

Source data field *

query

Enrichment granularity level ⓘ

Source field (default)

Checked items below require a field name.

<input checked="" type="checkbox"/> Text Cognitive Skills	Parameter	Field name
<input type="checkbox"/> Extract people names		people
<input type="checkbox"/> Extract organization names		organizations
<input type="checkbox"/> Extract location names		locations
<input checked="" type="checkbox"/> Extract key phrases		keyphrases
<input type="checkbox"/> Detect language		language
<input type="checkbox"/> Translate text	Target Language <input type="text" value="English"/>	translated_text
<input type="checkbox"/> Extract personally identifiable information		pii_entities

Previous: Connect to your data


Next: Customize target index

Step 9: Customize target index

In this step, the *index name* will be required when calling the API later on. The column *retrievable* is what will be shown in the output of your search query. The column *searchable* is what internally azure search service use to select the output of your search query.

Import data ...

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

 We provided a default index for you. You can delete the fields you don't need. Everything is editable, but once the index is created, deleting or changing existing fields will require re-indexing your documents.

Index name * 

demo-index





Key * 

AzureSearch_DocumentKey

Suggester name

Search mode 

+ Add field + Add subfield  Delete

Field name	Type	Retrievable	Filterable	Sortable	Facetable	Searchable	Analyzer	Suggester
query	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luc... 	...
url	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luc... 	...
grade	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		...
 AzureSearch_DocumentKey	Edm.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		...
metadata_storage_content_type	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		...
metadata_storage_size	Edm.Int64	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			...
metadata_storage_last_modified	Edm.DateTi...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			...
metadata_storage_content_md5	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		...
metadata_storage_name	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		...
metadata_storage_path	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		...
metadata_storage_file_extension	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		...
keyphrases	Collection(E...	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input checked="" type="checkbox"/>	Standard - Luc... 	...

Previous: Add cognitive skills (Optional)

Next: Create an indexer

Step 10: Create an indexer

The last step of importing data will be to customize the refresh of your indexer. If your data source is in live production and would have new data coming in hourly or daily, you should select *hourly* or *daily* schedule refresh in order for your search service data to stay updated. If your dataset is fixed and will not change any time sooner, you can select *once* schedule refresh.

Import data ...

* Connect to your data

Add cognitive skills (Optional)

* Customize target index

Create an indexer

Indexer

Name *

azureblob-indexer

Schedule ⓘ

Once

Hourly

Daily

Custom

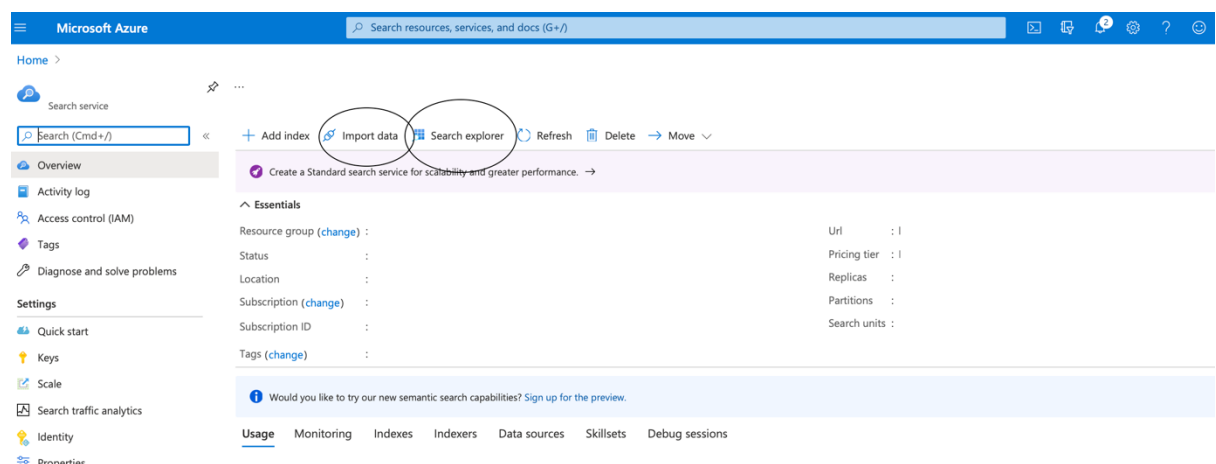
Description

(optional)

Advanced options

Step 11: Search Explorer

After creating the index, you will be brought back to the front page. Click on the search explorer which will produce a UI page that you can type your query.



