ELO Sync

Synchronisation

Inhaltsverzeichnis

Synchronisierung **3**

Konzepte 3

Konfiguration **8**

Sync job Konfigurationseinstellungen über die Benutzeroberfläche 8
Metadaten 9
Unterstützte Target-URLs für SharePoint 14
Exchange - Datenübersicht 15

Synchronisierung **16**

Job-Protokoll 16
Standard-Pipeline für die Synchronisierung 18
Konflikt Handling 26
Freigaben Handling 27
Bekannte Probleme und Beschränkungen 28

Synchronisierung

Konzepte

Dieser Abschnitt beschreibt die Konzepte, die für die Synchronisierung der Daten in ELO Sync verwendet werden.

Pipeline

Für die Synchronisierung wird derzeit eine zweistufige Pipeline verwendet, bei der mehrere Middlewares registriert werden.

Jede Stufe der Pipeline muss abgeschlossen sein, bevor die nächste Stufe gestartet wird. Damit soll sichergestellt werden, dass sich verschiedene Änderungen in mehreren Systemen nicht gegenseitig beeinträchtigen können und somit weniger Komplexität bei der Verarbeitung dieser Änderungen erforderlich ist.

Ein großer Nachteil dieses Ansatzes ist, dass alle Änderungen im Speicher gespeichert werden, bis die nächste Stufe bearbeitet werden kann, was zu einem höheren Speicherverbrauch führt.

Im Allgemeinen funktioniert die gesamte Verarbeitung für die Synchronisierung von oben nach unten, d. h. Ordner werden vor ihren Inhalten verarbeitet, wodurch sichergestellt wird, dass alle Änderungen an Ordnerinhalten darauf vertrauen können, dass ihre übergeordneten Ordner bereits vollständig synchronisiert sind.

Die erste Stufe umfasst die Analyse und Durchführung von Änderungen, die nicht die Löschung betreffen. Die Daten aller synchronisierten Systeme werden abgerufen und mit den gespeicherten Metadaten in ELO Sync verglichen, um festzustellen, welche Änderungen seit der letzten Synchronisation vorgenommen wurden. Alle erkannten Änderungen werden in Modifikationen umgewandelt, die für eine spätere Ausführung geplant werden.

Werden Löschungen gefunden, so werden diese zunächst normal wie andere Änderungen eingeplant, dann aber in die zweite Stufe der Pipeline übertragen. Dies ist notwendig, um zu verhindern, dass nicht gelöschte Einträge gelöscht werden. Ein Beispiel: Ein Benutzer verschiebt eine Datei von Ordner A nach Ordner B und löscht dann Ordner A. Wenn die Löschung nicht zu einem späteren Zeitpunkt geplant wird und gewährleistet ist, dass die Eltern vor ihren Kindern synchronisiert werden, würde beim Löschen der Ordner A (einschließlich der verschobenen Datei) gelöscht.

Weitere Einzelheiten finden Sie im Kapitel Pipeline.

System

Ein *System* beschreibt eine Software, einen Server oder einen Teil davon, dessen Daten synchronisiert werden sollen.

Dabei kann es sich um ein SaaS-System wie OneDrive oder SharePoint Online oder um ein lokales ELO Repository handeln.

System Provider

Eine Softwarekomponente, die die Implementierung für den Zugriff auf ein System bereitstellt. Wird nur manchmal verwendet, wenn eine Delegation der Systemerstellung erforderlich ist.

Collection

Collections sind eine Teilmenge der Inhalte eines Systems, die einen abstrakten Zugriff auf diese Inhalte ermöglichen.

Einige Systeme verwenden mehrere Collections (vor allem, wenn die Inhalte nicht eindeutig identifizierbar sind) oder nur eine einzige Collection.

Die genauen Einzelheiten hängen von den vom System bereitgestellten Daten ab.

Subcollection

Der Begriff *Subcollection* wird für Collections verwendet, die in einer anderen Collection enthalten sind oder auf die von einem Element in einer Collection verwiesen wird.

Synchronization Entry

Ein *Synchronization Entry* oder einfach *Entry* ist eine abstrakte Einheit, die eine Sache beschreibt, die zwischen mehreren Systemen synchronisiert wird.

Ein Entry hat keine eigenen Daten, sondern ist die Obermenge aller Items, die zusammen synchronisiert werden.

Wenn es notwendig ist, ein Entry eindeutig zu identifizieren, wird empfohlen, die Identität der zugehörigen Datenzuordnung zu verwenden. Diese Identität bleibt so lange erhalten, wie die einzelnen Elemente des Entrys existieren.

Synchronization Item

Ein *Synchronization Item* beschreibt ein bestimmtes Item in einem synchronisierten System. Dabei kann es sich zum Beispiel um eine Datei oder einen Ordner handeln, muss es aber nicht. ELO Sync könnte auch abstraktere Daten wie E-Mails oder Chats synchronisieren.

Data Mapping

Das *Data Mapping* beinhaltet die Metadaten eines synchronisierten Entrys. In der Regel bedeutet dies, dass das Data Mapping Hash-Werte oder ähnliche Elemente enthält, damit Änderungen zu einem späteren Zeitpunkt erkannt werden können.

Die Daten im Mapping können lange gespeichert werden und dürfen daher keine potenziell geschützten Informationen (wie PII) enthalten. Dies ist ein weiterer Grund für die Verwendung von Hash-Algorithmen (oder vergleichbaren Verfahren) zur Speicherung aller Daten in dem Mapping.

Structure Mapping

Die *structure mapping* eines Entrys speichert seine relative Position in der synchronisierten Struktur.

Im Gegensatz zum Data Mapping werden im Structure Mapping keine Daten gehasht oder verschlüsselt, da diese bei späteren Synchronisationsläufen wieder rekonstruierbar sein müssen.

Field

Auf die Daten der synchronisierten Elemente wird über *Fields* zugegriffen. Jedes Field ermöglicht den Zugriff auf eine einzige, nicht aufteilbare Information über das Element.

Nicht aufteilbar bedeutet in diesem Fall, dass die Synchronisation diese Information vollständig oder gar nicht synchronisieren soll. Das bedeutet nicht, dass die Informationen nicht aufgeteilt und zusammengeführt werden können.

System Field

Die *System Fields* sind ein Sonderfall, der von allen synchronisierten Systemen bereitgestellt werden muss, um die Synchronisierung überhaupt zu ermöglichen.

Die derzeit definierten System Fields sind: *Name*, *Content* und *Parent*.

Von diesen Feldern muss nur das Feld *Name* einen Inhalt haben.

Die beiden Felder *Content* und *Parent* müssen angegeben werden, können aber auch leer sein.

Field Mappings

Durch die Verwendung von *Field Mappings* werden einzelne Felder über verschiedene Systeme hinweg abgeglichen, sodass der Synchronisationsprozess weiß, welche Daten verglichen und ausgetauscht werden müssen.

Normalerweise konfiguriert der Benutzer die Field Mappings bei der Erstellung eines Synchronisierungsauftrags.

