SLIM Applications

BulkMetadataManager 2.9.1 - Release Notes

2026/10/07

1. Introduction 5
2. About BulkMetadataManager 5
2.1 Extract metadata 5
2.2 Detect duplicates 5
3. Supported Environments 6
3.1 Supported Operating Systems 6
3.2 Supported SharePoint Versions 6
4. Releases 7
4.1 Feature and bug release (2.9.1) 7
4.2 Feature and bug release (2.9.0) 7
4.3 Feature and bug release (2.8.1) 9
4.4 Feature release (2.8.0) 10
4.5 Bug release (2.7.2) 10
4.6 Bug release (2.7.1) 10
4.7 Feature and bug release (2.7.0) 10
4.8 Bug release (2.6.4) 10
4.9 Feature release (2.6.3) 10
4.10 Feature release (2.6.2) 11
4.11 Feature release (2.6.1) 11
4.12 Feature release (2.6.0) 11
4.13 Bug release (2.5.5) 11
4.14 Feature release (2.5.4) 11
4.15 Feature release (2.5.3) 12
4.16 Feature release (2.5.2) 12
4.17 Bug release (2.5.1) 12
4.18 Feature and bug release (2.5.0) 12
4.19 Bug release (2.4.6) 13
4.20 Bug release (2.4.5) 13
4.21 Bug release (2.4.4) 13
4.22 Bug release (2.4.3) 13

4.23 Feature and bug release (2.4.2) 13
4.24 Feature release (2.4.1) 14
4.25 Feature and bug release (2.4.0) 14
4.26 Bug release (2.3.3) 15
4.27 Feature and bug release (2.3.2) 15
4.28 Feature and bug release (2.3.1) 15
4.29 Feature and bug release (2.3.0) 16
4.30 Feature and bug release (2.2.2) 16
4.31 Feature and bug release (2.2.1) 16
4.32 Feature and bug release (2.2.0) 16
4.33 Feature and bug release (2.1.4) 17
4.34 Bug release (2.1.3) 17
4.35 Feature and bug release (2.1.2) 18
4.36 Feature release (2.1.1) 18
4.37 Feature release (2.1.0) 19
4.38 Bug release (2.0.7) 19
4.39 Bug release (2.0.6) 19
4.40 Feature release (2.0.5) 19
4.41 Bug release (2.0.4) 19
4.42 Feature and bug fix release (2.0.3) 19
4.43 Bug fix release (2.0.2) 20
4.44 Bug fix release (2.0.1) 20
4.45 Feature release (2.0.0) 20
4.46 Bug fix release (1.0.13) 20
4.47 Bug fix release (1.0.11) 20
4.48 Bug fix release (1.0.10) 21
4.49 Bug fix release (1.0.9) 21
4.50 Bug fix release (1.0.8) 21
4.51 Bug fix release (1.0.7) 21
4.52 Feature release (1.0.6) 21
4.53 Feature release (1.0.5) 21

4.54 Feature release (1.0.4) 21
4.55 Initial release (1.0.0) 21
5. Known Issues 23

1. Introduction

This document describes the release notes for BulkMetadataManager 2.9.1 and provides the latest information on the SharePoint application.

2. About BulkMetadataManager

BulkMetadataManager is an application that uses REST and/or Graph API to interact with SharePoint. The application supports different scenarios.

2.1 Extract metadata

BulkMetadataManager can extract properties from documents stored in SharePoint and capture the values into SharePoint columns. This allows for optimal use of the metadata for searching, filtering, sorting etc., … and this results in increased user productivity and adoption. It also allows for properly managing the document life cycle (records management).

The documents that need to be processed are selected using a CAML query. The application is optimized for performance by using parallel requests, caching, and batch requests. It also caters for SharePoint throttling (429/503 responses) and the application can handle large lists exceeding the list view threshold. The use of a CAML query allows for using a condition to avoid processing the same document multiple times.

The “WhatIf” option allows for running the application without the final step to add the metadata into SharePoint columns. The extracted metadata is saved into a csv for further review and analysis.

2.2 Detect duplicates

BulkMetadataManager can also detect duplicate documents in SharePoint. Exact matches are identified using SHA256 for file types that are not affected by property demotion (such as pdf, jpg, mp4, and txt). For emails, the email’s Message-ID is used. For files affected by property demotion (for example docx and xlsx), SHA256 is calculated over the document content rather than the full file binary.

Downloads are limited through an efficient pre-filter, so not every document needs to be retrieved. Results are stored in a local database for power users and can be exported to CSV for broader access and reporting.

