

Ingest

Content

[Package specifications](#)
[The ingest process](#)
[Process diagramm METS deposit](#)
[Process diagramm OAI-PMH deposit](#)
[Process diagramm CSV deposit](#)

Further information

[Package structures](#)

[Specifications for submission information packages \(SIP\)](#)

[Transforming transfer information packages to SIPs and AIPs](#)

Package specifications

Ingest can be carried out automatically or manually. During ingest, the transfer information package is transformed to a pre-ingest SIP using a submission application. If the pre-ingest SIP is transferred to Rosetta, it is transformed to a post-ingest SIP in the deposit process. Following successful transfer to permanent archival storage, an AIP is created.

The various package structures are explained in the overview graphic [package structures](#).

The ingest process

Ingest workflows are configured individually for each library team. Workflows vary according to the transfer information package and metadata that are added during the ingest process:

- stored license texts
- Assigned access rights
- Identification of the type of inventory (IE Entity Type)
- Identification for the collection and the production path (User Defined Field A)
- Identification of whether password-protected files are present in the package (User Defined B)
- Identification of whether the content originates from a data carrier that was recognized as defective and was made accessible again as part of a data recovery (User Defined Field C)
- Identification of the possibility of access from outside Rosetta (Status)
- Creating collections in Rosetta
- Capturing metadata via the OAI harvesting process (Web Harvesting Section)

The processing steps in Rosetta are identical at the process level. The graphic Transforming transfer information packages to SIPs and AIPs shows in detail the process involved in transforming the [transfer information package](#) to a [pre-ingest SIP](#), a [post-ingest SIP](#) and an [AIP](#).

At the technical level, objects are handled according to their format requirements using appropriate plugins. Identifying, structural and descriptive metadata are recorded in the METS file during the ingest process.

During deposit, Rosetta creates a checksum for every file.

Various processes are performed in the validation stack during ingest:

- Format identification using DROID
- Format validation using JHOVE
- Crosscheck of three checksums
- Validation of supplied MD5 checksums for METS- and CSV-Deposit
- Virus check
- Extraction of technical metadata using JHOVE, mediainfo or the NLNZ Metadata Extraction Tool
- Validation of the METS file (if METS-deposit is selected)

New identification tools and validation tools can be integrated in Rosetta as plug-ins at any time.

The results of the validation stack are written to the METS file and are indexed.

The extracted technical metadata are written as [Significant Properties](#).

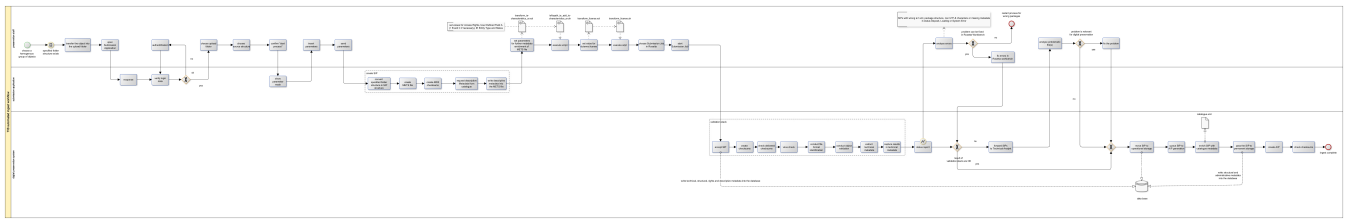
Ingest is carried out automatically for all holdings.

After every automatic ingest of a collection, the Digital Preservation team conducts matching with the results of the preliminary analysis, which involves the following steps:

- Matching the results of the preliminary analysis with the results of identification and validation in Rosetta
- Checking the IEs, representations and files in Rosetta for completeness
- Checking objects in Rosetta for completeness and the correct assignment of descriptive, administrative, legal, technical and structural metadata.

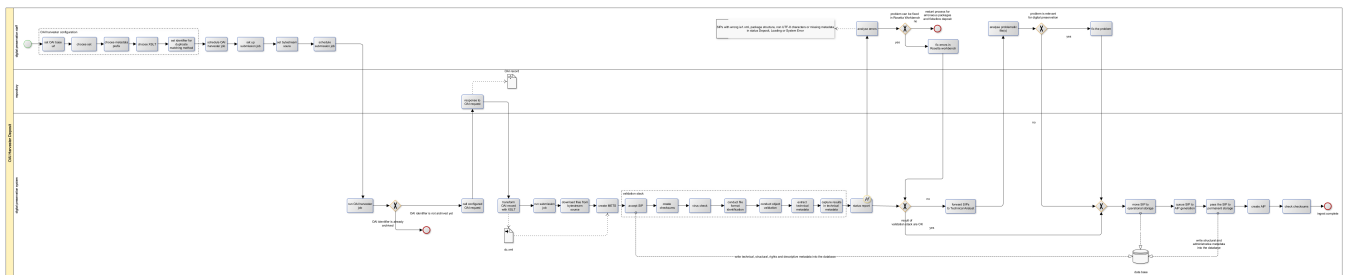
Process diagramm METS deposit

[Click to enlarge.](#)



Process diagramm OAI-PMH deposit

Click to enlarge.



Process diagramm CSV deposit

Click to enlarge

