

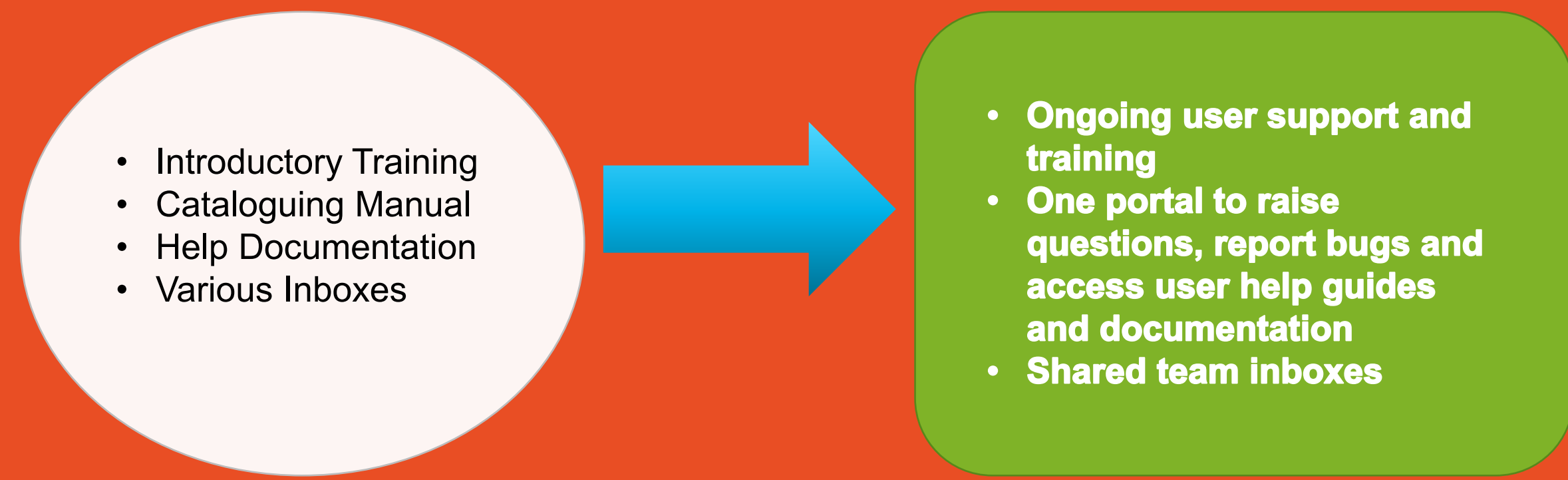


# Documentation at the BFI

Lucy Wales and Jennifer Macmillan

## How we got to where we are...

How did we manage user support and documentation before?

- 
- ```
graph LR; A((Old methods)) --> B[New methods];
```
- Introductory Training
  - Cataloguing Manual
  - Help Documentation
  - Various Inboxes

- **Ongoing user support and training**
- **One portal to raise questions, report bugs and access user help guides and documentation**
- **Shared team inboxes**

## Why is digital preservation documentation important to you and your organization?

- 1) WHY & HOW**
- 2) SUPPORT**
- 3) STANDARDS**

## What tools or platforms do you use to create and provide access to your documentation?

- **User Knowledge Base and Admin Knowledge Base**
- **CID and Digital Preservation Systems Service Desk**
- **JIRA Dashboard and Projects**

- **Axiell Collections**
- **GitHub Repository**
- **Videos**

# User Knowledge Base and Admin Knowledge Base

CID and DPI Systems / Service Desk / Digital Preservation Infrastructure (DPI) overview

## Digital Preservation Infrastructure (DPI) overview

### What is the Digital Preservation Infrastructure?

The Digital Preservation Infrastructure - also known as DPI - is the suite of integrated systems that the BFI Knowledge and Collections directorate has developed to preserve and provide access to the digital collections held in the BFI National Archive.

