

**Δquaforest**

Searchlight Tagger  
**Release Notes**



# Searchlight Tagger

## Release Notes



Version 1.2  
April 2022

# Content

1	VERSION 1.2.2204.08.....	4
1.1	Bug Fixes.....	4
2	VERSION 1.2.2201.18.....	4
2.1	Bug Fixes.....	4
3	VERSION 1.2.2111.30.....	4
3.1	Bug Fixes.....	4
4	VERSION 1.2.2111.22.....	4
4.1	Enhancements.....	4
4.2	Bug Fixes.....	4
5	VERSION 1.1.200615.....	4
5.1	Enhancements.....	4
6	VERSION 1.1.200511.....	4
6.1	Upgrading from previous versions.....	4
6.2	Enhancements.....	4
6.3	Bug Fixes.....	6

# 1 Version 1.2.2204.08

## 1.1 Bug Fixes

- [AT-57] “This implementation is not part of the Windows Platform FIPS validated cryptographic algorithms” fix
- [AT-58] Tagger fails to tag whole SharePoint On-Premises site collection when running in VMWare

# 2 Version 1.2.2201.18

## 2.1 Bug Fixes

- [AT-55] Root SharePoint URLs are not processed
- [AT-56] ‘UnlimitedLengthInDocumentLibrary’ is not initialized exception

# 3 Version 1.2.2111.30

## 3.1 Bug Fixes

- [AT-54] Delete Cert keys in ProgramData\Microsoft\Crypto\RSA\MachineKeys after use

# 4 Version 1.2.2111.22

## 4.1 Enhancements

- [AT-46] Add support for multi-line text SharePoint field
- [AT-48] Add option to tag terms with the highest frequency in a document  
You will need to set `tagItemsWithTheHighestFrequencyFirst` setting in `Tagger.config` to `true` and restart the Tagger service.
- [AT-52] Add support for SharePoint Online interactive authentication
- [AT-53] Add support for SharePoint app-only authentication

## 4.2 Bug Fixes

- [AT-45] Filter documents by creation date does not work
- [AT-49] Terms with ampersand are not tagged
- [AT-50] Missing results when a term only appears once in a document and `tagItemsWithTheHighestFrequencyFirst` is set to `true` and/or `minWordFrequency` is set to `1`
- [AT-51] Tagger scan barcodes that have not been selected

# 5 Version 1.1.200615

## 5.1 Enhancements

- [AT-39] Replaced the barcode engine with a newer one, which significantly improves barcode recognition.

# 6 Version 1.1.200511

## 6.1 Upgrading from previous versions

We highly recommend users creating a backup of their database before attempting an upgrade. The database can be found at `[installation path]\data\Tagger.db`. The database can be very large, depending on the number of runs.

## 6.2 Enhancements

- [AT-13] Added support for modern authentication. The solution involves using Azure Active Directory App-Only authentication.

In summary, you need:

1. Create a self-signed certificate
2. Register a dummy Web App in Azure (no coding involved)
3. Give the Web App permissions to access the SharePoint tenant
4. Connect the certificate created in step 1 to the Web App

Full instructions can be found in the following link:

<https://docs.microsoft.com/en-us/sharepoint/dev/solution-guidance/security-apponly-azuread>

- [AT-20] Updated Microsoft NLP service to use the latest Microsoft Cognitive services. The previous version could only be used to retrieve keywords from documents using the “Keywords” NLP entity in Tagger. Now entities can also be retrieved.

Additionally, new config settings have been added in the `Tagger.config` file:

- `microsoftNLPServiceEndpoint`  
If a custom Azure Cognitive Service have been created, update this setting with the endpoint.
- `microsoftNLPServiceLanguage`  
To set the language of the documents being processed

- Retries for SharePoint queries are now only performed for specific `HttpStatusCodes` by default, and is defined in the `httpStatusCodesToRetry` setting in the `Tagger.config` file. The number of times to retry and the amount of time to wait before each retry is specified in the `webRequestRetries` setting.

See the comments in the `Tagger.config` file for more information:

```
<!-- Start Retries -->
<!-- value="x,y" (where x is the maximum number of retries to perform and y is the time (in millisecond)
<add key="databaseRetries" value="5,1000" />
<add key="webRequestRetries" value="5,5000" /> <!-- For SharePoint and Office 365 libraries, retry will
<add key="httpStatusCodesToRetry" value="503" /> <!-- Add specific HttpStatusCodes to retry (separated
<!-- End Retries -->
```

**NOTE:** The Searchlight Tagger service must be restarted after making the change.

- Tagger was unable to connect to SharePoint servers where only TLS 1.1 and/or 1.2 was enabled. The `Tagger.config` file now has additional configuration options to enable the various cryptographic protocols supported by SharePoint.
- Accessing servers with invalid certificates - If Tagger tries to access a site that has an invalid SSL certificate, it will fail with an error message like “*Could not establish trust relationship for the SSL/TLS secure channel.*”

There are 2 settings in the `Tagger.config` file that can be used to fix this.

```
<!--
To enable Searchlight to process sites that have invalid certificates or certificates with errors,
enter a comma separated list of all thumbprints of certificates that you recognize
in the setting below to make Searchlight ignore them and continue processing. Alternatively, set
"ignoreAllCertificateErrors" to "true" (use at your own risk!).
-->
<add key="recognizedCertificateThumbprints" value="" />
<add key="ignoreAllCertificateErrors" value="false" />
```

- `recognizedCertificateThumbprints`  
Use this setting to add the thumbprint/fingerprint (SHA-1) of the certificate that is causing the issue. More thumbprints can be added by separating each one with a comma. This is recommended way as it instructs Tagger to only ignore errors of recognized certificates.

Check the **Troubleshooting Guide** to see how to retrieve the thumbprint or fingerprint of the certificate.

- `ignoreAllCertificateErrors`  
Set this setting to true to ignore all errors for all certificates.

**NOTE:** The Tagger UI and service must be restarted after making the change.

- SharePoint CSOM version has been updated from version 15 to 16
- Barcode reader version had been updated from version 2.2 to 2.3

## 6.3 Bug Fixes

- [AT-11] Fixed 'Retain Modified By'
- [AT-16] Read-only SharePoint columns were being ignored if used in "Filter Documents by Regular Expression"
- [AT-21] Updated "Tika" to the latest version to fix issue with getting text from certain documents
- [AT-25] Fix for issue where new locations could not be added to existing jobs
- [AT-26] Fixed decoding PDF417 barcode
- [AT-31] Fixed issue with the scheduler where Tagger would immediately run all scheduled jobs after a reboot or after restarting the service, regardless of time since last run.
- Empty Library Tab - the **Job** tab will now be greyed out if no jobs have yet been created. This tab was navigable before and users were using it to add locations to a non-existent job, thus causing errors. New jobs should be set up by clicking the **Add new job** button on the Dashboard instead.