The target audience for this document is SharePoint architects, SharePoint administrators and site collection administrators.

3. Supported Environments

This section provides information on the supported environments. The application is a node.js application that is packaged as an exe and runs on a separate system, it does not require installation of software on SharePoint. The installation instructions are described in the BulkMetadataManager Administrator Guide.

3.1 Supported Operating Systems

BulkMetadataManager is supported on the following operating systems.

Operating system	Supported	Notes
Windows 11	✓	
Windows Server 2016	✓	
Windows Server 2019	✓	
Windows Server 2022	✓	
Windows Server 2025	✓	

Operating systems not listed are not supported.

3.2 Supported SharePoint Versions

The following SharePoint versions are supported.

Versions	Supported	Supported Authentication methods
SharePoint Online (Office 365)	✓	Entra ID Certificate-based Authentication
SharePoint Server SE	✓	NTLMv2
SharePoint 2019	✓	NTLMv2

SharePoint versions not listed have not been tested and are not supported.

4. Releases

This section provides the details of each release.

4.1 Feature and bug release (2.9.1)

1. Support for processing different file types with different content types in a single operation. Previously this required a separate run for each content type. Add a line to BulkMetadataManagerProperties.txt, for example:
   DocumentContentTypeName=Email(msg/eml),BinaryOfficeFile(doc/xls/ppt)

   This maps msg and eml files to the Email content type, and binary Office files (doc, xls, ppt) to the BinaryOfficeFile content type.
   Note: a separate WhatIf log file is created for each content type, because the columns differ.
2. Support for extracting tokens from Office attachments without creating a temporary file on the system running BulkMetadataManager.
3. Fixed a bug where the email body in the WhatIf log file broke Excel parsing.
   Fields that contain line breaks are now quoted and truncated to 256 characters, so the row stays on one line.
4. Fixed a bug handling metadata update responses with “Invalid text value”.
5. Fixed a bug handling metadata update responses containing the generic error code -2146232832.
6. Minor changes:
   - omit the start time in the log file to allow for easy parsing of the csv file.
   - minor textual change in the error message when the extraction of tokens from an email attachment fails.

4.2 Feature and bug release (2.9.0)

1. Support for detecting and deleting duplicate documents in SharePoint.
   To detect duplicate documents with BulkMetadataManager, set the following property in BulkMetadataManager.ini:
   Mode=Duplicates
   That runs the full process: inventory the documents, store the results in a local database, then compare candidates and list duplicates.
   It is also possible to run the phases on their own (inventory only, or duplicate check only).
   To delete files, set Mode=Delete. BulkMetadataManager reads the document URLs from a csv file and deletes them in batches per site. By default, files go to the recycle bin; an optional property (UseRecycleBin=false) permanently deletes the documents.

Mode	What it does
Duplicates	Inventory + candidate check (full run)
DuplicatesInventory	Inventory only (write results to the database)
DuplicatesCheck	Candidate duplicate check only (uses the existing inventory)

Delete	Delete the documents specified in a csv file

If Mode is omitted, the default mode is ExtractMetadata which extracts metadata from SharePoint documents and captures the values into SharePoint columns.
Note: the BulkMetadataManager app registration in Entra ID needs to have the SharePoint / Sites.ReadWrite.All permission (plus admin consent granted) to delete documents.

2. Support for detecting all sites using Graph API instead of search (only for SharePoint Online)
   BulkMetadataManager detects all sites when the BulkMetadataManager.ini files contains *Sites=**
   Using search skips sites with "Allow this site to appear in search results = No" and anything not crawled yet.
   To use Graph API, add the following property to the BulkMetadataManager.ini file
   SiteDiscoveryMethod=Graph
   The default mode is to use search.
   Note: the BulkMetadataManager app registration in Entra ID needs to have the Microsoft Graph / Sites.Read.All permission (plus admin consent granted)
3. Added support for the new replacement tag <SentDateSite> for emails.
   This tag inserts the email's Sent Date converted to the site's local time zone.
   To use it, update the EmailNameSyntax property with the new tag:
   EmailNameSyntax=<Subject>_<SentDateSite>

