

Übernahme (Ingest)

Inhalt

[Paketspezifikationen](#)
[Ingest-Prozess](#)
[Prozessdiagramm METS-Ingest](#)
[Prozessdiagramm OAI-PMH-Deposit](#)
[Prozessdiagramm CSV-Ingest](#)

Weiterführende Informationen

[Paketstrukturen](#)

[Spezifikation für Ablieferungspakete \(SIP\)](#)

[Transformation von Eingangspaketstrukturen zu SIPs und AIPs](#)

Paketspezifikationen

Der Ingest erfolgt automatisch. Im Ingest werden die verschiedenen Eingangspaketstrukturen mit einer dafür passenden Applikation in Pre-Ingest-SIPs transformiert. Wird das Pre-Ingest-SIP an Rosetta übergeben, wird dieses im Deposit-Prozess zu einem Post-Ingest-SIP transformiert. Daraus entsteht nach der erfolgreichen Weiterleitung in den permanenten Archivspeicher ein AIP.

Die verschiedenen Paketstrukturen sind in der Übersichtsgrafik [Paketstrukturen](#) erläutert.

Ingest-Prozess

Die Ingest-Workflows werden für jeden Bestand eines Teams/einer Datenquelle individuell konfiguriert. Die Workflows unterscheiden sich hinsichtlich der Paketeingangsstruktur, der Anzahl der Repräsentationen, des Ingestverfahrens, der Kataloganbindung sowie einiger Metadaten, die im Ingest-Prozess hinzugefügt werden:

- hinterlegte Lizenztexte
- vergebene Access Rights
- Kennzeichnung der Art der Bestände (IE Entity Type)
- Kennzeichnung für die Sammlung und den Produktionsweg (User Defined Field A)
- Kennzeichnung, ob passwortgeschützte Dateien im Paket vorhanden sind (User Defined B)
- Kennzeichnung, ob die Inhalte von einem als defekt erkannten Datenträger stammen und im Rahmen einer Datenrettung wieder zugänglich gemacht wurden (User Defined Field C)
- Kennzeichnung der Zugriffsmöglichkeit von außerhalb von Rosetta (Status)
- Bilden von Collections in Rosetta
- Erfassen von Metadaten über den OAI Harvesting Vorgang (Web Harvesting Section)

Die Bearbeitungsschritte in Rosetta sind auf Prozessebene identisch. Die Grafik [Transformation von Eingangspaketstrukturen zu SIPs und AIPs](#) zeigt detailliert den Transformationsprozess der [Eingangspaketstruktur](#) zu einem [Pre-Ingest-SIP](#), [Post-Ingest-SIP](#) und [AIP](#).

Auf technischer Ebene werden die Objekte entsprechend ihrer Formatanforderungen durch geeignete Plugins behandelt. Identifizierende, strukturelle und beschreibende Metadaten werden im Ingestprozess in der METS-Datei erfasst.

Während des Deposit bildet Rosetta eine Prüfsumme für jede Datei.

Während des Ingest werden im sogenannten Validation Stack verschiedene Prozesse durchgeführt:

- Formatidentifikation mit DROID
- Formatvalidierung mit JHOVE oder veraPDF
- Erstellen von drei Checksummen
- Gegenprüfen mitgelieferter MD5-Prüfsummen bei METS_ und CSV-Deposit
- Viruscheck
- Extraktion technischer Metadaten mit JHOVE, mediainfo, dem NLNZ Metadata Extraction Tool , exiftool oder veraPDF
- Validierung der METS-Datei (nur bei gewähltem METS-Deposit)

Neue Identifizierungs- und Validierungstools können jederzeit als Plug-ins in Rosetta eingebunden werden.

Die Ergebnisse des Validation Stack werden in die METS-Datei weggeschrieben und sind indexiert.

Die extrahierten technischen Metadaten werden als [Signifikante Eigenschaften](#) weggeschrieben.

Für alle Bestände erfolgt der Ingest automatisch.

Nach jedem Ingest eines Bestandes führt das Team Langzeitarchivierung einen Abgleich mit den Ergebnissen der Voranalyse durch, der folgende Schritte umfasst:

- Abgleich der Ergebnisse der Voranalyse mit den Ergebnissen der Identifizierung und Validierung in Rosetta
- Kontrolle der IEs, Repräsentationen und Files in Rosetta auf Vollständigkeit
- Kontrolle der korrekten Vergabe der workflowspezifischen deskriptiven, administrativen, rechtlichen, technischen und strukturellen Metadaten

Im Folgenden werden verschiedene Ingestverfahren beschrieben.

