

Spezifikation zur Übertragung von Transferpaketen in ein Hotfolder

Version 2.2

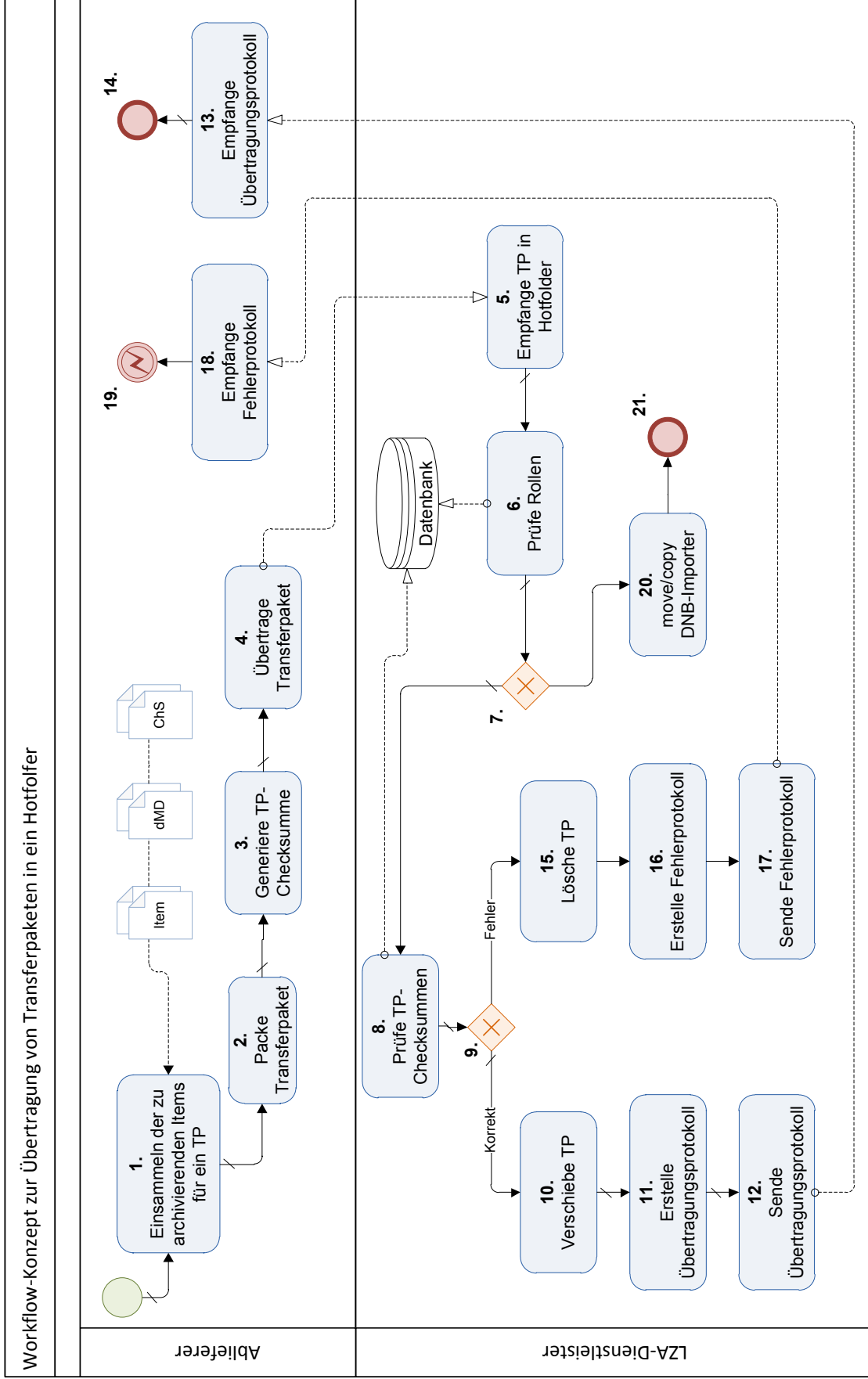
01.09.2010

Das vorliegende Dokument definiert die Übertragung von Transferpaketen an einen LZA-Dienstleister mittels eines Hotfolders.

Bitte beachten Sie, dass es sich bei diesem Dokument lediglich um die Spezifikation der Transferpakete und deren Übertragung von der abliefernden Institution an den LZA-Dienstleister handelt. Die nachfolgende Weiterverarbeitung der digitalen Objekte zur Aufnahme in das digitale Langzeitarchiv, findet in der separat definierten Spezifikation: „Spezifikation zur Übernahme von Transferpaketen in ein Langzeitarchiv“ [Übernahme 2010] statt.

Aufbau

Es wird ein von DP4lib entwickeltes Workflow-Konzept dargestellt, das die Übertragung eines Transferpaket, deren Prüfung sowie das geforderte Protokollwesen schematisch aufzeigt. Es soll vor allem als Diskussionsgrundlage dienen und verdeutlichen, welche Komponenten beim Transfer berücksichtigt werden müssen.



Beschreibung des Workflows

Schritt 1: *Einsammeln der zu archivierenden Items für ein Transferpaket*

Der Ablieferer stellt zunächst die zu archivierenden Paket-Items sowie deren zugehörigen Checksummen und die Metadaten-Datei zusammen.