   Note 1: The date format automatically adapts to the site's DateFormat setting:
   - DateFormat = 0 (US): MM-DD-YYYY_hh-mm
   - DateFormat = 1 (International / Default): DD-MM-YYYY_hh-mm
   - DateFormat = 2: YYYY-MM-DD_hh-mm
   Note 2: The existing <SentDate> tag continues to use the Sent Date in UTC using YYYYMMDD_hhmmss.
4. Implemented a robust detection algorithm for identifying encrypted pdf attachments with improved accuracy and reliability.
   The property hasEncryptedAttachmentYesNo (type Yes/No) captures if an email contains one or more encrypted attachments. Add the following line to the BulkMetadataManagerProperties.txt file:
   hasEncryptedAttachmentYesNo=Email:Email_x003A_hasEncryptedAttachment
   Note: replace the content type (Email) and the internal SharePoint column name by values appropriate for your environment.
5. Fixed incomplete retrieval of sites associated with a hub.
   The Search query now uses DepartmentId instead of RelatedHubSites, which did not return all associated sites.

4.3 Feature and bug release (2.8.1)

1. Enhanced email attachment metadata extraction:
   - extract SHA-256 hashes of attachments (for deduplication and integrity)
   - extract file sizes of attachments
   - extract the full text content of attachments for improved search and analysis and use by Microsoft 365 Copilot.
2. Changed format for attachment content mapped using the property AttachmentContents.
   Previously
   [[attachmentname1]]
   <unique words in attachment1>
   [[attachmentname2]]
   <unique words in attachment2>

   New
   attachmentname1
   Contents: < unique words in attachment1>

   attachmentname2
   Contents: < unique words in attachment2>

   The rationale for this format change is to keep the format consistent with the other attachment related properties.
3. The words in the AttachmentContents column are converted to lowercase to ensure consistent and reliable search results.
   Note: The full text content stored in the AttachmentFullText column preserves the original case for better readability.
4. Fix bug extracting To, CC and BCC properties from eml files.
5. Fix bug extracting unique words from files in zip attachments.
6. Fix bug that resulted in double counting the msg attachments in msg emails.
7. Fix bug for capturing content from Office attachments into the AttachmentContents and/or AttachmentsFullText fields:
   - avoid bug that results in unnecessary spaces in the extracted text (docx only)
   - include the content from word/comments.xml (docx only)
   - include the content from ppt/notesSlides/notesSlide*.xml (pptx only)
   - sort the slides (pptx only)
   - include content from xl/comments1.xml (xlsx only)
   - include content from visio/documents.xml, visio/shapes/shape*.xml and visio/masters/masters.xml (vsdx only)
8. Fix bug with the conversion of dates extracted from binary Office files resulting in dates with incorrect months.

4.4 Feature release (2.8.0)

1. Support for extracting properties (e.g., Subject, Create Time/Date, Keywords, …) from binary Office files (doc, xls, and ppt) stored in SharePoint.
2. Support for extracting the Thread ID from msg and eml emails present in SharePoint.

4.5 Bug release (2.7.2)

1. Fix bug throwing an error updating metadata when the original ModifiedBy account has been deleted.

4.6 Bug release (2.7.1)

1. Fix bug affecting the extraction of properties from mp4 and mov files.

4.7 Feature and bug release (2.7.0)

1. Support for authentication to SharePoint Online using Microsoft Entra ID app-only authentication with certificates.
   This authentication mechanism is required because ACS-based Client ID and Secret authentication will no longer be supported after April 2, 2026. See section 3.1.1 in the BulkMetadataManager Administrator Guide 2.7.
2. Fix bug showing an authentication error when trying to determine the date format for <tenant URL>/contentstorage/ when using search to list all sites.
3. Improved structure of the console log output.
4. The static properties file is now the only supported method for configuring BulkMetadataManager. Support for the following configuration methods has been removed as they were not used by any customers:
   - SharePoint tenant properties (created and managed via PowerShell)
   - Properties file stored in the SharePoint site under /sites/BulkMetadataManager

4.8 Bug release (2.6.4)

1. Fix bug with the error reporting when the metadata extraction from Office files fails.
2. Fix bug preventing the extraction of unique words from pdf attachments in eml emails.
3. Fix bug preventing the extraction of metadata from eml emails with very large header sections (e.g., 200 kB) containing the proprietary and internal header X-MS-Exchange-Forest-IndexAgent.
4. Fix bug extracting the CC from eml emails.

4.9 Feature release (2.6.3)

1. Support for extracting the References, In-Reply-To and Normalized Subject properties from emails. To map these properties to SharePoint columns, add the following lines to the

BulkMetadataManagerProperties.txt file
References=Email:Email_x003A_References
ConversationGUID=Email:Email_x003A_ConversationGUID
NormalizedSubject=Email:Email_x003A_NormalizedSubject
where the Email:Email_x003A_xyz are examples of the email content type and internal SharePoint column name.