#### Further Reading

If you would like further information about Digital Preservation theory and practice, the Digital Preservation Coalition has some very useful resources:

- general - the Handbook: <https://www.dpconline.org/handbook>
- general - Novice to Know-How online course: <https://www.dpconline.org/digipres/train-your-staff/n2kh-online-training>
- audiovisual specific - Ashley Blewer's 'Pragmatic Audiovisual Preservation' guide: <https://www.dpconline.org/news/pragmatic-audiovisual-preservation-public-release>

A brief description of the systems involved:

#### Documentation: CID

- No ingest to DPI is possible without a CID record to describe the preservation file in its context
- Descriptive and some technical metadata about the digital objects (moving image, special collections, library) recorded in CID records
- All technical metadata for each digital file recorded in CID media record (coming soon)

#### Media Asset Management:

- Originally, Imagen was our Media Asset Management (MAM). Imagen was removed from DPI in March 2021.
- The DPI Ingest process is now managed by Python scripts. These are created and managed by the Data and Digital Preservation department.
- The functionality to download preservation files from data tape is being provided by the EON Browser tool on the BK-CI-Access server.
- A permanent replacement for Imagen's core functionality will be developed in Winter 2022 - Spring 2023.

Instant access storage: NAS servers - Isilon, QNAP and G-Rack

## Adding websites to the Allowed list in the Firewall



Owned by Collections Systems, created with a template

Last updated: just a moment ago by [Digital Preservation Systems](#) • 4 min read • 3 people viewed

As part of the Cyber Security being applied to DPI network, access to the Internet is being restricted. Basically all sites are blocked unless they have been added to the list of acceptable sites in the Firewall.

The instructions below detail how to add a site to the list.


### Instructions

#### Stage 1

1. A request will come in to the help desk asking for a site to be authorised. This request will need to be authorised by the Head of Data & Digital Preservation (Stephen McConnachie) or Digital Preservation Manager (Lucy Wales).
2. It is unlikely that the person requesting the website will know the URL. A google search outside of DPI will usually find the website.
3. To identify the exact URL for the licencing server, go to the support pages and search on Firewall.

# Service Desk

CID and DPI Systems / Service Desk












DPI Internet destination access request

Raise this request on behalf of\*

Collections Systems (collectionssystems@bfi.org.uk)

Reason for request\*

Normal text

**B** *I* ...         

Please describe the context for the request in relation to BFI National Archive work. For exam audio processing application, required by the application at start up. Application used for BFI I Show more...

Specific URLs\*

Please specify the full destination (including URLs) for approval. If not known, please provide qnap.com or support.qnap.com

Priority\*

Medium



Assign a priority to your request.

Share with\*


Share with BFI Users


Send


Cancel



Projects /  User Issues /  UI-5201

DPI Internet destination access request

 Link issue

 Add Checklist




  raised this request via **Portal**

Hide details

View request in portal


Description






File download for  restoration

Record number


[https://silversalt.fromsr...

Priority

 Medium


  1   

Waiting for customer

 Actions


Your pinned fields

Team member



 None

Details


Assignee

 Digital Preservation Systems


Reporter





Request Type

 DPI Internet destination access request

Knowledge base

 15 related articles

Approvers

# JIRA Dashboard and Projects

Default dashboard

Pie Chart: User Issues - Open issues

| Status               | Total Issues: 78 |
|----------------------|------------------|
| Waiting for support  | 37               |
| Open                 | 24               |
| Pending              | 11               |
| Waiting for customer | 5                |
| Waiting for approval | 1                |

Filter Results: User Issues - Open issues (Unassigned)

| T | Key     | Summary                                 | P | Status                | Components | Created   | Updated   |
|---|---------|-----------------------------------------|---|-----------------------|------------|-----------|-----------|
|   | UI-5201 | DPI Internet destination access request |   | WAITING FOR APPROV... |            | 13/Sep/23 | 13/Sep/23 |

