Long Term Preservation – Info Session

Belspo – Eric Danon

belspo

Belnet
dedicated connectivity

.be

Ingest workflows

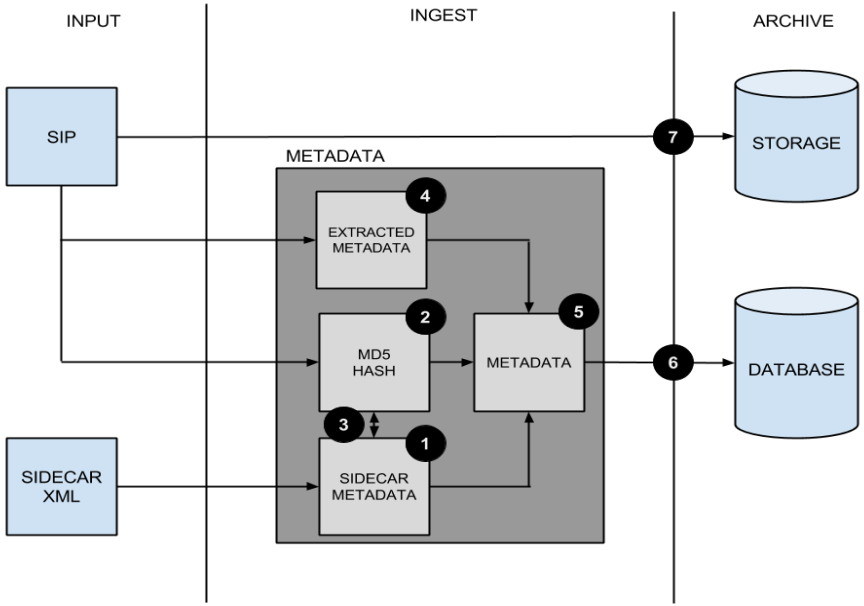
- Standard workflow
  - Digital File
  - Sidecar Metadata File
  - Online ingest

Default

- Advanced workflow
  - Archive packages
  - METS Sidecar Metadata File
  - Online ingest

Lots of small files

Standard Workflow: File or Archive + Sidecar



1. Parse & validate sidecar
2. Calculate MD5
3. Validate MD5
4. Extract the metadata
5. Merge the metadata
6. Create file & save metadata
7. Copy to the long term storage

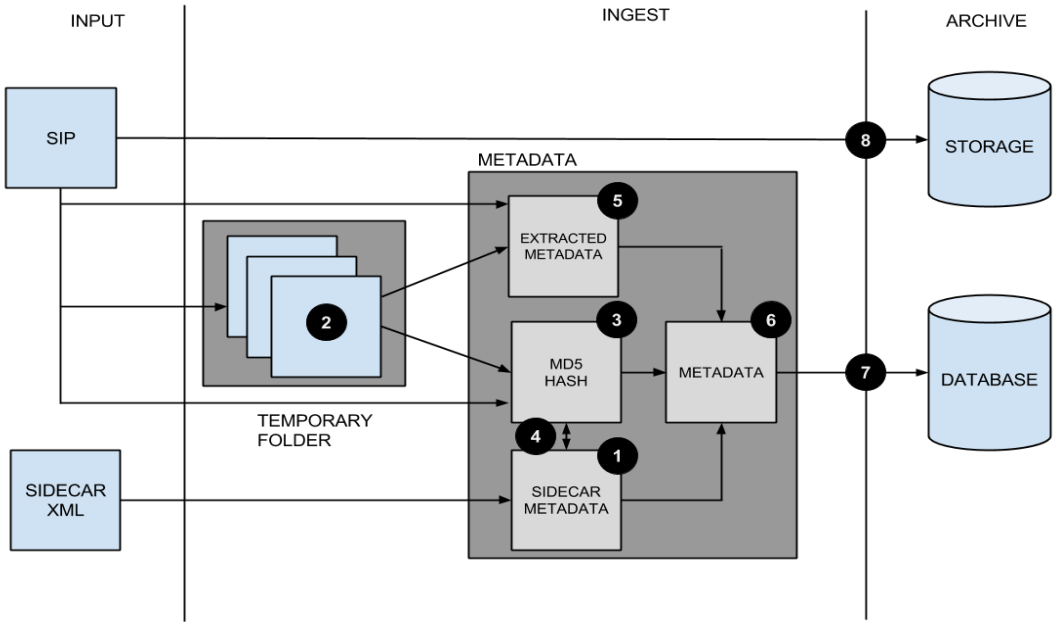
Standard Workflow: Metadata

- XML format
- Metadata
    - Title (or use the original filename as title)
    - MD5 checksum