Note: this feature allows for extracting the thread details from existing SharePoint emails. The email threads can then be presented in the browser using the Email Manager App.

4.10 Feature release (2.6.2)

1. Enables TLS/SSL certificate validation.
   Self-signed certificates will throw an error and prevent execution of the programme.
2. Support for extracting the Conversation Index (tag 0x00710102) and the Conversation GUID from msg files.
   The Conversation Index is a unique identifier of an email used by email systems to track the threading and hierarchy of related messages within a conversation.
   The Conversation GUID serves as a consistent identifier for all emails within a thread, enabling reliable grouping and organization of messages that belong to the same conversation.
   Note: the Conversation property cannot be used to reliably construct an email thread.

4.11 Feature release (2.6.1)

1. Disables TLS/SSL certificate validation.
   This allows for using BulkMetadataManager with SharePoint farms using self-signed certificates.

4.12 Feature release (2.6.0)

1. Support for SharePoint Server SE and SharePoint 2019 using NTLMv2 authentication.

4.13 Bug release (2.5.5)

1. Fix very infrequent bug with showing msg email properties as Chinese characters. This only occurred for email properties having a specific length (e.g., 192 B) and specific internal storage structure (i.e., duplicate sector ID for the mini-FAT and FAT).

4.14 Feature release (2.5.4)

1. Support for configuring the number of items used in the dynamic CAML query when the library contains 5000+ items. The default is 5000. Use a smaller value like 1000 or 2000 to avoid LVT errors when processing large lists. For example,

LVTChunkSize=2000

4.15 Feature release (2.5.3)

1. Support for vertical scaling using ini files from a separate secure location.
   The ini files must be present in the secure location with the naming syntax:
   BulkMetadataManager.ini
   BulkMetadataManager_2.ini
   BulkMetadataManager_3.ini
   etc…
   To configure BulkMetadataManager to use a specific ini file (e.g., BulkMetadataManager_2.ini) use
   BulkMetadataManager --instance=2
   Note: the instance parameter is optional. Starting BulkMetadataManager without parameter will use the BulkMetadataManager.ini file

4.16 Feature release (2.5.2)

1. Support for vertical scaling by running multiple instances of BulkMetadataManager on a single system, with configuration via an .ini file located in a subfolder of each instance.
2. Support for configuring the delay between successive requests to retrieve item details from large lists that exceed the list view threshold (5,000 items). Add the following line to the BulkMetadataManager.ini file with your custom wait time in ms:
   WaitTimeListViewThresholdRequests=500
   The default value is 200 ms.
3. Support for configuring the delay between successive download requests. Add the following line to the BulkMetadataManager.ini file with your custom wait time in ms:
   WaitTimeFileDownloads=500
   The default value is 200 ms.
4. Support for configuring the delay between successive requests to update metadata. Add the following line to the BulkMetadataManager.ini file with your custom wait time in ms:
   WaitTimeUpdateMetadata=500
   The default value is 200 ms.

4.17 Bug release (2.5.1)

1. Fix bug with attachment names incorrectly containing Chinese characters.

4.18 Feature and bug release (2.5.0)

1. Support for SharePoint columns of type “multiple lines of text” configured with unlimited length.
2. Fix bug with extracting email properties (e.g., Subject) having a length > 64 bytes where the content is stored in non-consecutive 64-byte sectors. This bug may have resulted in the presence of Chinese characters in extracted properties. Uploading these extracted properties may result in

errors about invalid characters or exceeding the maximum size of the mapped SharePoint columns.

4.19 Bug release (2.4.6)

1. Fix bug with extracting GPS coordinates in decimal format from jpg files.

4.20 Bug release (2.4.5)

1. Fix bug with extracting properties from small mp4 files (e.g., 1 MB) resulting in "RangeError20Offset is outside the bounds of the dataview".
2. Fix bug showing an extracted date as 1 January 1904 when the mp4/mov file does not contain a date value.
3. Add timeout of 60 s for the document download request.

4.21 Bug release (2.4.4)

1. Fix bug with extracting GPS coordinates from files without GPS coordinates resulting in "Invalid number value" errors.
2. Skip extracted dates containing 00:00:00 as yyyy:mm:dd.

