

Wisconsin Statewide Digital Archival Storage Instruction Manual

Review the Wisconsin Statewide Digital Archival Storage Procedures [Service Model](#) before proceeding.

- I. Choose files to store
- II. Prepare files for storage
- III. Package files for storage
- IV. Move files into storage
- V. Check files in storage
- VI. Retrieve or remove files from storage

Definitions

Depositor: An entity that can load, check, retrieve, and remove files from the storage system; for example, a public library system, a collaboration of multiple public library systems, a resource library, or Recollection Wisconsin.

Submitter: An entity that partners with a depositor to store files in the system; for example, a public library system's member libraries, or Recollection Wisconsin content partners. Each depositor determines how it will work with its submitters.

Before You Start

1. **Designate a single person to act as your storage administrator.**

Each depositor (e.g. public library system or collaboration of multiple systems) will be provided with access credentials for a single user who is authorized to upload, access, and remove files from the depositor's designated bucket in the storage system. Contact SCLS Technology staff to request, update, or transfer those credentials.

2. **Identify a processing workstation for your storage administrator to use.**

This workstation should have a fast, reliable Internet connection and at least 200GB of space available for temporary loading and processing of files.

Install these tools on the workstation:

- DART (see [section III](#))

- VPN client (see [section IV](#))
- S3 Browser (see [section IV](#))

3. Determine how you will track what you're storing.

A key piece of a well-managed digital storage program is maintaining up-to-date documentation of the files you have stored, who they belong to, and where they are located. Each depositor using the storage system is responsible for creating and updating their own documentation.

We recommend a two-part approach to documentation:

- **Digital Archival Storage Log.** Use a log to record information about each bag (package of files and related metadata) you move into storage. The [Digital Archival Storage Log template](#) can be downloaded and modified to meet your needs. This spreadsheet contains several different tabs for tracking files at different stages.
- **Folder for storage manifests.** Each bag you create will come with several text files, known as manifests, that contain important information about what's in the bag. Designate a location outside of the Dell ECS storage system where you'll save copies of these manifests for future reference.

I. Choose Files to Store

General Requirements

All content placed in the storage system should align with the [Recollection Wisconsin Collection Policy](#). Formats may include, but are not limited to, photographs, postcards, maps, letters, diaries, articles, books, artifacts, artwork, audio, film, and video. Content is not required to be made available online through Recollection Wisconsin at the time it is stored, but it should support Recollection Wisconsin's overall mission to provide access to digital collections from Wisconsin cultural heritage organizations.

Ineligible Content

The following types of content are out of scope:

- Content that is not included in the [Recollection Wisconsin Collection Policy](#), specifically:
 - Data-only records, such as birth, death or other genealogy indexes
 - Finding aids or EADs

- Institutional repository content, such as student dissertations, theses, or research data
- Content that is being preserved as part of a records management program (electronic records, meeting minutes, etc.). All content placed in the storage system is assumed to be stored indefinitely, and should not be subject to retention schedules.
- [Restricted or sensitive](#) content (HIPAA, personally identifiable information, etc.)
- Encrypted files

Depositors are responsible for communicating these requirements to their member libraries or other submitters. Questions or concerns about content eligibility will be reviewed by the Digital Archives Backup Collaboration Steering Committee of the Wisconsin Public Library Consortium. The WPLC Committee reserves the right to reject content or require its removal from the storage system.

Other Considerations

- **What are the local needs and priorities?**
If you're a public library system working with member libraries, what are those libraries' expectations and needs? For example, a grant-funded digitization project might require a specific approach to storage and preservation.
- **Are the files “done” and ready to store?**
A digitization project should be as “done” as possible before moving it into the dark archive. The storage system is not intended to be an active backup - while a project is in progress, be sure to **back up your files regularly in a separate location**, so you don't lose your work in the event of a hardware failure, accidental deletion, etc.
- **What is the risk if files are lost?**
Consider where your files fall on the Data Criticality Scale:

Data Criticality		Examples
1 (highest priority)	Only digital and we hold the only copy - if we lose it, it's gone forever.	<ul style="list-style-type: none"> ● Oral history interviews recorded digitally ● “Digital donations” (i.e. materials loaned for scanning by a patron, patron keeps the originals)

2	We have a digital copy, but physical versions are at high risk.	<ul style="list-style-type: none"> Digitized cassettes, VHS tapes, or other analog audiovisual media
3	We have a digital copy, but physical versions reside elsewhere.	<ul style="list-style-type: none"> Collaborative projects with partners, i.e. a local historical society
4	We have a digital copy <i>and</i> digital copies reside elsewhere.	<ul style="list-style-type: none"> Institutional partnerships, i.e. materials digitized for public libraries by University of Wisconsin Digital Collections, or historical newspapers scanned from microfilm by the Wisconsin Historical Society
5 (lowest priority)	We have a digital copy <i>and</i> still hold the original physical item.	<ul style="list-style-type: none"> Materials that can be readily located and re-scanned, if necessary.