1-1 of 1

🕒 1 minute ago

Filter Results: User Issues - Assigned to me (Active)

| T | Key     | Summary                                                         |
|---|---------|-----------------------------------------------------------------|
|   | UI-5200 | Record deletion request                                         |
|   | UI-5199 | Record deletion request                                         |
|   | UI-5198 | Workflow crashes when trying to add fields to view              |
|   | UI-5197 | Barcode not correctly recorded in acquisition record            |
|   | UI-5196 | Unable to edit compression code field on Axiell                 |
|   | UI-5193 | Record deletion request                                         |
|   | UI-5192 | Collections Search Crashing Causing WF Jobs to Fail on creation |
|   | UI-5188 | Record deletion request                                         |
|   | UI-5159 | Article Type not showing                                        |
|   | UI-5157 | Data change request                                             |

1-10 of 25

🕒 1 minute ago

Projects / DPI Systems Roadmap

Timeline

🔍

SR DS 1 +3

Status category

Epic

Label

Type

Sprints

August

2023 September

DSR-3 2023 Routine Tasks

DSR-16 2024 Routine Tasks

DSR-80 IIIF Universal Viewer - small scal...

DSR-85 Transition to Archive Cybersecu...

DSR-104 Confluence pages

+ Create Epic

DSR-104

Confluence pages

Description

Add a description...

Child issues

Order by

50% Done

DSR-32 Confluence Page (Admin) Re...

DSR-34 Confluence Page Firewall a...

DSR-2 Remove referen...

DSR-105 Confluence Pa...

DSR-108 Confluence Page Admin Cr...

DSR-31 Confluence Page (Admin) Se...

DSR-36 Confluence Page Prep...

DSR-33 Confluence Page Vide...

Assignee

Unassigned

Assign to me

Labels

None

Start date

None

Due date

None

Development

Create branch

Create commit

Releases

Add deployment

Reporter

Digital Preservation Systems

# Axiell Collections

Record details

Action

Clipboard

Occurrence

Features

–Identification–

Number

C-1575337

Item type

Digital

Can ID

5250755A

Status

Viewing

Usage

Digital file can be accessed on request

Collection

–Title–

| Title article | Title                              | Language | Type |
|---------------|------------------------------------|----------|------|
|               | PRESTON'S EMPIRE DAY MAY 22ND 1909 |          |      |

1

1 OCCURRENCES

–Conditions governing access and reproduction–

Access conditions

Before reusing BFI National Archive digital collections please ensure the required clearances from copyright holders, contributors or other stakeholders have been obtained for the specific use.