4.22 Bug release (2.4.3)

1. Fix bug with the calculation of GPSDateTimeStampUTC when only GPSTimeStamp is present and GPSDateStamp is missing.

4.23 Feature and bug release (2.4.2)

1. Changed the precedence order (high to low) for GPSDateTimeStampUTC, DateTimeUTC and DateTime.
   This allows for mapping multiple date properties to a single column of type date and the property with the highest precedence and non-zero value will be captured into the SharePoint column.
   NB: the order of the properties in the BulkMetadataManagerProperties.txt file does not matter.
2. Fix bug affecting sites using SharePoint's DateFormat=1 where modified date are retrieved as dd/mm/yyyy hh:mm.
3. Fix bug with the generated GPSDateTimeStampUTC using single digit hour or minute values.
4. Minor changes:
   - limit the value of GPS altitude to 2 decimals

4.24 Feature release (2.4.1)

1. Support for exposing file properties of type date in UTC. Specifically,
   EXIF dates:
   DateTimeUTC
   DateTimeOriginalUTC
   DateTimeDigitizedUTC
   GPSDateTimeStampUTC

   IPTC dates:
   Date Created UTC
   Time Created UTC

   XMP dates:
   xmp:ModifyDateUTC
   xmp:MetadataDateUTC
   xmp:CreateDateUTC
   xap:ModifyDateUTC
   xap:MetadataDateUTC
   xap:CreateDateUTC
   photoshop:DateCreatedUTC

   The above date fields ending with UTC contain the original date values converted to UTC using the time zone information (like +02:00) or the offset time property in the file. The UTC fields are only available if the original date contains time zone information (like +02:00) or when the offset time is available. GPSDateTimeStampUTC combines the GPSDateStamp and GPSTimeStamp values if both values are present.
2. Minor changes:
   - added 3x additional EXIF labels:
   OriginalDefaultFinalSize, OriginalBestQualityFinalSize and OriginalDefaultCropSize
   - warnings are only shown in the console log when Debug=true in the properties file.
   - show filename when an “Offset is outside the bounds of the DataView” error occurs

4.25 Feature and bug release (2.4.0)

1. Support for jfif files.
   jfif files use the existing grouping “JPG” and there is no need to add additional mappings.
2. Support for tiff files.
   Use the grouping “TIF” for mapping tiff properties.
3. Support for CR2 files.
   Use the grouping “TIF” for mapping cr2 properties. For example,
   TIF:GPSCoordinatesDD=Video:GeoLoc
4. Fix bug extracting the EXIF GPSVersionID value without a trailing . character like “2.3.0.0”.

5. Fix intermittent bug resulting in missing document properties in the WhatIf log file.

4.26 Bug release (2.3.3)

1. Fix bug extracting the EXIF GPSVersionID value.
   The value read (e.g., 2300) is extracted as 2.3.0.0.

4.27 Feature and bug release (2.3.2)

1. Support for setting static SharePoint column values. For example, to set a client name for a set of documents add the following line to the BulkMetadataManagerProperties.txt file
   STATIC_1=Matter Document Set:Client=SLIM Applications B.V.
   where STATIC_1 is one of the 15 static properties that can be mapped, "Matter Document Set" is the name of content type and "Client" is the internal SharePoint column name. The value after the last = character will be applied to all documents selected using the CAML query.
   The following column types are supported: Single line of text, Multiples lines of text, Date and Time, Lookup, Choice, Currency, Number, Person or Group, YesNo, Managed Metadata and Location:Coordinates. See section 3.2.8 in the Administrator Guide (version 2.3) for further information.
2. Include the "ModifiedBy" and "Size" values into the WhatIf log file.
3. Fix bug with setting metadata resulting in "Cannot read properties of undefined (reading 'ServerRelativeUrl')" errors.
4. Fix bug with extracting GPS Latitude and GPS Longitude values as .000 and 0.000
5. Fix bug with the date/time stamp used in the naming of the log files.
   Instead of naming the WhatIf log file like "whatif_2062_129.csv" create a log file named "whatif_20241028_165935.csv".
6. Remove CR (0x0D) and LF (0x0A) characters in single line of text fields.
   This circumvents errors renaming emails when the email's subject contains a LF character.
7. Cater for 503 responses ("Service Unavailable") when downloading documents.