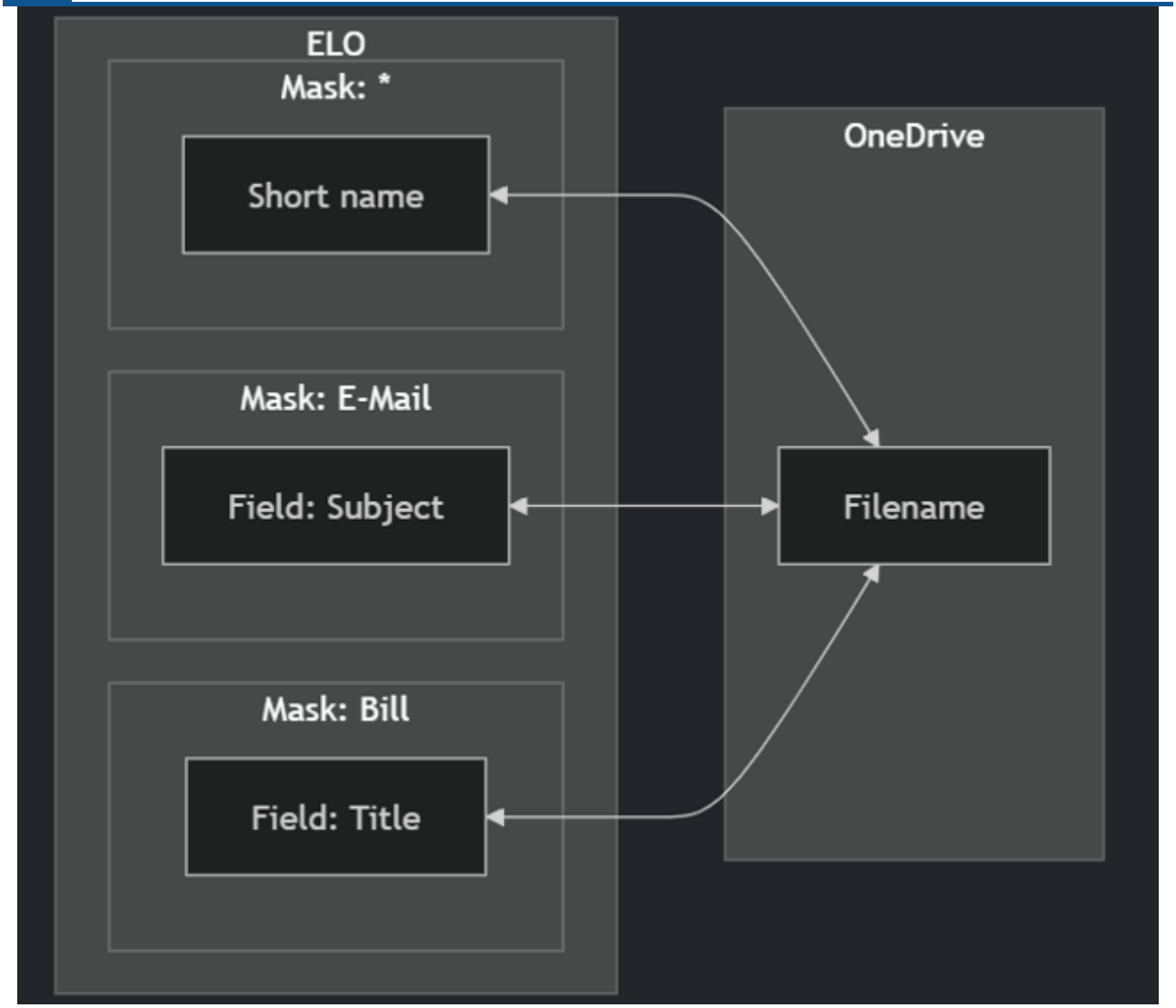
Die Systemfelder sind von den vom Benutzer vorgenommenen Zuordnungen ausgenommen, diese Felder werden immer automatisch zugeordnet.

In ELO ist es möglich, mehrere Dokumente mit unterschiedlichen Masken in demselben synchronisierten Ordner zu erstellen. Folglich sind die Feldzuordnungen nicht einfach Eins-zu-Eins-Zuordnungen zwischen zwei Systemen.

Die Felder im Synchronisierungsprozess werden immer als n-zu-m-Beziehungen zueinander betrachtet, aber wenn bestimmte Elemente synchronisiert werden, sollten die Zuordnungsregeln zu einer Eins-zu-Eins-Zuordnung führen.

Beispiel

Als Beispiel sind hier die Zuordnungen für eine Synchronisierung zwischen einem ELO Repository und OneDrive definiert.

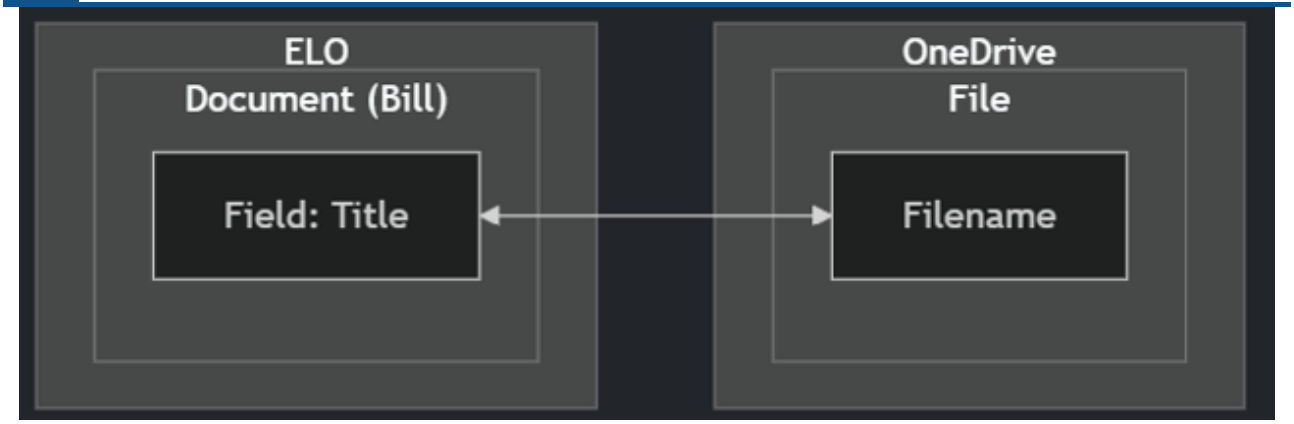


Die Datei in OneDrive hat nur den Dateinamen, aber im ELO Repository sind mehrere Masken verfügbar, mit denen dieser Dateiname synchronisiert werden kann.

Wenn jedoch ein bestimmtes Dokument mit der entsprechenden Datei in OneDrive synchronisiert wird, ist die Maske des Dokuments bekannt, sodass sich die Zuordnungen auf eine einzige Zuordnung reduzieren.

> **Information**
>
> Dies ist nur ein Beispiel.
>
> In einem Produktionsszenario ist der Dateiname ein Systemfeld und wird immer mit dem Systemfeld *Short name* in ELO synchronisiert.

Wenn das Dokument in ELO die Maske *Bill* hat, dann würde folgende Synchronisation ausgeführt werden:


ELO
Document (Bill)
Field: Title
OneDrive
File
Filename

Konfiguration

Sync job Konfigurationseinstellungen über die Benutzeroberfläche

Diese Konfigurationseinstellungen können über die Benutzeroberfläche global oder für jeden Synchronisierungsauftrag konfiguriert werden.

Metadatenzuordnung

Um die Metadaten der Spalten in den SharePoint-Bibliotheken und die ELO-Felder in eine Synchronisation einzubeziehen, müssen sie miteinander verknüpft werden. Dies geschieht über ein Metadaten-Mapping. Weitere Informationen finden Sie hier.

Metadaten

In diesem Kapitel wird beschrieben, wie die Synchronisierung von Metadaten von oder zu SharePoint Online funktioniert und konfiguriert werden kann.

Was sind Metadaten?

Bei SharePoint Online beziehen sich die Metadaten auf die nicht systemrelevanten Spalten der Dokumentenbibliotheken. Weitere Informationen zu SharePoint Online, z. B. zum Erstellen von Dokumentbibliotheken oder benutzerdefinierten Spalten, finden Sie in der SharePoint-Dokumentation.

In ELO sind die Metadaten die Felder der Index- oder Aspektmasken. Weitere Informationen zu ELO Indexmasken und Aspektmasken finden Sie in der Dokumentation.

Unterstützte Spalten

In SharePoint Online werden nur bestimmte Spalten unterstützt, bei denen eine Übertragung nach ELO möglich ist.