Schritt 2: *Packe Transferpaket*

Der Ablieferer packt die unter Schritt 1 zusammengetragenen Daten in ein ZIP/TAR-Archiv, dieses ZIP/TAR-Archiv stellt das Transferpaket (TP) dar. Der Aufbau dieses Transferpakets ist der Beschreibung aus dem Glossar zu entnehmen.

Schritt 3: *Generiere TP-Checksumme*

Der Ablieferer erzeugt aus dem unter Schritt 2 generierten Transferpaket eine für die Absicherung der Übertragung genutzte Checksumme (TP-ChS). Das verwendete Verfahren muss vorher abgesprochen werden; die Checksumme wird in einer wie unter Schritt 2 gleichnamigen Notation geschrieben mit der Erweiterung md5 oder sha1.

Schritt 4: *Übertrage Transferpaket*

Der Ablieferer beginnt nach erfolgreichem Login mit der Übertragung des Transferpakets und der Checksummendatei in den Hotfolder des LZA-Dienstleisters.¹

Schritt 5: *Empfange Transferpaket in Hotfolder*

Der LZA-Dienstleister empfängt beide Objekte aus Schritt 4.

Schritt 6: *Prüfe Rollen*

Anhand der Verzeichnisse, in denen das Transferpaket vom Ablieferer eingespielt worden ist, werden eine Erkennung der Rollen und die Zuordnung nachfolgender Schritte durchgeführt (vgl. Rollen-Kennzeichnung im Glossar). Das Ergebnis der Prüfung wird in die Verwaltungsdatenbank geschrieben.

¹ Die im Rahmen von DP4lib angestrebte Testphase, könnte bei der Ablieferung Infrastrukturbeschränkungen (z. B. Bandbreite der Datenleitung) aufdecken. Für diesen Fall muss ein noch zu definierendes Sheduling-Verfahren z. B. auf der Basis von Timeslot besprochen werden.

Schritt 7: *Exklusives Gateway*

Abhängig vom Ergebnis der Prüfung aus Schritt 6 können zwei Wege beschritten werden:

1. Im Normalfall
→ weiter bei Schritt 8
2. Bei Sammelpflichtigen Dokumenten
→ weiter bei Schritt 20

Schritt 8: *Prüfe TP-Checksumme*

Der LZA-Dienstleister prüft das eingetroffene Transferpaket anhand des abgesprochenen Checksummenverfahrens. Das Ergebnis der Prüfung wird in die Verwaltungsdatenbank geschrieben.

Schritt 9: *Exklusives Gateway*

Abhängig vom Ergebnis der Prüfung aus Schritt 8 können zwei Wege beschritten werden:

1. Erfolgsfall: Wenn die TP-Checksumme als erfolgreich geprüft wird
→ weiter bei Schritt 10
2. Fehlerfall: Bei einem Fehler der TP-Checksummenprüfung
→ weiter bei Schritt 15

Schritt 10: *Verschiebe Transferpaket*

Das Transferpaket wird in ein Arbeitsverzeichnis (TEMP-Verzeichnis) des Import-Service verschoben. Ab hier beginnt im nächsten Workflow die Übernahme des Transferpakets in ein Langzeitarchiv. Die Übertragung des Transferpakets ist abgeschlossen.

Schritt 11: *Erstelle Übertragungsprotokoll*

Der LZA-Dienstleister erstellt nach einem vorher festgelegten Intervall ein Übertragungsprotokoll über das Transferpaket.

Schritt 12: *Sende Übertragungsprotokoll*

Der LZA-Dienstleister übermittelt das unter Schritt 11 erstellte Übertragungsprotokoll an den Ablieferer.

Schritt 13: *Empfange Übertragungsprotokoll*

Der Ablieferer empfängt das vom LZA-Dienstleister gesendete Übertragungsprotokoll.

Schritt 14: *Terminierung*

Die Übertragung ist erfolgreich beendet.

Schritt 15: *Lösche Transferpaket*

Das Transferpaket wird gelöscht.

Schritt 16: *Erstelle Fehlerprotokoll*

Der LZA-Dienstleister erstellt nach ein Fehlerprotokoll für das Transferpaket.

Schritt 17: *Sende Fehlerprotokoll*

Der LZA-Dienstleister sendet ein Fehlerprotokoll mit Informationen der fehlerhaften Checksummenprüfung.

Schritt 18: *Empfange Fehlerprotokoll*

Das Fehlerprotokoll wird auf Seiten des Ablieferers empfangen und ausgewertet. Hierfür enthält das Fehlerprotokoll eine Auflistung, welche Dateien die Checksummenprüfung nicht bestanden haben.

Das betroffene Transferpaket sollte vom Ablieferer korrigiert und erneut in den Ingest-Prozess überführt werden.

Schritt 19: *Fehlerverfahren*

Bei der Übertragung ist ein Fehler aufgetaucht. Ein Eskalationsverfahren wird notwendig.

Schritt 20: *move/copy DNB-Importer*

Transferpakete, die unter die Sammelrichtlinie der DNB fallen werden an den DNB-Importer weitergegeben.

Schritt 21: *Terminierung*

Damit endet der gesamte DP4lib-Workflow.

Quellen / Referenzen

[Übernahme 2010] Spezifikation zur Übernahme von Transferpakete in ein Langzeitarchiv V. 0.9, DP4lib.