

Ingest User Manual and Workflow Guidelines

- [Overview](#)
 - [Preservation Events](#)
 - [Results](#)
- [Detailed Instructions](#)

Overview

This section provides information on the use of the BDPL Ingest Tool and overall workflow guidelines. Once the BDPL manager validates a manifest workbook, BDPL staff may begin using the BDPL Ingest Tool, a Python-based tool to assist with the initial ingest and creation of standardized Submission Information Packages (SIPs). The graphical user interface is designed to guide staff through transfer and analysis workflow steps as determined by the type of content. The tool generates log files for each preservation action and also records events in a basic Preservation Metadata: Implementation Strategies (PREMIS) XML file.

This project was inspired by and includes significant elements of [Brunnhilde](#) and [Disk Image Processor](#), both © by [Timothy Walsh](#) and released under an MIT License.

Preservation Events

The BDPL Ingest Tool employs a micro-service design to address four main job types:

- [Disk images](#): use cases involving digital material stored on physical media, including 5.25" floppies, 3.5" floppies, zip disks, and optical media.
- [Copy only](#): use cases where disk imaging is not appropriate or where content has arrived via email, network transfer, or download.
- [DVD](#): use cases where moving image content is stored as DVD-Video on optical media.
- [CDDA](#): use cases where sound recordings are stored as Compact Disk Digital Audio on optical media.

Each job type is comprised of two main steps: transfer and migration. Significant preservation events include:

- Transfer:
 - Disk imaging
 - ddrescue (production of raw images)
 - cdrdao (production of bin and cue files for CDDA use cases)
 - File replication
 - tsk_recover (file extraction from disk images with file systems that include ntfs, fat, exfat, hfs+, etc.)
 - unhfs (file extraction from disk images with file systems that include hfs and hfsx)
 - TeraCopy (replication of files in other use cases, including from optical media with ISO9660 or UDF file systems)
 - Normalization
 - cdparanoia (production of single .wav and cue files for CDDA use cases)
 - ffmpeg (production of one .mpeg per title for DVD-Video use cases, with content information provided by lsdvd)
- Analysis:
 - Virus scan: clamscan.exe
 - Sensitive data scan: bulk_extractor
 - Forensic feature analysis:
 - disktype (document disk image file system information)
 - fsstat (document range of meta-data values (inode numbers) and blocks or clusters)
 - ils (document allocated and unallocated inodes on the disk image)
 - mmls (document the layout of partitions on the disk image)
 - Format identification: Siegfried
 - Documentation of file directory structure: tree
 - Checksum creation: fiwalk or Python hashlib module (depending on use case)

Results

The BDPL Ingest Tool produces a standardized SIP as well as a report and documentation of ingest procedures. Each transfer item is identified by a unique barcode value, as described in the guidelines for [Shipments, Items, and Manifests](#).

A barcode folder has the following structure:

```
/[barcode]/
|
|__ /disk-image/ (if produced; will include bin and cue files for CDDA use cases)
|
|__ /files/ (will contain normalized versions of content for DVD-Video and CDDA use cases)
|
|__ /metadata/
|   [barcode]-dfxml.xml
|   [barcode]-premis.xml
|
|__ /logs/ (output of various preservation micro-services)
|
|__ /media-images/ (scanned photographs of media, if present)
|
|__ /reports/ (technical metadata related to original media and/or files; includes a version of
Brunnhilde html report)
```

High-level information about each object and its ingest process is also saved to the 'Appraisal' worksheet of the shipment's manifest workbook to assist collecting units with the review and appraisal of content before it is saved to the SDA to await final ingest and AIP creation procedures.

Detailed Instructions

Consult the following sections for more detailed information about the ingest process.

- [1. Receive Shipment](#)
- [2. Initiate an Ingest Session](#)
- [3. Transfer an Item](#)
- [4. Analyze an Item](#)
- [5. Scan the Physical Media](#)
- [6. New Item or Quit](#)
- [7. Making Notes on Items](#)
- [8. Troubleshooting](#)