

Publish to Azure Search App Bundle

The Publish to Azure Search application performs *content feeds* to a Azure Search of metadata and content of files extracted by Aspire connectors. The feed to the Azure Search can be customized by editing the JSON transformation file provided by the user.

Publish to ElasticSearch App Bundle	
Factory Name	com.searchtechnologies.aspire:app-publish-to-azure-search
Type Flags	job-input
Inputs	AspireObject from a connector's subjob with metadata and content extracted from a specific file/folder.
Outputs	A JSON transformation of the AspireObject sent to the Azure Search's bulk URL.

Configuration

This section lists all configuration parameters available to configure the Publish to Azure Search App Bundle component.

Element	Type	Default	Description
server	String	-	Indicates the name of the service endpoint to use..
index	String	-	Azure Search index where the jobs will be stored
apiVersion	String	-	Azure Search Api version of the REST Api.
apiKey	String	-	Azure Search Api Key used to connect to the REST Api.
aspireToAzureSearchGroovy	String	\${appbundle.home}/config/groovy/aspireToAzureSearchBulk.groovy	Location of the Groovy to transform the job data to a Azure Search feed.
debug	boolean	false	If <i>true</i> it will log debug information from the component

Example Configuration

With Host and Port

```
<application config="com.searchtechnologies.aspire:app-publish-to-elasticsearch">
  <properties>
    <server>corest.search.windows.net</server>
    <index>test</index>
    <apiVersion>2016-09-01</apiVersion>
    <apiKey>encrypted:
9804B36327DAF1E712E4E82301B6A276FCBBA459834EB15F6A94255B6B0BC32B20A3E7262DD2D3D74A6FE5A70A251FCD</apiKey>
    <aspireToAzureSearchGroovy>${appbundle.home}/config/groovy/aspireToAzureSearchBulk.groovy<
/aspireToAzureSearchGroovy>
    <debug>false</debug>
  </properties>
</application>
```

Edit Groovy

The default Groovy transformation file can be found in [aspireToAzureSearchBulk](#)

The default transformation Groovy file provided by the publisher expects metadata as described in [Connector Metadata](#)

Add metadata field

To add a new metadata field extracted by an Aspire Connector add an groovy element inside the **builder.\$object()** that is right after the **builder.flush()**.

```
metadata-name doc.metadatafield
```

Change the document ID

The **id** of a Azure Search document is used to uniquely identify a file in the index. By default, Publish To Azure Search will use the MD5 of the following fields from the Aspire document in order of precedence (if one is missing, then the next will be used):

- fetchUrl
- url
- displayUrl
- id

If you want to change this behavior, edit or create a new Groovy file which has the following element inside **builder.index()**:

```
'_id' value-for-id
```



For more information in how to create a Groovy file transformation please see [JSON Transformation](#)

Connector-specific fields

By default the connector specific fields of the document are not indexed, in order to enable the indexing of connector specific fields you have to add them at the map **connectorSpecificMap** that is at the start of the Groovy file

```
def connectorSpecificMap = [  
    'isContainer': 'is_container'  
]
```

The key of the map entry is the name of the connector specific field as is contained by the document and the value is the name that is going to be used for the indexing. Only the fields specified in this map will be indexed.