Date

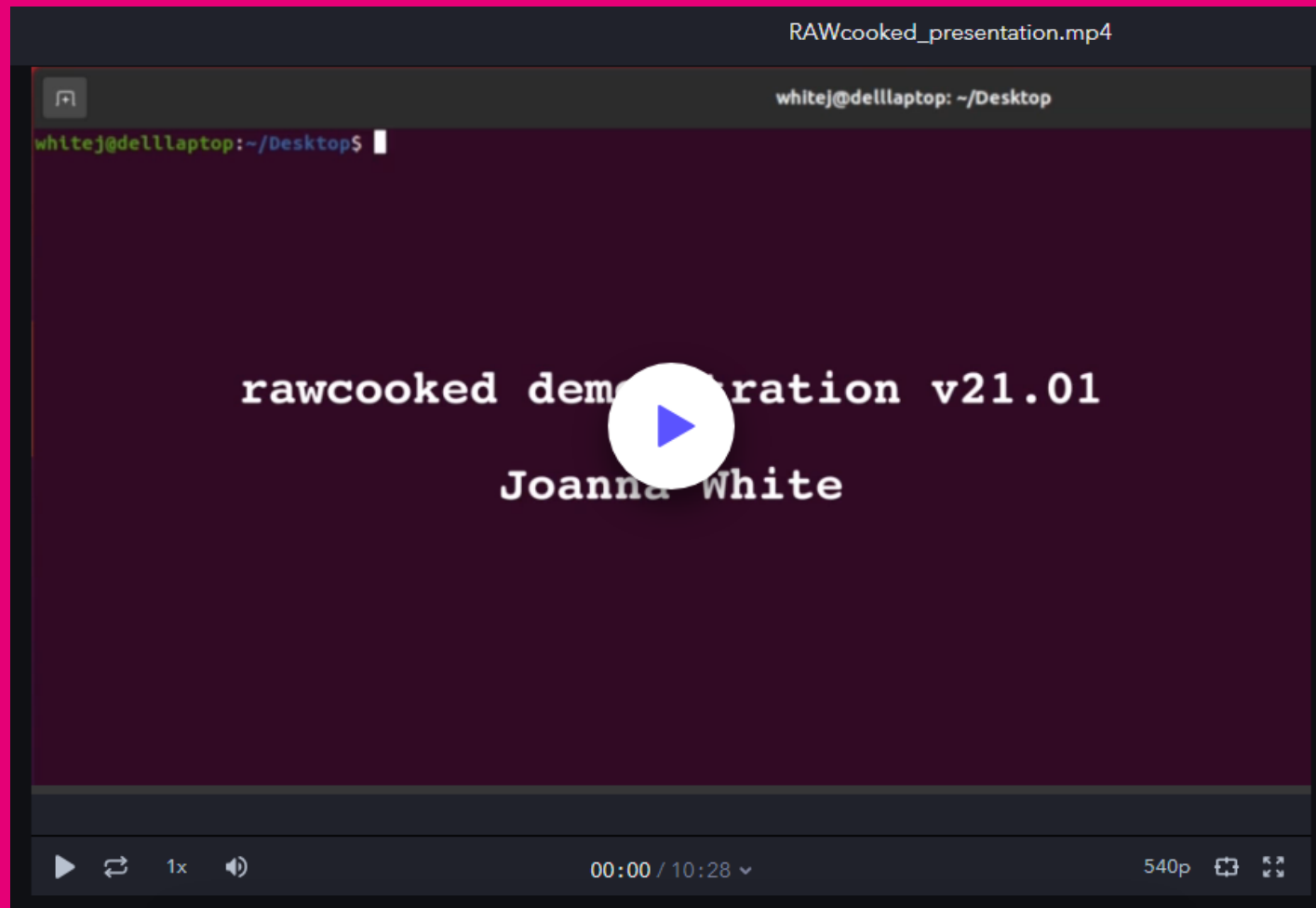
Help

Access conditions

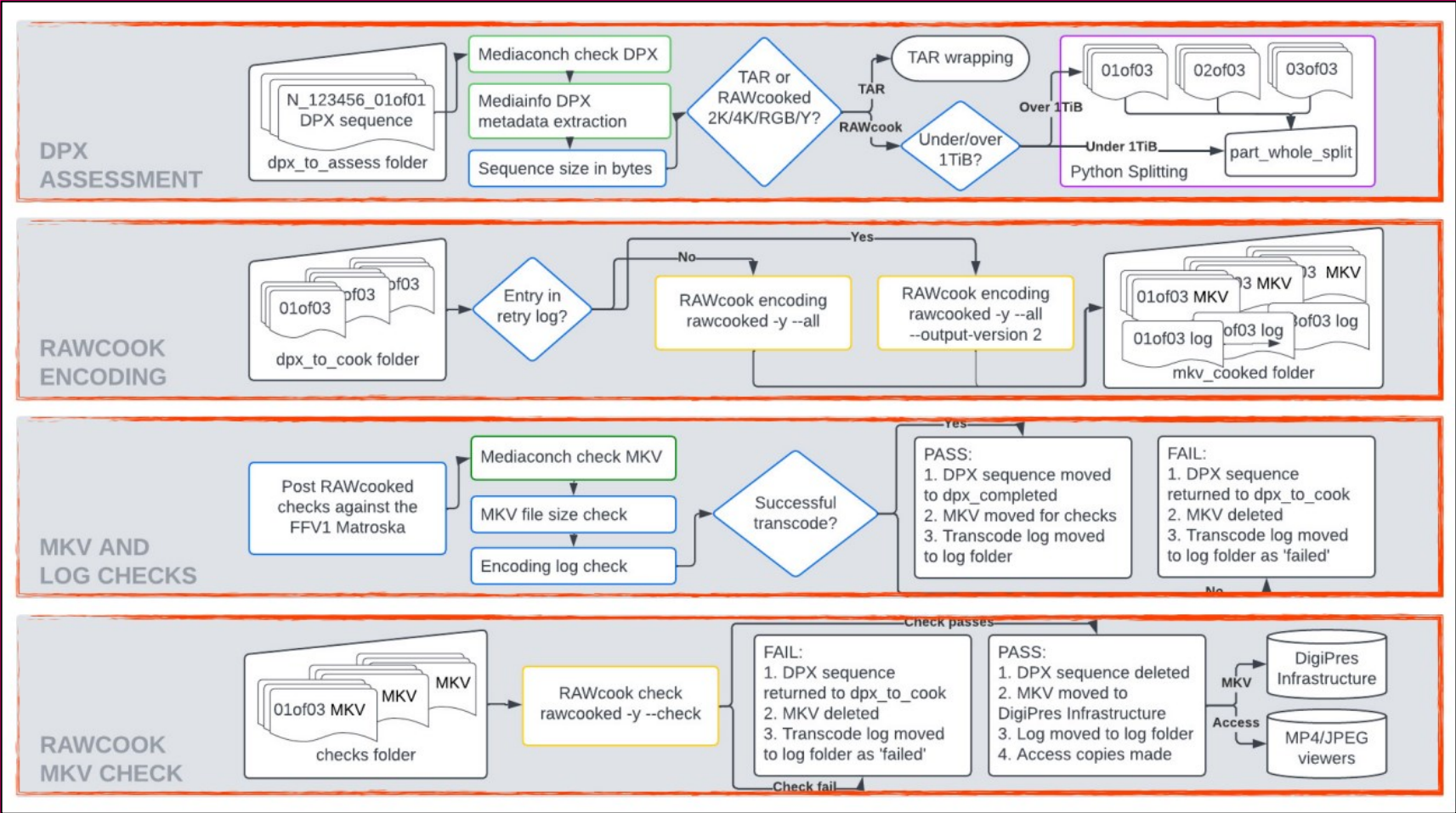
A statement on the level of access permitted to BFI staff and other stakeholders, as agreed by the Head of Collections Management. The access permissions are sometimes informed by the requirements of the lender, copyright holder or donor or are sometimes informed by the copyright conditions or condition of the item.

[https://github.com/bfidatadigipres/dpx\\_encoding/](https://github.com/bfidatadigipres/dpx_encoding/)

## Videos



# Lucid Chart



### Autoingest - Service Overview

Autoingest automates the verification, ingest, fixity and deletion of media from approved ingest locations

- About
  - Validation
    - Filenames must be correctly formatted
    - Media must be fully described in CID
    - Media must be valid
  - Ingest
  - Persistence
  - Visualisation
- Related articles

To learn how to use Autoingest, please read the [Autoingest - User Guide](#).

#### About

Autoingest is a digital preservation workflow service. It automates the validation, ingest, persistence and generate access proxies. It can handle files individually or in bulk.

Autoingest connects BFI Digital Preservation Infrastructure systems, processes and hardware together to

As the Autoingest service is always running, media is continuously queued for ingest and deletion. Media – is left in place, and will be re-queued and reprocessed repeatedly. Successfully ingested and persisted f

Autoingest records all media validation, ingest status and persistence status messages in a log file.

Media has to be documented in CID before it can be processed by Autoingest.