4.28 Feature and bug release (2.3.1)

1. Support for extracting properties from mov files recorded using an iPhone .
2. Support for extracting GPS coordinates from png files.
3. Optimization of the metadata extraction from large video files (4 MB+).
4. Support for using local BulkMetadataProperties.txt (i.e. in the same location as the BulkMetadataManager exe file)
   This simplifies the deployment process by not having to create a separate support site for BulkMetadataManager and avoids the creation of separate client/secret IDs for the support site.
   BulkMetadataManager will automatically use the local BulkMetadataProperties.txt if the tenant credentials AND support site credentials are not specified.

4.29 Feature and bug release (2.3.0)

1. Support for extracting properties from jpg, dng, heic and heif images and capturing the values into SharePoint columns.
2. Support for extracting properties from mp4, mov and wmv videos and capturing the values into SharePoint columns.
3. Support for Picture libraries (template ID 109).
4. Cater for emails containing encrypted zip attachments.
   The names of the files in the encrypted zip attachment are captured in the column with the attachment names but the contents of the files is not extracted.

4.30 Feature and bug release (2.2.2)

1. Support for hybrid content types where the content type is defined at site level, deployed to a list and then additional list columns are added.
2. Fix bug with priorities when multiple content types are configured via the DocumentContentTypeName property and more than 1 of the content types are present in a single library. The bug resulted in the order of the content types not being respected.
   For example, if a document library contains 2 content types (e.g., "*Matter Document Set*" and "*Document*") and the *DocumentContentTypeName=Matter Document Set,Document* the bug would have resulted in using the content type "*Document*" being used whereas the "*Matter Document Set*" has a higher priority.

4.31 Feature and bug release (2.2.1)

1. Support for capturing the version of BulkMetadataManager into a SharePoint column.
   To use this feature add the following line to the BulkMetadataManager
   BulkMetadataManagerVersion=<ContentType name>:<Internal SharePoint column name>
   For example,
   BulkMetadataManagerVersion=Email:Email_x003A_UpdatedByBMMVersion
   The column needs to be created a single line of text.
2. Fix bug in BulkMetadataManager 2.2.0 that fails to update the content type.

4.32 Feature and bug release (2.2.0)

1. Support automatic selection of all sites associated with a hub site.
   In the BulkMetadataManager.ini file set the GUID of the hub site as value for the "Sites" property. BulkMetadataManager will process all sites that are associated with the hub site (including nested hub sites).
   Note #1. Make sure the SiteClientID or tenantClientID have 'write' access to each of the associated sites. This is not necessary when using the tenantClientID.
   Note #2. Separate credentials are required when using the hub sites' GUID because the associated sites are retrieved via search. This is described in section 3.1.4 of the Administrator Guide

2. Support for using an ini file from a different location (e.g., C:\Apps\BulkMetadataManager\ini)
   This allows for placing the ini file in a separate directory with controlled access.
   Note. To use this feature provide us with the full path to the ini file.
3. Support for specifying the locations for the log files and the temp folder used for processing Office files. *This is controlled by the following properties in the ini file:*
   LogFilePath=C:/Apps/mylogs
   TempFilePath=C:/Apps/mytemp
   The default behavior is to use the subfolders logs and temp in the current location of the BulkMetadataManager exe file.
4. Support for using multiple content types.
   Note #1. The content types need to be specified in the BulkMetadataManagerProperties.txt file in the property DocumentContentTypeName. The values are separated by commas.
   For example,
   DocumentContentTypeName=Email,EmailExt
   Note #2. Each of the mapped metadata fields need to be prepended by the content type name plus a : character as separator:
   *Subject=**Email:**Email_x003A_Subject*
   *Subject=**EmailExt:**EmailExt_x003A_Subject*
   Each available metadata property can be mapped to multiple content types.
   Note #3. Renaming documents using the DocumentNameSyntaxBulkMetadataManager property is independent from the email content types used. The replacement tags refer to the BulkMetadataManager properties like <Subject> and not the internal name of the SharePoint column.
5. Minor changes:
   - support for using the replacement tags <CurrentDate> and <CurrentTime> in the property DocumentNameSyntaxBulkMetadataManager

4.33 Feature and bug release (2.1.4)

1. Cater for http 503 responses when retrieving documents or updating the metadata.
2. Handle processing of corrupt zip attachments without crashing the application.
3. Minor changes:
   - include the SharePoint document ID in the log and Whatif log files
   - allow for configurable separator characters for the log file and the Whatif log file. For example, add the following lines to the ini file to use this feature:
   SeparatorCharacterLog=#
   SeparatorCharacterWhatIF=~
   The default is to use a tab.