- Restriction: XML must be encoded as UTF-8

```
<?xml version="1.0" encoding="UTF-8"?>
<MediaHAVEN_external_metadata>
   <title>My title</title>
   <md5>8e0781f932b527b4cbe24ff1a076a3b5</md5>
</MediaHAVEN_external_metadata>
```

Advanced Workflow: Archive + METS



1. Parse & validate METS
2. Unpack archive to temp folder

Perform steps 3→7 for each unpacked file + archive itself

3. Calculate MD5
4. Validate MD5
5. Extract metadata
6. Merge metadata
7. Create file & save metadata

8. Copy file to long term storage

Advanced Workflow: Uses (storage strategy)

- For each file in the archive you can provide a USE in the METS:
  - Virtual:
    - A file is indexed and validated in the platform
    - Can be found in the search
    - NOT stored as a separate copy
  - Preservation
    - A file is indexed and validated in the platform
    - Can be found in the search
    - Stored as a separate copy (i.e. can be exported without exporting the full zip file)
  - Fixity
    - A file is validated by the platform (checksum)
    - Can not be found in the search
    - No previews are generated

belspo
English
Welcome Zeticon Management module Edit password Logout
Start new search
Collections
Map
Favorites
Upload
My Basket (2)
Search terms:
Refine your search
Group (1)
16.1 (3)
Empty Group
Date (3)
2016 (1)
2015 (2)
2014 (2)
Media type (2)
document (4)
video (1)
GPS Coordinates (1)
0 (5)
Uploaded by (1)
bulkupload@16.1 (5)
Empty Uploaded by
Show: 50
Sort by: by Set
5 Results
PREVIEW NOT AVAILABLE
t14th8dfw
Creation date 28-09-2015
Rights Owner © amsab
Filename t14th8df5w.zip
Archive date 15-02-2016
Description Blaxxxxxx
t14th8dfw_pdf.pdf
Creation date 28-09-2015
Rights Owner © amsab
Filename t14th8df5w_pdf.pdf
Archive date 15-02-2016
Description Blaxxxxxx
t14th8df5w_mets.xml
Creation date 15-02-2016
Rights Owner © 16.1
Filename t14th8df5w_mets.xml
Archive date 15-02-2016
t14th8df5w_0001.dpx
Creation date 14-01-2014
Rights Owner © 16.1
Filename t14th8df5w_0001.dpx
Archive date 15-02-2016
t14th8df5w_0002.dpx
Creation date 14-01-2014
Rights Owner © 16.1
Filename t14th8df5w_0002.dpx
Archive date 15-02-2016
© Zeticon 2016 • Contact: info@mediahaven.com • Copyright • Manual • To Top
ARCHIVE
FILE IN ARCHIVE (USE: VIRTUAL)
FILE IN ARCHIVE (USE:PRESERVATION)
FILE IN ARCHIVE (USE:PRESERVATION)
FILE IN ARCHIVE (USE: VIRTUAL)
OTHER FILEs IN ARCHIVE (USE: FIXITY)
Original file
Searchable, not exportable (only as full archive zip file)
Searchable, exportable without downloading full zip
Not Searchable

Uploading: FTPS

The pair of data and sidecar file needs to be uploaded using FTPS to the LTP platform.

Server	LTP Ingest server
User	Standard Worklow: standard@<organisation> Advanced Workfloaw: advanced@<organisation>
Password	Sent to the local administrator of the organisation
Mode	Passive FTPS: Allow outbound FTP connections from your FTP client to LTP Ingest Server with port 21 (command) and between ports 50000 and 51000 (data) -> FW rules must be adapted!

Uploading: file name convention

The data file and the sidecar file must the same file root filename (filename without extension)

Example:

Data file	Sidecar file
KB_JB1156_1940-06-13_01.zip	KB_JB1156_1940-06-13_01.xml
FOD_NX_2000_NO_EXT	FOD_NX_2000_NO_EXT.xml

Uploading: limitation

The LTP platform has a maximum file size determined by the capacity of an LTO-7 tape medium.

!!! The maximum file size is ~ 5.5 TB. The ingest server will abort the connection if this limit is exceeded !!!

belspo

Belnet
dedicated connectivity

.be