Step 12: Search Explorer UI

In the Search explorer, you would be able to choose the API version that you would like and enter a query string that you would like to submit. The request url will be generated and results will be shown below.

Search explorer

csidemailsearch

Index demo-index API version 2020-06-30-Pre...

Query string

powershell

Search

Request URL

https://csidemailsearch.search.windows.net/indexes/demo-index/docs?api-version=2020-06-30-Preview&search=powershell

Results

```
1  {
2    "@odata.context": "https://csidemailsearch.search.windows.net/indexes('demo-index')/$metadata#docs(*)",
3    "value": [
4      {
5        "@search.score": 2.5463753,
6        "query": "powershell",
7        "url": "https://docs.microsoft.com/en-us/powershell/scripting/install/installing-powershell",
8        "grade": "7",
9        "AzureSearch_DocumentKey":
10       "aHR0cHM6Ly9jc2lkZW1haWxkYXRhLmJsb2IuY29yZS53aW5kb3dzLm5ldC9lbWpGRhdGEvZGVtb9tc2Z0X2RvY3NfbGFiZWxzLnNzdjs00",
11       "keyphrases": [
12         "powershell"
13       ],
14     },
15     {
16       "@search.score": 2.5463753,
17       "query": "powershell",
18       "url": "https://docs.microsoft.com/en-us/powershell/scripting/install/installing-windows-powershell",
19       "grade": "6",
20       "AzureSearch_DocumentKey":
21       "aHR0cHM6Ly9jc2lkZW1haWxkYXRhLmJsb2IuY29yZS53aW5kb3dzLm5ldC9lbWpGRhdGEvZGVtb9tc2Z0X2RvY3NfbGFiZWxzLnNzdjs10",
22       "keyphrases": [
23         "powershell"
24       ],
25     },
26     {
27       "@search.score": 2.461978,
28       "query": "powershell",
29       "url": "https://docs.microsoft.com/en-us/powershell/",
30       "grade": "9",
31       "AzureSearch_DocumentKey":
32       "aHR0cHM6Ly9jc2lkZW1haWxkYXRhLmJsb2IuY29yZS53aW5kb3dzLm5ldC9lbWpGRhdGEvZGVtb9tc2Z0X2RvY3NfbGFiZWxzLnNzdjsy0",
33       "keyphrases": [
34         "powershell"
35       ],
36     }
37   ]
38 }
```

Step 13: Creating Azure Search Service Config file

In order to call the API from azure search, you will need details of the service. The credentials you would need would be:

1. service_name (name of azure search service)
2. endpoint [the variable url in main page of azure search service shown below]
 - a. starts with https://<service_name>.search.windows.net
3. api_version (choose one from search explorer UI in step 12)
4. api_key (follow picture below)
5. index_name (the name you created in step 9)

The credentials should be stored as a json format.

```
{
  "service_name": "<search service name>",
  "endpoint": "https://<search service name>.search.windows.net",
  "api_version": "2020-06-30-preview",
  "api_key": "<primary admin key>",
  "index_name": "<demo-index>"
}
```

Endpoint location

The screenshot shows the Microsoft Azure portal interface for a Search service. The 'Essentials' section is visible, displaying various properties. The 'Url' property is circled in blue, indicating its location for the endpoint. The 'Url' field is currently empty, with a placeholder text 'Url : '.

API key location

The screenshot shows the Microsoft Azure portal interface for a Search service, specifically the 'Keys' section. The 'Primary admin key' field is circled in blue, indicating its location for the API key. The 'Primary admin key' field is currently empty, with a placeholder text 'Primary admin key'. A 'Copy to clipboard' button is visible next to the field.