

Publish to Elasticsearch How to configure

Follow this procedure to learn how to publish to Elasticsearch.

On this page

- [Step 1. Launch Aspire and Open the Content Source Management Page](#)
- [Step 2. Add a New Content Source](#)
- [Step 3. Add a New Publish to Elasticsearch to the Workflow](#)
 - [Step 3a. Specify Publisher Information](#)

Step 1. Launch Aspire and Open the Content Source Management Page

? Unknown Attachment

Launch Aspire (if it's not already running). See:

- [Launch Control](#)
- Browse to: <http://localhost:50505>. For details on using the Aspire Content Source Management page, please refer to [Admin UI](#)

Step 2. Add a New Content Source

- For this step please follow the step from the Configuration Tutorial of the connector of your choice, please refer to [Connector list](#)

Step 3. Add a New Publish to Elasticsearch to the Workflow

To add a Publish to Elasticsearch drag from the **Publish to Elasticsearch** rule from the *Workflow Library* and drop to the *Workflow Tree* where you want to add it. This will automatically open the Publish to Elasticsearch window for the configuration of the publisher.

Step 3a. Specify Publisher Information

In the Publish to Elasticsearch window, specify the connection information to publish to the index.

1. **Name:** Enter the name of the publisher. (*This name must be unique*).
2. **ElasticSearch URL:** Select how you want to enter the ElasticSearch URL
 - a. **Host and port**
 - **ElasticSearch Host:** Enter the ElasticSearch host.
 - **ElasticSearch Port:** Enter the ElasticSearch port (*9200 by default*)
 - b. **Complete Url**
 - **ElasticSearch URL:** Enter the url for the ElasticSearch bulk index endpoint, it must have this format `<protocol>://<host>:<port>/_bulk`
3. **ElasticSearch Index:** Enter the index to which the jobs are going to be publish.
4. **Groovy Transform:** set to the default parameter to use the default JSON transformation file. To use a custom file, follow the instructions in [JS ON Transformation](#)

HINT: The default file is located at `${appbundle.home}`. if user is using other file in the `/config/xls` folder, the value should be change to `${aspire.home}/config/xls`

The user is responsible for ensuring that the location typed in that box is the correct one of the file being used.

Publish to Elasticsearch

Coordinates: com.searchtechnologies.aspire.app-publish-to-elasticsearch

Name: Publish to ElasticSearch

ElasticSearch URL: Host and port

ElasticSearch Host: localhost

ElasticSearch Port: 9200

ElasticSearch Index: index1

Groovy Transform: S(appbundle.home)/config/groovy/aspireToElasticsearchBulk.groovy

Index dump options

Max Results per request: 1000000

Page Size: 10000

Id field: hits_id

Uri field: hits.fields.url

Timestamp field: hits.fields.submitTime

Additional Options

Debug

Add Cancel

For Elasticsearch 6.3.0 there is a new transformation file that must be used called "aspireToElasticsearch6Bulk.groovy".

5. **Max Results per request:** How many documents can be fetched by the search engine for the same query

6. **Page size:** How many documents to fetch per page
7. **Id field:** Field used to store the id in elasticsearch. Used to compare against the content source audit logs
8. **Url field:** Field used to store the url in elasticsearch
9. **Timestamp field:** the name of the timestamp field holding the index timestamp of every document
10. **Authentication:** Indicates the authentication credentials for the Elasticsearch url, if you are using protocol https, then you must provide authentication credentials
 - a. **None:** use this setting if you do not need to provide credentials
 - b. **Basic:** use this setting if your Elasticsearch url requires authorization credentials.
 - **Username:** enter the username required to access the Elasticsearch url.
 - **Password:** enter the password required to access the Elasticsearch url.
11. **Debug:** Check if you want to run the publisher in debug mode.
12. **Enable Delete by Query:** Check if you want to disable the Delete by Query Feature. **(3.1.1 only)**
Note: The Delete by Query feature does not work on Elasticsearch **5.x** onward
13. Click on the *Add* button.

Once you've clicked **Add**, it will take a moment for Aspire to download all of the necessary components (the Jar files) from the Maven repository and load them into Aspire. Once that's done, the publisher will appear in the Workflow Tree.



For details on using the Workflow section, please refer to the [Workflow](#) introduction.