Benutzerdefinierte Spalten sind unterstützt, sofern diese einen der folgenden Typen haben:

- Text
- Datum und Uhrzeit
- Mehrzeiliger Text
- Zahl
- Ja/Nein
- Hyperlink
- Währung

Von den Systemspalten, die SharePoint selbst bereitstellt, werden die folgenden Spalten unterstützt:

Bezeichnung	Metadaten-Regel
Complianceobjekt-ID	`*\*\ComplianceAssetId`
Titel	`*\*\Title`
Erstellt	`*\*\Created`
Geändert	`*\*\Modified`
Kommentar zum Einchecken	`*\*\_CheckinComment`
Untergeordnete Elementanzahl	`*\*\ItemChildCount`
Untergeordnete Ordneranzahl	`*\*\FolderChildCount`
Bezeichnungseinstellung	`*\*\_ComplianceFlags`
Aufbewahrungsbezeichnung	`*\*\_ComplianceTag`
Aufbewahrungsbezeichnung angewendet.	`*\*\_ComplianceTagWrittenTime`
Bezeichnung angewendet von	`*\*\_ComplianceTagUserId`
Anzahl der Kommentare	`*\*\_CommentCount`
Anzahl der Likes	`*\*\_LikeCount`

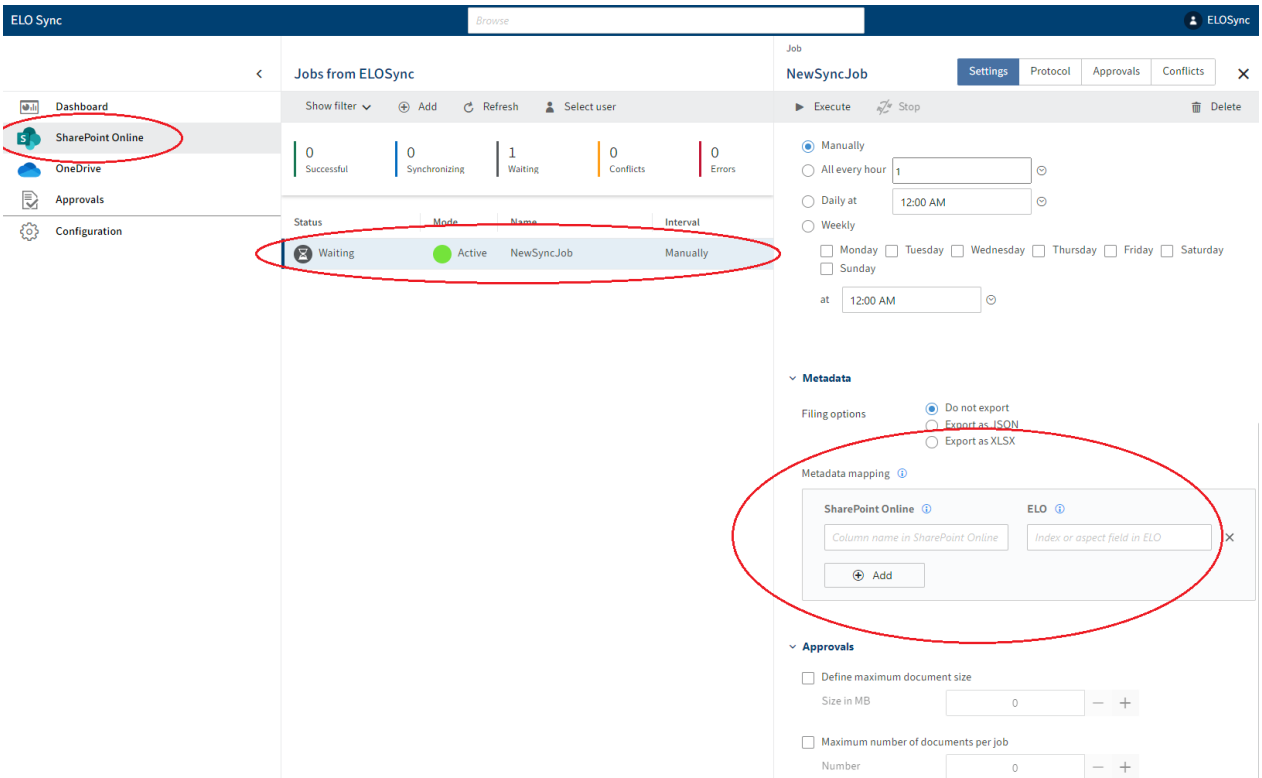
Bezeichnung	Metadaten-Regel
App erstellt von	`*\*\AppAuthor`
App geändert von	`*\*\AppEditor`

Diese Spalten sind nur lesbar, d.h. es können keine Metadaten aus ELO nach SharePoint übertragen werden.

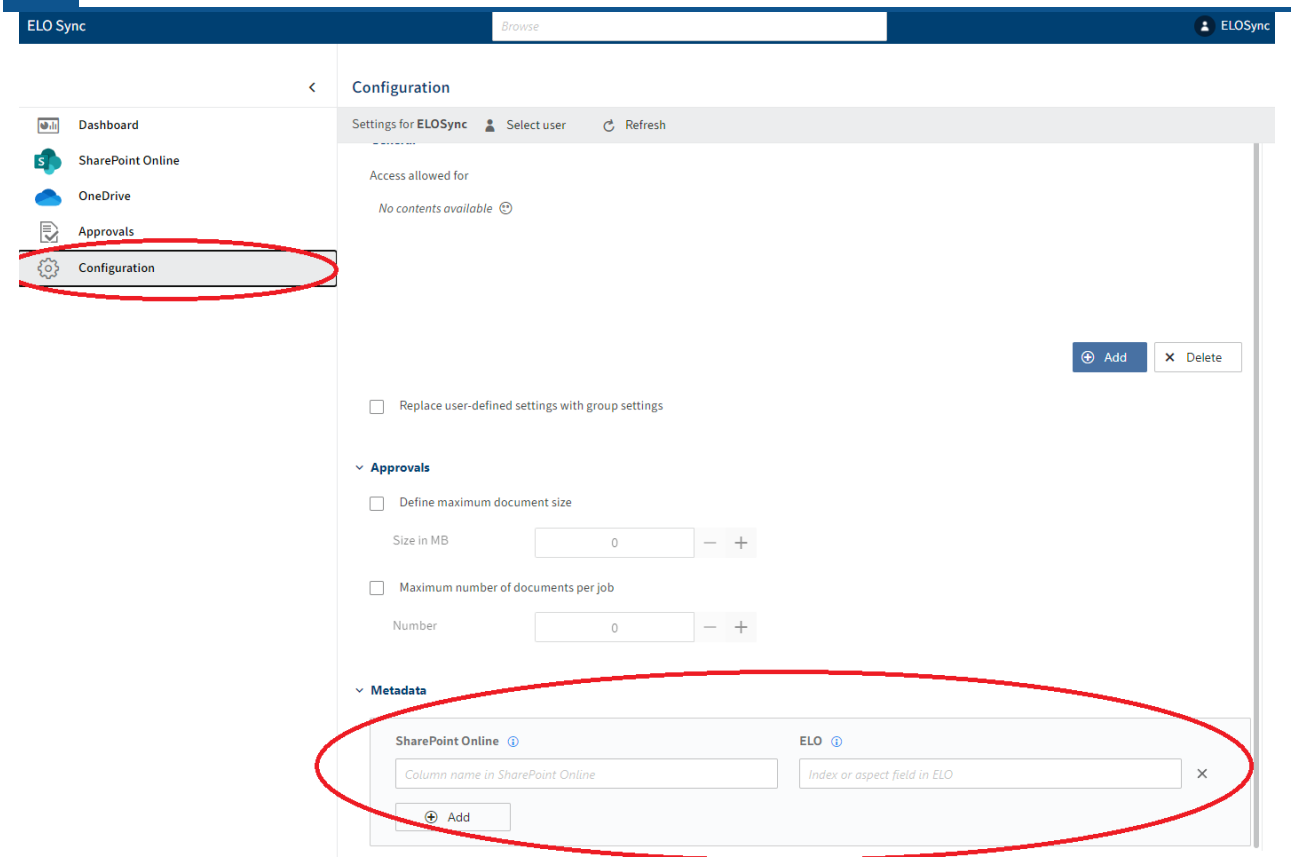
Metadatenzuordnungen

Um die Metadatenfelder zwischen SharePoint Online und ELO zu verknüpfen, muss ein Mapping der Felder definiert werden. Diese kann entweder pro SyncJob in der jeweiligen Konfiguration oder global in den Einstellungen erstellt werden, wobei die Verknüpfungen der SyncJob-Konfiguration Priorität haben.