4.34 Bug release (2.1.3)

1. Add missing display labels such as “truncated to” and “characters”.

4.35 Feature and bug release (2.1.2)

1. Fix bug with saving incorrect date values into SharePoint columns when email metadata extraction is used and the SharePoint site uses a time zone other than UTC.
   In previous versions the extracted date values (e.g. Sent Date, Received date, ...) were uploaded in UTC without considering the time zone of the site.
2. Fix bug affecting the extraction of emails with multiple pdf or Office attachments resulting in *ERR_UNHANDLED_REJECTION* an error.
3. Include the pdf properties (e.g., keywords) into the list with extracted words.
4. Minor changes:
   - extend error handling for the metadata update response
   - automatically create folders for logs and temp if they do not exists
   - use subfolder (temp) for temporary saving Office attachments
   - remove multiple dots (e.g., ……) from the extracted tokens of attachments
   - insert the text “pdf attachment without tokens" into the extracted “Attachment Content” column for pdf attachments without unique words (e.g., scans)
   - show “Average time per document” in ms
   - when a batch request to update metadata fails log the details of the affected emails
   - limit the maximum number of characters extracted from attachments to 100k characters

4.36 Feature release (2.1.1)

1. Support for extracting properties defined as list columns.
   BulkMetadataManager 2.1.0 and earlier versions only extracted properties into columns defined at site level. This limitation has been removed. Add the following line to the BulkMetadataManager.ini file to use columns defined at list level:
   ContentTypeSource=list
   The default value is site.
2. Support for combining the username plus smtp email address. For example, this will show Paul de Jong <paul@slimapplications.com>
   Combining fields is supported for the From, To, CC and BCC fields. To use the feature, add one or more of the following lines to the BulkMetadataManager.ini file (or update the tenant properties):
   AppendAddressFrom=FromAddress
   AppendAddressTo=ToAddress
   AppendAddressCC=CCAddress
   AppendAddressBCC=BCCAddress
   Note: copy the lines as is.
3. Use of a separate support site (https://<tenant>.sharepoint.com/sites/bulkmetadatamanager) for storing the metadata mapping file. The file “BulkMetadataManagerProperties.txt” needs to be uploaded to the support site’s siteassets library.

4.37 Feature release (2.1.0)

1. Support for indexing attachments in msg emails (optional).
   SharePoint Search does not crawl all attachments in msg emails. This results in blind spots in the search index that affects users and security/compliance officers using search or advanced tools like Purview to find content.
   BulkMetadataManager extracts unique words from pdf, docx, xlsx, pptx, vsdx, csv and txt email attachments and captures the values into a SharePoint column. This feature also caters for the above file types in (nested) zip attachments.
   Searching for a word in an email attachment will then include the msg email in the search result list. This feature allows users and security/compliance officers to search across all existing email content in SharePoint.
2. Support for extracting the number of attachments (optional).
   This number includes the files present in zip attachments and msg attachments. It excludes inline images like image001.png that are typically used for logos and small icons.
3. Support for extracting the attachment names (optional).
   The inline image names like image001.png are excluded.
4. Simplified deployment by packaging the BulkMetadataManager application, node.js and required modules as a single exe file.

4.38 Bug release (2.0.7)

1. Fix bug affecting the extraction of Create and Modify dates from email, Office and PDF files.

4.39 Bug release (2.0.6)

1. Fix bug affecting the generation of the log file.

4.40 Feature release (2.0.5)

1. Support for extracting Create and Modify dates from email, Office and PDF files and using the values as SharePoint creation and modification dates.

4.41 Bug release (2.0.4)

1. Fix bug affecting the metadata extraction from multiple document properties into a single common SharePoint column where one or more of the document properties are empty.
   For example, when both the email's From and FromAddress properties are mapped to a single SharePoint column named Sender and the From property has a value and the FromAddress is empty the resulting value for the Sender was empty.

4.42 Feature and bug fix release (2.0.3)

1. Allow trimming of X500 addresses by using the optional parameter TruncateX500Address=true
   For example, instead of using the X500 value */O=OEXCH082/OU=EXCHANGE ADMINISTRATIVE*

GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=6C83EA5270B446858D39D8713DE2D5A1-someuser the value someuser is extracted.