Every file proceeds through the following three stages before it is deleted: validation, ingest and persiste

#### Validation

Media must pass these three sets of validation tests before it is ingested:



Copy

Move

Export

Archive

Delete

Presenter mode

Analytics

Change page owner

Page history

Attachments

Resolved comments

Advanced details

Slack Notifications

Page Information

View Source

Hide related pages...

View Storage Format



Search

Enter

Code snippet

Display code with syntax highlighting

Info panel

Highlight information in a coloured panel

Quote

Insert a quote or citation

Decision

Capture decisions so they're easy to track

Divider

Separate content with a horizontal line

Expand

Insert an expand

View more

# Examples

## Error Log

The error log is created at 8am every morning. The log is called `current_errors` and it is located in the top level of the `autoingest` folder tree, in all operational shares in Isilon - Audio Operations, Digital Operations, Film Operations, Video Operations, Special Collections - as well as in the ingest share used by the Data team.

There is a CSV version. Check the error log to find out which files have not successfully ingested. You can filter on `path` to view only those rows that relate to the Isilon share where you work - ie excluding the other operational areas where you have no ingest responsibility.

This table below gives the error message, the meaning and actions to take to address the problem and progress the file in the next cycle of Autoingest.

| Error Message                                                                               | Meaning                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Action                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|---------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cannot parse partWhole from filename                                                        | This indicates poorly formed partWhole statement. Some common examples: <ul style="list-style-type: none"><li>missing underscore after Object Number</li><li>missing leading zero before a number below 10 - 1of6 instead of 01of06</li><li>presence of extra text such as 'Reel1' or 'R1' - R1ofR6 instead of 01of06</li><li>presence of trailing version increment - 01of01-2, 01of01-3, etc</li><li>part greater than whole - 05of03, 10of01, etc</li><li>first part is 00, and not 01</li></ul> | Change filename to use correct partWhole syntax making sure to use an underscore to separate Object Number, not a hyphen or space. Numbers up to 4 digits long are accepted, eg 0001of1900.<br><br>Also make sure to use a leading zero for a number below 10, and no extra characters such as Reel, R, Part, Pt.<br><br>Also make sure the part is not greater than the whole - ie 04of05 and 05of05 are fine, but 05of04 will fail.<br><br>Please make sure the 'of' is lowercase letters with no gaps between. |
| Extension does not match <file_type> in record                                              | This indicates a mismatch between the <code>file_type</code> in the CID record and the file extension. Some common examples: <ul style="list-style-type: none"><li>CID Item states MP4, file extension is .mov</li><li>CID record states JPG, file extension is .tif</li></ul>                                                                                                                                                                                                                      | In some cases, this error may indicate incorrect Object Number use in filename, so that should be checked first. Most commonly, the CID record must be changed, to document the file type accurately.                                                                                                                                                                                                                                                                                                             |
| ffprobe failed to read file: [1] status<br>or<br>ffprobe failed to read file: [None] status | This indicates a serious problem with the file's metadata headers, at a level that prevents it from being preserved in DPI. Some common examples: <ul style="list-style-type: none"><li>No video stream within the container</li><li>Fundamental metadata missing from container or video stream, which would compromise the file's usability</li></ul>                                                                                                                                             | In most cases this indicates a poorly encoded file, which must be re-encoded or re-acquired, as it cannot be ingested to DPI. Use MediaInfo to examine the file's metadata, and re-encode / re-acquire as appropriate.                                                                                                                                                                                                                                                                                            |