> **Information**
>
> In den Feldern werden Zahlen auf 15 Stellen gerundet.



Metadatenzuordnungen in der SyncJob-Konfiguration



Metadatenzuordnungen in der globalen Konfiguration

Wenn ein Feld nicht zugeordnet werden konnte, werden die Metadaten für das Feld nicht übertragen.

Struktur einer Metadatenzuordnung

Die Zuordnungen (im Folgenden als *Mappings* bezeichnet) können aus mehreren Teilen bestehen und werden durch ein "" getrennt.

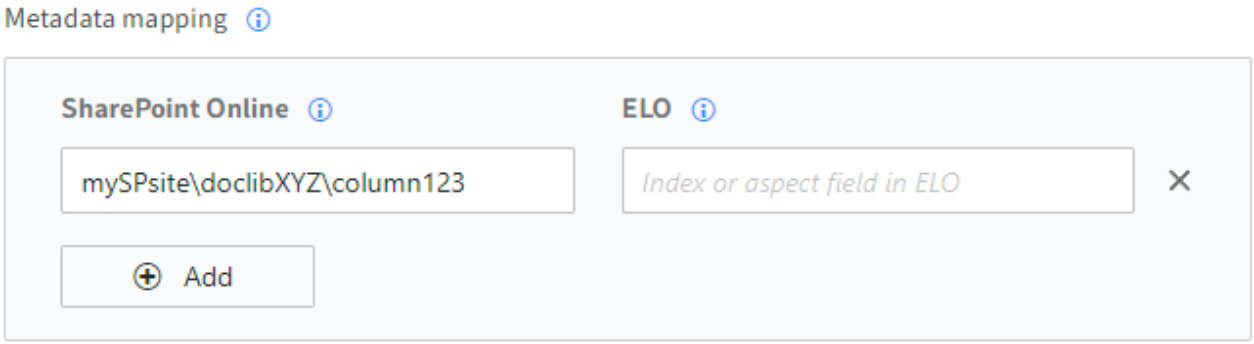
Struktur der Mappings für SharePoint Online

Die Struktur eines Mappings für SharePoint Online kann wie folgt aussehen:

SiteName\ListName\ColumnName

Seiten- und Listennamen sind hier optional. Um Verwechslungen zu vermeiden (z. B. eine Spalte aus einer anderen Liste), sollten die Zuweisungen so viele Informationen wie möglich enthalten.

Beispiel:



Struktur der Mappings für ELO Index- oder Aspektfelder

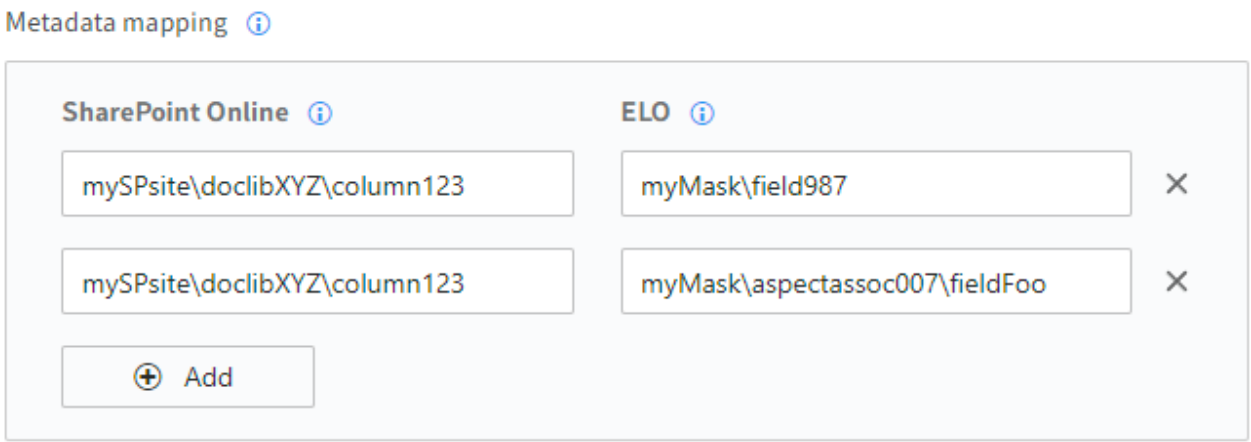
Die Struktur eines ELO-Mappings kann wie folgt aussehen:

Indexfelder: *MaskName\FieldName*

Aspektmasken: *MaskName\AspectAssociation\FieldName*

Maskennamen und Aspektzuordnungen sind hier optional. Um Verwechslungen zu vermeiden (z. B. eine Spalte aus einer anderen Liste), sollten die Zuordnungen so viele Informationen wie möglich enthalten.

Beispiel:



Wildcards

Darüber hinaus können die beiden Platzhalterzeichen "*" und "?" in den Zuordnungen verwendet werden. Dabei steht "?" für ein beliebiges Zeichen und "*" für eine unbegrenzte Anzahl beliebiger Zeichen.

Beispiel:

mySPsite\doclib*\column1?3 myMask*\fieldF??

Zuordnung von Feldern

Die Zuordnungen der Felder können sowohl in den Einstellungen der jeweiligen SyncJobs als auch in den globalen Einstellungen vorgenommen werden. Die Zuweisung in den SyncJob-Einstellungen hat eine höhere Priorität, d.h. wenn eine Zuweisung eines Feldes in den SyncJob-

Einstellungen und in den globalen Einstellungen definiert wurde, wird nur diejenige in den SyncJob-Einstellungen berücksichtigt.

Die jeweiligen Zuweisungen werden von der kleinsten zur größten Einheit, d. h. von hinten nach vorne, aufgelöst. Eine Zuweisung wird also zunächst auf den Spalten- oder Feldnamen ausgeführt.

Anhand des Beispiels aus dem Abschnitt Struktur der Mappings für SharePoint Online `mySPSite\\doclibXYZ\\column123` wird die Zuordnung in der folgenden Reihenfolge aufgelöst:

1. *column123*
2. *doclibXYZ*
3. *mySPSite*

Unterstützte Target-URLs für SharePoint

- URLs, die über die Freigabefunktion von SharePoint erstellt werden, meist durch Rechtsklick auf eine Datei oder einen Ordner und anschließende Auswahl von *Link kopieren*
- Die Browser-URLs der meisten SharePoint-Seiten

Detailierte Liste der unterstützten URLs

Dies ist eine Liste von SharePoint-URLs, die unterstützt werden, wobei Wildcards (*/?) für die variablen Teile verwendet werden.

Die Pattern werden von oben nach unten getestet. Wenn ein Pattern übereinstimmt, wird es verwendet und die restlichen Pattern werden übersprungen.