2. Fix bug preventing the use of sp-request module 3.0.0

4.43 Bug fix release (2.0.2)

1. Same logic as 2.0.1 with the exception that the Office and PDF metadata extraction is not supported. The 2.0.2 version can only be used for extracting properties from msg and eml emails. This allows for using existing configurations working with BulkMetadataManager 1.x
   The following modules are therefore not required:
   - npm install xml2js
   - npm install adm-zip
   - npm install dom-parser

4.44 Bug fix release (2.0.1)

1. Use separate credentials for root and tenant credentials.
   See section 3.1.4 in the Administrator Guide.
2. Extract sent date and received date from emails accepting a meeting.
3. Fix bug affecting the use of Sites=* that depends on search to list all sites present in the tenant.

4.45 Feature release (2.0.0)

1. Support for using a single set of Client Id and Client Secret at tenant level.
2. Support for granular control of sites and libraries via wild cards, using specific site/library names and exclusions.
3. Support for extracting metadata from additional document formats:
   - Office files (docx, xlsx, pptx and vsdx)
   - PDF
4. Support for specifying parameters via an .ini file.

4.46 Bug fix release (1.0.13)

1. Change in parameter naming convention to circumvent conflicts with Email Manager. The optional field EmailNameSyntax containing the email naming syntax needs to be changed to EmailNameSyntaxBulkMetadataManager.

4.47 Bug fix release (1.0.11)

1. Support for using a default user account as “Modified By” when the item’s existing “Modified By” value is “SharePoint App”. This feature also circumvents errors when updating the “Modified” and “Modified By” columns using the “Modified By” value app@sharepoint. This username cannot be resolved by SharePoint and will result in errors for both columns. The

default user account needs to be specified via the tenant properties or the static properties.txt file and the key is named “ModifiedByDefault”.

2. Fix bug preventing the extraction of email metadata in emails with attached msg files.

4.48 Bug fix release (1.0.10)

1. Circumvent bug in SharePoint’s $batch requests where URLs are case sensitive. This avoids http 401 responses with “invalid request”.

4.49 Bug fix release (1.0.9)

1. Fix bug affecting msg files where metadata is present as non-Unicode characters.

4.50 Bug fix release (1.0.8)

1. Use the tenant properties (via StorageEntity) as primary source for configuration settings. If the tenant properties are not available, the application will then try to open the static BulkMetadataManagerProperties.txt file in the SiteAssets library in the root of the tenant.
2. Improved error handling when the specified Content Type is not available.
3. Omit the Client Id’s and Client Secrets from the summary log file.

4.51 Bug fix release (1.0.7)

1. Fix bug with processing large number of emails (5000+) in large lists.

4.52 Feature release (1.0.6)

1. Increase the maximum file size for msg file from 50 MB to 75 MB.

4.53 Feature release (1.0.5)

1. Increase the maximum file size for msg file from 25 MB to 50 MB.

4.54 Feature release (1.0.4)

1. Support for keeping the original “Modified” and original “Modified By” details.
   Note: this requires an extra request per item and negatively impacts end-to-end throughput. It should only be used when it adds value.
   Include “Modified”, “Modified By” and size details of the original items in the log file.

4.55 Initial release (1.0.0)

1. Support for SharePoint Online with App Only authentication (using Client Id and Secret Id).

2. Support for extracting email properties (From, Subject, Sent, …) from emails (msg and eml) stored in SharePoint Online and capturing the values into SharePoint columns.
3. WhatIf functionality to test the application without updating the email metadata in SharePoint. The extracted email metadata is written to a csv file for further review and analysis.
4. Select documents for processing using CAML.
5. Optimized performance by using parallel processes, caching, and batch requests.
6. Control performance by setting the number of download and metadata threads.
7. Support for large lists (exceeding list view threshold).
8. Support for continuing after being throttled by SharePoint.
   The requests will automatically start after the specified retry period has expired.
9. Logging.
10. Support for Windows Server 2016 and Windows Server 2019.

5. Known Issues

This section lists known issues.

1. The optimum number of parallel threads for downloads and setting metadata depends on several factors:
   - sizes of the documents involved
   - latency between the system running node.js and SharePoint
   - available bandwidth to the SharePoint system
   - time of the run.
   The throttling limits seems to depend on the load on SharePoint. During non-Office hours the limits seem to be less strict.

   It is recommended to start the application with 1x download, 1x metadata thread and 1x listview thread and adjust these values based on the findings.
2. Extracting metadata from encrypted PDF files is not supported.
3. The log file log_yyyymmdd_hhmmss.csv contains duplicate entries when a value (e.g., length 500 characters) exceeds the length of a single line of text (255 characters) and is automatically truncated.

SLIM Applications B.V.
Rotterdamseweg 402H
2629 HH Delft, The Netherlands
www.slimapplications.com
info@slimapplications.com