

1

"Shotlist" Spreadsheet


	A	B	
1	unique_identifier	collection	item_title
2	32101047381338_1	Rita Guibert Collection of La Asturia:	
3	32101047381338_2	Rita Guibert Collection of La Asturia:	
4	32101047381346_1	Rita Guibert Collection of La Asturia:	
5	32101047381346_2	Rita Guibert Collection of La Asturia:	
6	32101047381353_1	Rita Guibert Collection of La Asturia:	
7	32101047381353_2	Rita Guibert Collection of La Asturia:	
8	32101047381361_1	Rita Guibert Collection of La Borges,	

A spreadsheet is prepared for the vendor by RBSC Tech Services. The purpose of this spreadsheet is to provide:

- Asset Inventory
Used by both Vendor and PUL staff
- Local ID (usually a barcode) of each asset
Used by vendor for filenames convention
- Component ID
Used to sync assets with parent component
- Description (optional)
Used in Quality Control process to make sure content matches description
- Item Title (optional)
For readability
- Collection Title (optional)
For readability

2


Vendor Returns Assets + HD




The vendor returns all physical assets, which are checked against the inventory "shotlist" to ensure all physical assets were returned.

The harddrive contains **one "Bag" per collection.**
The bag contains the following:

- bagit_info.txt file
- bagit.txt file
- manifest.sha256
- tag_manifest.sha256
- datafolder (no hierarchy)



Datafolder Contents



shotlist spreadsheet (copy of original)

ii. did-not-digitize spreadsheet

iii. 1 PBCore file per asset (Required for AV)

iv. 3 content files per asset side

a. tape1barcode_a.wav

b. tape1barcode_a.i.wav

c. tape1barcode_a.mp3

d. tape1barcode_a_wav_sha256.txt

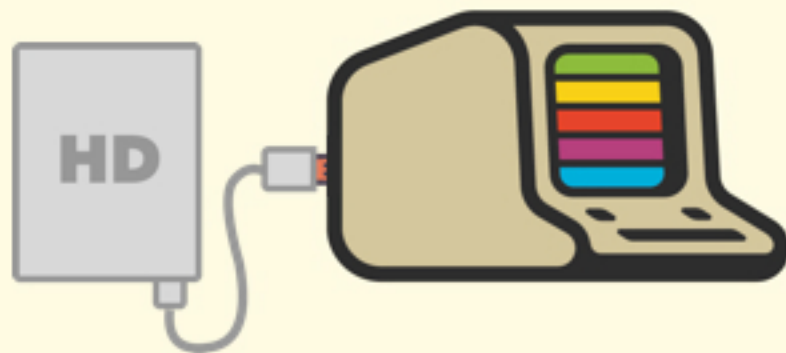
e. tape1barcode_a_i_wav_sha256.txt

f. tape1barcode_a_mp3_sha256.txt

v. 1 jpg per asset side (optional 2nd photo)

3

Copy to Staging



Step One

Run virus scan on HD

Step Two

Run rSync command that copies from HD to Isilon Staging



Step Three

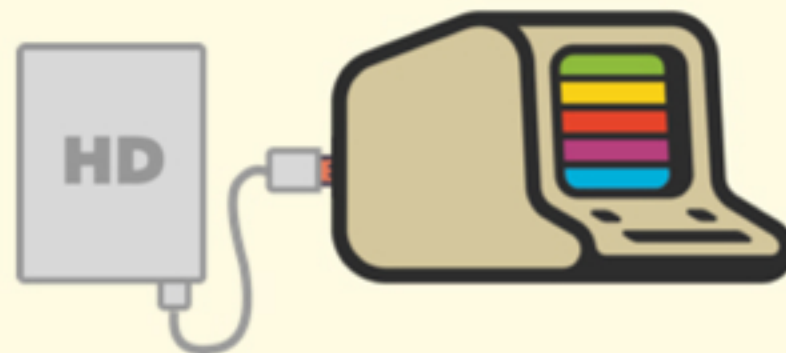
Staff runs Bagger/Bagit validation on Isilon Staging Bag copy via command line or Bagger GUI app.

Don't touch the staging bag!



4

Quality Control



To be completed before vendor deletes files, typically within 60 days:

- Physical QC - were assets all returned?
- Check for expected # of files for each asset.
(Consult the did-not-digitize list in bag)
- Verify that files are named by convention.
(Can be scripted using MediaInfo)
- Visually inspect waveforms for anomalies.
(Overkill? Scriptable?)
- Manually "spotcheck" of Preservation Master.
(Playback beginning, middle, end.)
- Verify that content matches descriptive metadata in PBCore.
- Check to make sure volume levels are OK.
(if not, check Transfer Notes in PBCore)
- Verify that Master and Access copies match.

If QC fails, go back to the vendor for new bags; delete bags from Isilon Staging, and begin again with Step 3.

5

Generate New EAD

After reviewing content, archivists may need to update descriptions or create new components.

- Update Description in "Shotlist" Spreadsheet (optional)
- Create new componenets, if necessary (optional)
- Generate new EAD from spreadsheet with <dao> links pointing to the filename of each asset. Multiple assets may or may not appear within a <daogrp>. PULFA will suppress any content that do not have <dao> links starting with "https://"
- Manually cut and paste new/modified components into existing EAD
- Run Normalizer to generate any new component ids
- Commit to SVN repo. The updates will be published to production via existing cron.

6

Import Collection into Figgy

Sync the EAD with the files in Isilon Staging so that components are linked with files. Import PBCore data for each asset.

Figgy Flow

Collection ID

AC045

☒ Sync from Pulfa

Path to bag

path/to/bag/on/isilon/staging

Sync

Preview for error checking

Ingest

7

Edit File-level Attributes in Figgy

Manage file permissions, as well as any structure, track listing/labeling, or reordering in Figgy.

NOTE: EAD structural changes to Figgy's published digital content is out of scope for the MVP. For example, a component with 50 tapes gets published, so we won't be able to break it up into multiple components without reingesting in the MVP.

8

Update EAD with ARKs

Manually run a script that updates the collection component <dao> links. The script will take a list of ARKs / ComponentIDs (generated by Figgy from published resources) as its input.

The updated EAD gets checked in and updated on PULFA via the cron. The viewer/content will appear on any components that have published