URL Pattern	Hinweise
https://tenant.sharepoint.com/:*:/s/*	Sharing Link; wählt die Bibliothek/Liste aus, wenn sie auf eine Datei abzielt
https://tenant.sharepoint.com/*/r/sites/*	OneNote-Notebook; wählt die enthaltende Bibliothek aus
https://tenant.sharepoint.com/sites//.aspx	URL des Formulars; möglicherweise wird nicht das direkte Ziel ausgewählt, sondern ein übergeordnetes Ziel
https://tenant.sharepoint.com/sites//*Lists/*	URL der Listenansicht; möglicherweise wird nicht das direkte Ziel ausgewählt, sondern ein übergeordnetes Ziel
https://tenant.sharepoint.com/sites//*_layouts/*	URL der Bibliotheksansicht; möglicherweise wird nicht das direkte Ziel ausgewählt, sondern ein übergeordnetes Ziel
https://tenant.sharepoint.com/sites//_/*	URL der Site-API; wählt die zugehörige Site aus
https://tenant.sharepoint.com/sites/*	Andere URLs; möglicherweise wird nicht das direkte Ziel ausgewählt, sondern ein übergeordnetes Ziel davon
https://tenant.sharepoint.com/	Wählt die Root-Seite aus
https://tenant.sharepoint.com/_/	SharePoint API; Wählt die Root-Seite aus

Es gibt Einschränkungen, wenn die Root-Site ausgewählt wird, da die von Microsoft bereitgestellte API keine Inhalte für die Root-Site selbst liefert.

So ist es zwar möglich, eine URL für die Root-Site einzugeben, aber es ist nicht möglich, ein tatsächliches SharePoint-Synchronisierungsziel auszuwählen.

Dies kann sich in Zukunft ändern, wenn Microsoft die API ändert, wird aber derzeit nicht unterstützt.

Exchange - Datenübersicht

In diesem Kapitel erfahren Sie, welche Daten aus Exchange Online mithilfe von ELO Sync übertragen werden können.

Elemente

Folgende Elemente können aus Exchange Online synchronisiert werden:

- E-Mails: E-Mail werden immer im .eml-Format synchronisiert, das .msg-Format aus Outlook wird nicht unterstützt.
- E-Mail-Ordner
- E-Mail-Anhänge: Bei Anhängen können derzeit nur Dateianhänge synchronisiert werden, sogenannte `fileAttachment`. Andere Arten von Anhängen, wie z. B. `itemAttachment` oder `referenceAttachment`, werden derzeit nicht unterstützt.

Metadaten

Für die Metadaten ist keine manuelle Konfiguration vorgesehen, stattdessen werden die Standard-Metadatenmasken von ELO mit den entsprechenden Exchange-Metadaten verknüpft.

Synchronisierung

Job-Protokoll

Das Protokoll des Auftrags wird in der Datenbank aufgezeichnet und kann in der Konfiguration eines Auftrags über die Schaltfläche *Protokoll* angezeigt werden.

Zusammenfassung

Dort wird zunächst eine Zusammenfassung der letzten Ausführung des Auftrags angezeigt. Es ist auch sofort ersichtlich, ob es Fehler bei der Synchronisierung von individuellen Dokumenten oder Ordner gab.

> **Information**
>
> Wenn ein Fehler nicht eindeutig einem System zugeordnet werden kann, wird dieser Fehler bei beiden Systemen gezählt.
>
> Dadurch kann es vorkommen, dass die Summe der Fehler größer ist als die Gesamtzahl der Einträge, die synchronisiert wurden.

Details

Unterhalb der Zusammenfassung können die Details zur letzten Ausführung des Auftrags aufgerufen werden.

In der zusätzlichen Ansicht sind alle Dokumente bzw. Ordner und deren Synchronisierungsstatus aufgelistet.

Wenn der Status erfolgreich ist, zeigt die Spalte *Aktionen* an, welche Aktion durchgeführt wurde.

Wenn der Status fehlerhaft ist, zeigt die Spalte *Aktionen* an, welche Aktionen versucht wurden.

> **Information**
>
> Wenn es mehrere Aktionen gibt, kann derzeit nicht unterschieden werden, ob eine oder mehrere Aktionen fehlerhaft waren.
>
> Dies wird in zukünftigen Versionen verbessert.

Detailliertes Protokoll

Das Protokoll der Aufträge kann im ELO Repository abgelegt werden, indem in der Konfiguration des Benutzers die Checkbox *Job-Protokoll im ELO Repository ablegen* aktiviert wird.

Zusätzlich ist es dort auch möglich, das Protokoll-Level zu konfigurieren. Bei Auswahl von *Debug* wird auch das technische Protokoll im ELO Repository abgelegt.

Im Job-Protokoll selbst sind alle Zeiten mit einer Referenz-ID versehen (bspw. `entry_Abasjh890gfd76gfd87`). Mithilfe dieser Referenz-ID können alle zusammenhängenden Protokolleinträge eines Dokuments oder Ordners identifiziert werden.

Standard-Pipeline für die Synchronisierung

Dieses Dokument beschreibt die Standard-Pipeline, die ELO Sync für die Datensynchronisation verwendet.

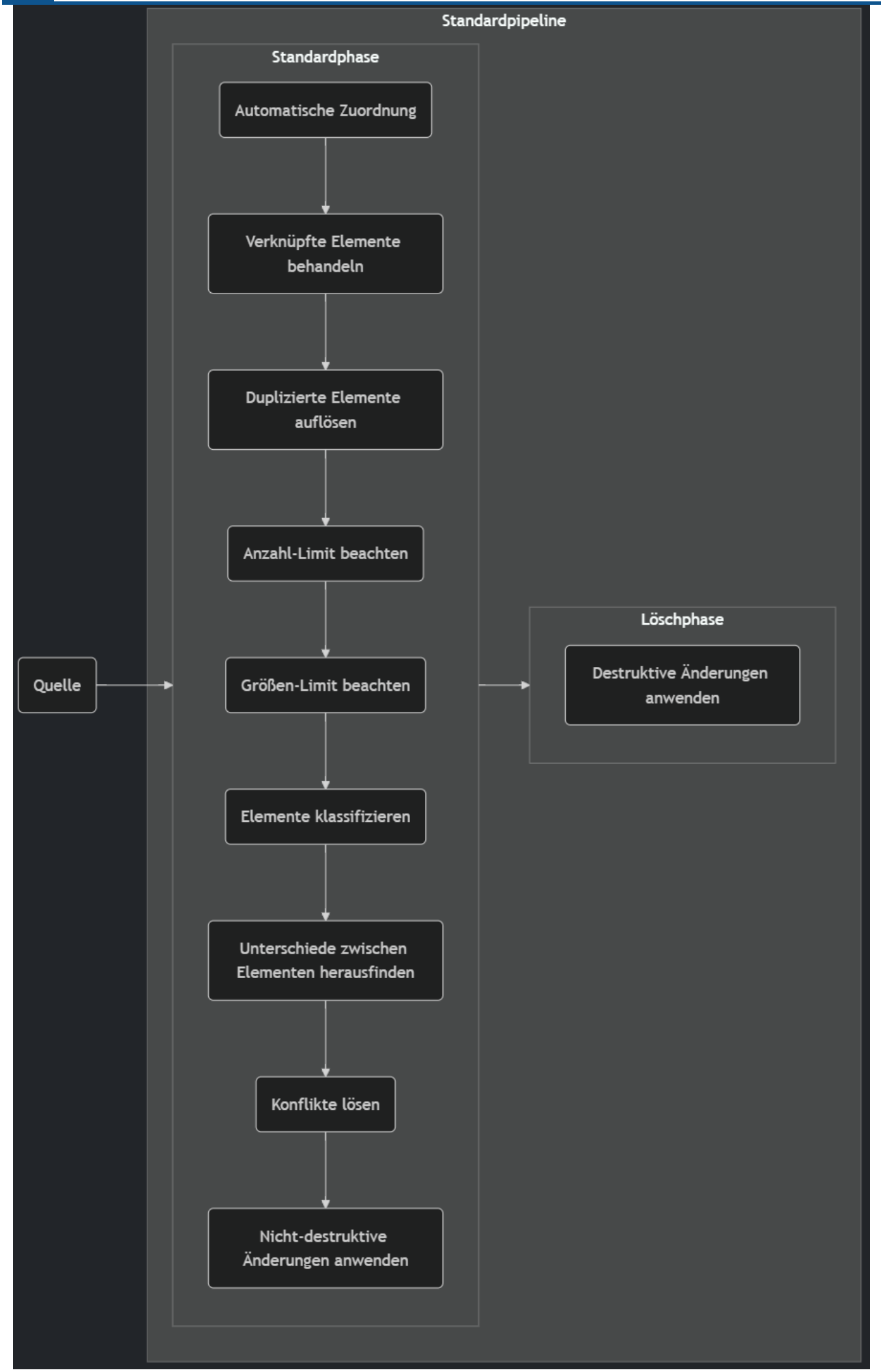
Je nach Art und Richtung des Synchronisierungsauftrags werden Elemente zur Pipeline hinzugefügt oder aus ihr entfernt.