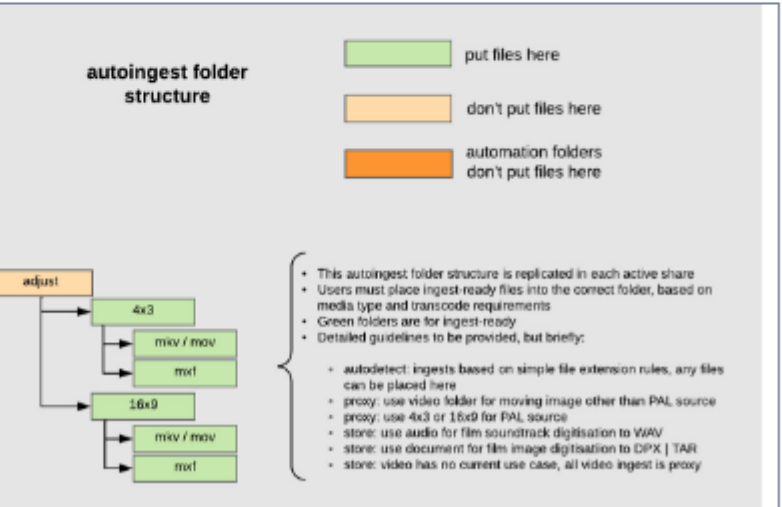
queued for ingest. Files which are successfully ingested are persisted as exact copies; additionally, access proxies

are supported. Files may simply be stored; or stored with an additional proxy made available; or stored with an

cks contains an `autoingest` directory structure. The preservation workflow used by Autoingest is no longer ingest directory structure.

the `video/adjust` or `proxy/image` subfolders, viewing proxies for video or image files will be made successfully. However, existing automated workflows still use different ingest folder paths so their structure is being retained for to continue their existing ingest practises, or simplifying ingest by using the `autodetect` folder.

files. The directory tree is mapped to `autoingest` by a configuration file. The tree is organised hierarchically and options for storing, transcoding and adjusting media.



## Examples

### Admin Knowledge Base

- Heritage 2022 Processes
  - Adding Imagen URLs to H22 Google Sheets (redundant)
  - Adding Proxy URLs to H22 RNA Google Sheets
  - Aspera - Building Aspera Reports in Aspera Console
  - Aspera Onboarding
  - Aspera System Support
  - BFI file checking sampling (MN WIP)
  - BFI Supplier Invoices
  - CID manual data input processes for H22 BFI Film Digitisation BFIPlayer titles
  - Delivering Files for BFI Replay (WIP)
  - Exceptions - Editing CID Item records
  - File Workflow
  - Film Digitisation Workflow and Processes
  - Framework Supplier allocation spreadsheet
  - Google Sheets - Data gathering for Google Sheets
  - Items digitised through shared container
  - Logging - WIP MN
  - MediaConch - FSN File Validation with MediaConch - WIP
    - MediaConch Failures - WIP
  - MediaConch Policies
  - Mediatheque Restrictions
  - Monthly Stats for H22 Access strand
  - Monthly Stats - NEW
  - Proxy supply to rightsholders (WIP)
  - Removing Items from Containers post digitisation
  - Returning files to RNAs
    - Calculating filesize for multiple HDDs:
    - CID import data
    - Copying files to HDD with 7-Zip Encryption (WIP)
    - Generating RNA filelist with checksums: (WIP MN)
  - Rights mp4 requests - download, rewrap and supply
  - RNA file requests - download and supply
  - Splitting WARNING / INFO Reports
  - Troubleshooting - Processing; Segmentation; Autoingest
  - Updating the Google Data Studio H22 Dashboard

# When do you create documentation and how often is it reviewed and updated?



Inception (2010), dir. Christopher Nolan

## What is the update process and how do you manage versions?



## What is next for digital preservation documentation at your organization?



## What tips do you have for people starting out on documenting their digital preservation activities?

- **Engage + research your audience**
- **Research systems for documentation**
- **Formatting is key**
- **Link to other articles**
- **Make documentation a key part of all of your processes, so it's regular and sustainable**



# Thank you

## Any Questions?

[lucy.wales@bfi.org.uk](mailto:lucy.wales@bfi.org.uk)

[jennifer.macmillan@bfi.org.uk](mailto:jennifer.macmillan@bfi.org.uk)