Übersicht

Eine Synchronisationspipeline kann aus einer unbegrenzten Anzahl von Stufen bestehen, die aktuelle Standardpipeline verwendet zwei Stufen.

Jede Stufe enthält eine Reihe von Middlewares, die jeweils einen einzelnen Schritt für einen synchronisierten Eintrag ausführen.

Die nächste Stufe beginnt erst, wenn alle Einträge in den vorangegangenen Stufen bearbeitet wurden. Dadurch wird sichergestellt, dass einige Vorgänge immer nach anderen ausgeführt werden.


Standardpipeline
Standardphase
Automatische Zuordnung
Verknüpfte Elemente behandeln
Duplizierte Elemente auflösen
Anzahl-Limit beachten
Größen-Limit beachten
Elemente klassifizieren
Unterschiede zwischen Elementen herausfinden
Konflikte lösen
Nicht-destruktive Änderungen anwenden
Quelle
Löschphase
Destruktive Änderungen anwenden

Die Standard- und Löschphasen werden seriell und nicht parallel ausgeführt.

Dadurch wird sichergestellt, dass Änderungen immer in der richtigen Reihenfolge ausgeführt werden und dass destruktive Änderungen die erfolgreiche Ausführung anderer Änderungen nicht verhindern.

Darüber hinaus vereinfacht es die Anwendung von nicht-destruktiven Änderungen, da diese keine zusätzlichen Einschränkungen überprüfen müssen.

Gründe für die zweistufige Pipeline

Derzeit gibt es nur einen Grund für die Verwendung einer zweistufigen Pipeline:

Die Kombination aus dem Verschieben eines Elements und dem Löschen seines (indirekten) übergeordneten Elements zwischen zwei Synchronisationen.

In allen derzeit unterstützten Systemen wird beim Löschen eines Ordners auch sein gesamter Inhalt gelöscht. Dies führt zu einem Problem, wenn ein Element aus einem Ordner verschoben und der Ordner anschließend gelöscht wird.

Aus Performancegründen wird der Synchronisationscode immer parallel für jeden Eintrag ausgeführt, sodass es möglich ist, dass der Ordner im Zielsystem gelöscht wird, bevor das Element verschoben wurde.

> **Information**
>
> Man könnte meinen, dass eine offensichtliche Lösung darin bestünde, die aufgezeichneten Änderungen in der Reihenfolge ihres Auftretens zu ordnen, aber dies kann in Grenzfällen zu kleinen Fehlern führen:
>
> 1. Es ist nicht garantiert, dass die synchronisierten Systeme genaue Uhren verwenden.
> 2. Die Synchronisierung kann weit nach den Änderungen erfolgen, und bei einer Zwei-Wege-Synchronisierung wird das Objekt in einem System verschoben und der Ordner im anderen System gelöscht.

Da sich der Synchronisierungscode nicht auf die Genauigkeit der gemeldeten Zeitstempel verlassen kann, muss eine andere Lösung gefunden werden, damit diese Vorgänge erfolgreich durchgeführt werden können.

Die derzeitige Lösung besteht darin, alle destruktiven Änderungen nur in einer Warteschlange zu erfassen, die nach der Verarbeitung aller anderen Änderungen ausgeführt wird.

Beispiel

Gehen Sie in beiden Systemen von der folgenden Ausgangsstruktur aus:

- Folder *A*
  - File *F1*
  - File *F2*
- Folder *B*

Zwischen zwei Synchronisationsläufen verschiebt ein Benutzer die Datei *F1* nach *B* und löscht anschließend den Ordner *A*.

Um diesen Vorgang korrekt zu wiederholen, müssen die Änderungen in der gleichen Reihenfolge durchgeführt werden.

Bitte beachten Sie, dass die Verschiebe- und Löschaktion nicht auf demselben System ausgeführt werden muss und daher nicht garantiert werden kann, dass die Zeitstempel (oder ähnliche) vergleichbar sind.

> **Information**
>
> Zeitstempel können zwischen Systemen unterschiedlich sein, z. B. weil ein System einfach die falsche Zeit eingestellt hat oder ein NTP-Zeitserver vorübergehend unerreichbar war.
>
> Es gibt viele Ursachen für unzuverlässige Zeitstempel.
>
> Eine allgemeine Empfehlung für zuverlässige Systeme lautet, bei der Entscheidung über Aktionen keine Zeitstempel zu verwenden.

Source

Die Quelle für die Synchronisation holt die Positionen aus den einzelnen Systemen zur Verarbeitung und stellt sie anderen Komponenten zur Verfügung.

Alle lesbaren Systeme werden in die Verarbeitungswarteschlange für die Quelle eingefügt, und dann wird ein Generatorblock erstellt, der die Elemente abruft und sie für die weitere Verarbeitung bereitstellt. Der Generatorblock startet eine Anzahl von parallelen Workern zwischen 1 und der Anzahl der zu synchronisierenden Systeme.

Jeder Worker holt Elemente aus der Verarbeitungswarteschlange und führt je nach Art der Elemente unterschiedliche Aktionen durch.

Für ein System holt der Worker seine Root Collection und fügt alle von dieser Collection bereitgestellten Root-Elemente in die Verarbeitungswarteschlange ein. Die Elemente werden in Blöcken von bis zu acht Elementen gruppiert, sodass die spätere Verarbeitung mehrere Abrufe oder Änderungen zusammenfassen kann.

Handelt es sich bei dem verarbeiteten Element um ein Provider-Element einer anderen Sammlung, wird diese Sammlung abgerufen und ihre Root-Elemente werden in die Verarbeitungswarteschlange eingefügt. Andernfalls werden die Kinder des Eintrags abgerufen, in Gruppen von acht Einträgen unterteilt und dann wie Einträge in einem System verarbeitet.

Executor

Der Executor führt jede konfigurierte Pipelinestufe seriell aus, und innerhalb jeder Stufe alle Einträge parallel.

Für die erste Stufe wird der Quellblock aus der Quelle entnommen und verarbeitet, spätere Stufen verarbeiten jeweils die Warteschlange für ihre Stufe als Quellblock.

Nach Erhalt eines Eingangskontextes zur Bearbeitung prüft der Executor, ob der aktuelle Kontext einen Vorgänger hat und wartet, bis er beendet ist, falls ein solcher gesetzt ist. Danach startet der Executor die eigentliche Ausführung der Middlewares, die für die Stufe konfiguriert wurden.

Middlewares

Automatische Zuordnung

Diese ordnet automatisch Objekte zwischen verschiedenen Systemen anhand verschiedener Kriterien zu.

Die Standardkonfiguration für diese Middleware ordnet Elemente auf der Grundlage ihres relativen Pfads innerhalb der synchronisierten Struktur zu.

Beispiel

Falls eine Dokumentenbibliothek in SharePoint Online eine Datei an dem Pfad *Project/Marketing/Presentation.pptx* besitzt, dann wird in dem ELO Repository nach einem Ordner *Project* gesucht, innerhalb des Ordners nach einem Ordner *Marketing*, und darin nach einem Dokument mit der Kurzbezeichnung *Presentation* mit einer aktuellen Version, die die Dateiendung *.pptx* besitzt.

Verknüpfte Elemente behandeln

Diese wird benötigt, wenn frühere Schritte in der Pipeline nur die Ids der Elemente bestimmen konnten, aber die Elemente selbst nicht abgerufen haben.

Diese Middleware sucht dann in der Datenbank nach den Metadaten dieser Elemente und fragt, wenn möglich, die Systeme, zu denen sie gehören, ab, ob diese Elemente existieren (oder gelöscht wurden).

Damit ist gewährleistet, dass alle späteren Middlewares zumindest Zugriff auf die Metadaten und den Status dieser Elemente haben. Wenn Elemente gelöscht wurden, werden diese Elemente durch entsprechende Markierungen ersetzt, die von anderen Middlewares untersucht werden können.

Duplizierte Elemente auflösen

Diese Middleware wird zur Optimierung verwendet, um zu verhindern, dass Elemente in demselben Synchronisierungslauf mehrfach synchronisiert werden.

> **Information**
>
> Das Worst Case Szenario ist die Verwendung des Deltamodus (nur Differenzen abrufen) und die angeschlossenen Systeme listen dann jedes Element für jede Änderung mehrfach auf.
>
> Wenn zum Beispiel ein Artikel seit der letzten Synchronisierung 100 Änderungen erhalten hat, könnte das System diesen Artikel 100 Mal auflisten.
>
> Dies würde zu mindestens 100 Synchronisierungen dieses Elements führen (von denen 99 unnötig wären).

Paket-Pipelines ausführen

Diese Middleware führt separate Pipelines für Pakete aus. Es gibt sowohl eine Pipeline, die für die Synchronisierung der Paketelemente ausgeführt wird, als auch eine, die nach der Synchronisierung aller Elemente im Paket ausgeführt wird.

Zurzeit wird nur die Nach-Paket-Pipeline verwendet, um E-Mail-Anhänge mit ihren zugehörigen E-Mails zu verknüpfen. Lesen Sie dazu den Abschnitt E-Mail-Anhänge mit E-Mail verknüpfen.

Anzahl-Limit beachten

Diese Middleware erzwingt das konfigurierte Entry-Limit für einen Sync-Auftrag.

Wenn kein Limit konfiguriert wurde, tut diese Middleware nichts.

Wenn das konfigurierte Limit erreicht ist, wird automatisch eine Freigabe generiert oder, wenn ein hartes Limit konfiguriert ist, werden alle weiteren Einträge übersprungen.

Solange die Freigabe nicht von einem Benutzer bestätigt wurde, werden alle zusätzlichen Elemente nicht synchronisiert.

> **Information**
>
> Alle Einträge werden aus den Systemen geholt, auch wenn das Limit bereits erreicht ist.
>
> Dies ist erforderlich, um die Gesamtzahl der Einträge genau anzugeben.

- Freigabebehandlung

Größen-Limit beachten

Diese Middleware erzwingt die konfigurierte Dateigrößenbegrenzung für synchronisierte Dateien.

Wenn kein Limit konfiguriert wurde, tut diese Middleware nichts.

Für jede Datei, die das konfigurierte Limit überschreitet, wird eine Freigabe generiert oder, wenn ein hartes Limit konfiguriert ist, wird die Datei nicht synchronisiert.

Solange die Freigabe nicht von einem Benutzer erteilt wurde, wird die Datei nicht synchronisiert.

- Freigabebehandlung

Elemente klassifizieren

Diese Middleware stellt sicher, dass die Zustandsklassifizierung von Gegenständen korrekt ist, oder bestimmt ihren Zustand, wenn er unbekannt ist.

Ähnlich wie die Verknüpfte Elemente behandeln wird diese Middleware verwendet, um sicherzustellen, dass spätere Middlewares einen konsistenten Zustand für alle synchronisierten Items haben.

Diese Middleware hat die folgenden Aufgaben:

-

Sicherstellen, dass neu erstellte Objekte als neu markiert werden
- Sicherstellen, dass geänderte Elemente tatsächlich als geändert markiert sind
- Sicherstellen, dass nicht geänderte Elemente tatsächlich als nicht geändert gekennzeichnet sind

Unterschiede zwischen Elementen herausfinden

Diese Middleware führt den Großteil der Analyse während der Synchronisierung durch.

Es wird ermittelt, welche Änderungen an Elementen vorgenommen wurden und ob diese Änderungen mit anderen Änderungen in Konflikt stehen.

Zunächst wird aus allen verbundenen Elementen ein Prototyp ermittelt.

Dieser Prototyp wird als Vergleichsvorlage für alle anderen Elemente verwendet.

Dann holt es alle Felder aus dem Prototyp und ermittelt dann für jedes Feld die verbundenen Felder.

Jeder Feldeintrag wird dann auf Änderungen geprüft, und wenn nötig und möglich, werden die Inhalte verglichen.

Wenn es keine in Konflikt stehenden Änderungen gibt, wird ein entsprechendes Objekt erstellt, das die spätere Bearbeitung der Änderung ermöglicht. Das Modifikationsobjekt enthält alle erforderlichen Informationen für die Durchführung der Änderung.

Wenn es mehrere, sich in Konflikt stehende Änderungen an verschiedenen Feldern gibt, wird ein Konflikt erzeugt. Es wird ein neuer Konflikt mit allen relevanten Daten erstellt, einschließlich der Elemente und Felder, die von diesem Konflikt betroffen sind.

Konflikte lösen

Diese Middleware wird verwendet, um Konflikte zu bearbeiten und sie aufzulösen.

Wenn ein Handler einen Konflikt auflösen kann, wird eine entsprechende Modifikation erstellt und mit den anderen Änderungen in eine Warteschlange gestellt.

Wenn kein Handler eine Lösung anbieten kann, wird der Konflikt in das technische Log geschrieben und die Synchronisierung dieses Eintrags abgebrochen, um sicherzustellen, dass keine fehlerhaften Änderungen vorgenommen werden.

- Für weitere Informationen siehe: Konfliktbehandlung

Nicht-destruktive Änderungen anwenden

Diese Middleware führt die eigentliche Synchronisation durch, nachdem alle Analysen abgeschlossen sind.

Diese Middleware wird zu mehreren Stufen hinzugefügt und führt je nach Stufe unterschiedliche Änderungsarten aus.

Die tatsächliche Ausführung von Änderungen wird an die registrierten Change-Handler delegiert, die ihrerseits Änderungen für spätere Phasen verschieben oder sogar ausstehende Änderungen in der aktuellen Phase neu anordnen können.

Die derzeit registrierten Handler haben das folgende Verhalten:

- In der Standard-Stage werden alle Änderungen, mit Ausnahme von Löschungen, direkt ausgeführt. Die Ausführungsreihenfolge richtet sich nach dem Intra-Entry-Staging der Änderungen, das zuvor von den Analyse-Middlewares festgelegt wurde.
- Alle destruktiven Änderungen werden für die Ausführung in der Lösch-Stage erneut eingeplant.
- In der Lösch-Stage werden alle verbleibenden geplanten Änderungen ausgeführt.

Optionale Middlewares

Je nach Art des Synchronisierungsauftrags können weitere Middlewares zur Pipeline hinzugefügt werden. Diese Misddlewares sind hier dokumentiert, aber nicht in der obigen Standard-Pipeline aufgelistet.

E-Mails markieren

Bei der Ablage von E-Mails ist es optional möglich, E-Mails zu markieren, um für den Benutzer anzuzeigen, dass sie bereits synchronisiert wurden.

Folgende Markierungen sind möglich:

- Markierung der E-Mail als gelesen
- Hinzufügen eines Präfix zum Betreff der E-Mail

Intern werden auch versteckte Metadaten gesetzt, die mit ELO XC kompatibel sind, damit E-Mails nicht erneut synchronisiert werden.

Diese Middleware wird nur für Exchange-Online-Aufträge verwendet und nach der Middleware Nicht-destruktive Änderungen anwenden in der Standard-Phase eingefügt.

> **Information**
>
> Falls die Ablage für eine E-Mail nicht erfolgreich war, werden keine Markierungen gesetzt.

E-Mail-Anhänge mit E-Mail verknüpfen

Diese Middleware wird in der Nach-Paket-Pipeline ausgeführt und verknüpft E-Mail-Anhänge mit ihren zugehörigen E-Mails im ELO Repository.

Falls die Anhänge eine entsprechende Aspektmaske verwenden, wird das Relationsfeld `EMAIL_CONNECTIONS¶0¶ATTACHMENT_TO` gesetzt.

Zusätzlich werden paarweise Verknüpfungen zwischen der E-Mail und ihren Anhängen erstellt.

Diese Middleware wird nur für Exchange-Online-Aufträge verwendet.

Konflikt Handling

Dieser Handler speichert Konflikte in der Datenbank und liefert die Auflösung, wenn sie in der Datenbank gespeichert ist.

Am Ende der Synchronisation werden alle irrelevanten und gelösten Konflikte aus der Datenbank entfernt, da diese nicht mehr benötigt werden.

Gelöste Konflikte sind Konflikte, die von einem Benutzer oder einem Programm unter Verwendung der REST-API gelöst wurden.

Irrelevante Konflikte sind Konflikte, die nicht mehr existieren. Dafür kann es verschiedene Ursachen geben, wie z.B. das manuelle Löschen aller betroffenen Elemente.

Die derzeitige Implementierung verwendet ein Versionsfeld in der Konfliktdatenbanktabelle, um irrelevante Konflikte zu finden. Zu Beginn der Synchronisation wird die Version hochgezählt und neue oder aktualisierte Konflikte speichern die neue Versionsnummer. Nach Abschluss der Synchronisierung werden alle Konflikte, die nicht die neue Versionsnummer haben, gelöscht.

> **Information**
>
> Es gab einen alternativen Ansatz, bei dem ein Zeitstempel verwendet wurde, um alle Konflikte zu löschen, die älter als der aktuelle Zeitstempel sind.
>
> Aus Gründen der Korrektheit musste dieser Zeitstempel aktualisiert werden, um zu verhindern, dass ein Konflikt, der bei einem früheren Synchronisierungslauf gefunden wurde und noch besteht, gelöscht wird.
>
> Problematisch war, dass dieser Zeitstempel auch den Benutzern angezeigt werden kann, und es kann verwirrend sein, wenn ein Konflikt plötzlich einen anderen Zeitstempel hat.

Freigaben Handling

Unter bestimmten konfigurierbaren Umständen können bei der Synchronisation Freigaben erforderlich sein.

Diese Freigaben müssen von einem Benutzer bestätigt werden, um die Synchronisierung der betroffenen Elemente oder Aufträge zu ermöglichen.

Die Idee dahinter ist, dass ein Administrator (oder die Benutzer selbst) Grenzen setzen kann, um eine versehentliche oder übermäßige Synchronisierung von Dokumenten zu verhindern.

Die aktuelle Implementierung speichert die Freigaben in der Datenbank und stellt von dort aus die konfigurierten Einstellungen bereit.

> **Information**
>
> Freigaben werden nie automatisch gelöscht.
>
> Lang laufende Synchronisierungsaufträge können im Laufe ihres Bestehens viele Freigaben erhalten.
>
> Der Grund dafür ist, dass die Benutzer bei künftigen Synchronisierungen nicht ständig dieselben Aktionen freigeben müssen.

Bekannte Probleme und Beschränkungen

Falsche Datumswerte

Der Synchronisationscode geht davon aus, dass alle Datumswerte ohne explizite Zeitzone in der UTC-Zeitzone angegeben werden.

Wenn es Quellen gibt, die Werte in der lokalen Zeitzone (oder einer anderen) liefern, kann dies zu falschen Umrechnungen und damit zu fehlerhaften Werten in anderen Systemen führen.

Keine eindeutigen Namen im ELO Repository

Derzeit erfordert die Synchronisierung, dass alle synchronisierten Ordner und Dokumente im ELO Repository eindeutige Namen haben.

Diese Einschränkung besteht, da die meisten Systeme, die synchronisiert werden können, selbst eindeutige Namen erzwingen, während ELO dies nicht tut.

Derzeit wird nur eines der Elemente mit identischem Namen mit einem anderen System synchronisiert, alle weiteren Elemente werden ignoriert. Welches Element genau synchronisiert wird, ist nicht vorhersehbar.

> **Information**
>
> Es ist geplant, die automatische Umbenennung in einer zukünftigen Version zu implementieren.
>
> Die genaue Methodik für die Umbenennung wird derzeit noch untersucht.

Sonderzeichen in Kurzbezeichnung

SharePoint Online und OneDrive haben Einschränkungen bei den Zeichen, die in Datei- und Ordnernamen verwendet werden können. Diese Einschränkung gibt es im ELO Repository nicht.

Um die Synchronisierung von Dateien und Ordnern mit diesen Sonderzeichen zu ermöglichen, wird der Kurzname auf problematische Zeichen geprüft und diese werden dann mit percent-encoding kodiert.

Die folgenden Zeichen werden immer kodiert: /, \, :, *, ?, ", <, >, |

Die Kodierung dieser Zeichen wird vom Systemprovider implementiert, nicht vom allgemeinen Synchronisationssystem, da nur der Systemprovider die Zeichenbeschränkungen kennt.

Führende/nachfolgende Leerzeichen im Kurznamen

Normalerweise werden führende und nachfolgende Leerzeichen im Kurznamen vom ELO Indexserver entfernt.

Da ELO Sync mit anderen Systemen synchronisiert, die führende und nachfolgende Leerzeichen in den Datei- und Ordnernamen zulassen, werden alle führenden oder abschließenden

Leerzeichen von externen Systemen mit dem Unicode-Zeichen `U+2420` (Symbol für Leerzeichen) kodiert und nicht aus dem Kurznamen entfernt.

Für weitere Informationen siehe FAQ-Beitrag Ungewöhnliche Zeichen im Kurznamen.

Konflikte mit unveränderten Office-Dokumenten in SharePoint

Wenn Sie die Metadaten eines Office-Dokuments in SharePoint Online ändern, werden die neuen Daten automatisch in den Dateiinhalt des zugehörigen Dokuments geschrieben.

Dadurch entsteht eine neue Dateiversion, die dann mit ELO synchronisiert werden muss.

Wird zur gleichen Zeit eine neue Arbeitsversion in ELO hochgeladen (z.B. von einem anderen Benutzer), wird ein Konflikt erkannt, der manuell gelöst werden muss.

Insbesondere sind die Metadaten innerhalb des Office-Dokuments in SharePoint Online unsichtbar, sodass es so aussehen kann, als ob das Dokument überhaupt nicht geändert worden wäre.

Referenzen in ELO

Aktuell werden Referenzen in ELO ignoriert, da es noch ungelöste Probleme im Umgang mit Änderungen von assoziierten Elementen gibt.

Konflikte mit dem automatischen Mapping

Wenn bei der Synchronisierung neue Elemente gefunden werden und diese automatisch auf bestehende Elemente abgebildet werden können, können für diese Elemente mehrere Konflikte entstehen, wenn sie keine übereinstimmenden Daten haben.

Dies ist eine Sicherheitsmaßnahme, da ELO Sync keine Ausgangsversion kennt und daher nicht vorhersagen kann, welche Daten erhalten bleiben und welche verworfen werden sollen.

> **Information**
>
> Es befinden sich verschiedene Methoden in Planung, die eine automatische Behandlung dieser Konflikte ermöglichen. Z.B. indem der neuere Datensatz verwendet wird.

Keine Dublettenprüfung bei der Archivierung

Werden neue Dokumente gefunden, die im ELO Repository archiviert werden sollen, so werden diese immer als einzelne Dokumente archiviert.

Es wird keine Dublettenprüfung durchgeführt und somit auch keine Referenzen erstellt.

Dies liegt zum Teil daran, dass es derzeit noch ungelöste Probleme in Bezug auf Referenzen in ELO und deren Handhabung in ELO Sync gibt.