Metadata Files

A file, such as a spreadsheet, csv, txt, or xml file, containing descriptive metadata for each item should be included with **every bag**. This is so that when a bag is retrieved in the future, the relevant metadata can be retrieved along with the digital files. The only time a bag will not have a metadata file is if metadata was not available at the time the materials were bagged.

Tip: For CONTENTdm users, TXT and XML metadata files downloaded from CONTENTdm should be packaged in every bag. See instructions for exporting metadata from CONTENTdm [here](#).

Whenever possible, content should be accompanied by item-level metadata in a csv file. You can download a template [here] to modify for your own use. Alternative but non-preferred formats include xls, txt, or xml file.

In addition to the csv file containing item-level metadata, you may want to describe the files you're submitting for storage on a broader level. Consider: what would you want someone fifty years from now to know about this set of files? This may be a collection-level metadata line you add to your csv file, or could be plain text in a .txt file that you bag and store with the files themselves.

Type and Format of Files to Store

Primary files

All file formats are accepted as long as they meet the criteria outlined above. The general recommendation is to store the highest-quality files you have available, in uncompressed, non-proprietary standard formats. Consider these your **primary files** (also known as archival files, preservation files, or master files).

In many cases, those primary files will be TIFF (for images and text), WAV (for audio), or MOV (for video). A good overview of recommended file formats and specifications is available from the [UC-Santa Cruz Libraries' Digital Initiatives Department](#).

Primary files and access files

Typically, access files, also known as derivatives, are copies that have been reduced or compressed for sharing online. In some cases, you might choose to store *both* **primary files** *and* **access files**. For example, if your primary and access files are already stored together locally, it might be more efficient to keep them together instead of reorganizing. Or, if your access files are cropped, enhanced, or otherwise edited, you might want to retain those alongside your unedited primary files. For more on primary vs. access files, see [Digitize: Bronze Level](#) in the Recollection Wisconsin Digital Readiness Toolkit.

Access files only

It's possible that you might not have primary files available for storing -- maybe they were created long ago and can't be located, or maybe they never existed. In that case, store the best version you have available to you. It's better to store a less-than-ideal file than to lose it entirely.

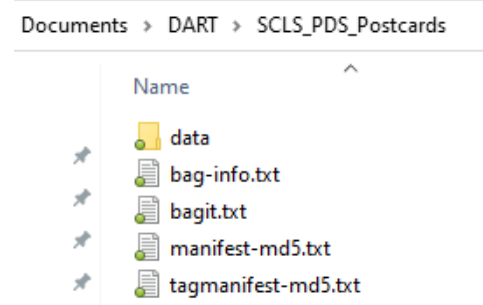
Unreadable, obsolete, or proprietary files

You may find that some of your files cannot be opened because they were created in formats that are now obsolete, incompatible with your current system, or dependent on a particular piece of software to read. These unreadable, obsolete, or proprietary files can still be stored! Again, it's better to store a less-than-ideal file than to lose it entirely. In the future, they can be converted on a case-by-case basis as time, resources, and tools allow. Indicate on your digital archival storage log if a file format is obsolete.

It is recommended to note if files are primary files or access files on your metadata spreadsheet, as well as if the files are unreadable, obsolete or proprietary.

II. Prepare Files for Storage

All files must be packaged into **bags** before they are placed in storage. A bag is a folder structured in a specific way (the [BagIt file packaging format](#)). Every bag includes a sub-directory, called “data,” that contains the files you will store, plus four text files that carry important information about the bag and its contents. You will bag up files and their metadata, then load that bag into the storage system.



Creating a bag is, more or less, the digital equivalent of placing physical archival documents in an acid-free box and making a record of the box’s contents. Packaging files into bags before moving them into storage will make it much easier to locate, check, move, and retrieve those files in the future.

This section describes the steps to organize your files so they’re ready for bagging. The process of creating the actual bags is described in section III.

This preparation step can be one of the most time-consuming parts of this process, but it’s also one of the most important. The better organized your files are before they go into storage, the easier it will be to locate a particular file needed in the future. The time invested in preparation now is a gift to your future self and colleagues. Keep in mind that if quality control of file naming, formats, and organization can be handled consistently during digitization, it will eliminate the need for a lot of preparation.

General Recommendations

- **Each bag should be around 100 GB in size, more or less.** This is for ease of uploading and file management, so a little over 100 GB is okay. If a submitter has files totaling more than 100 GB, split into multiple bags.
- **Group content in bags in a useful and logical way.** Remember that the bag is the unit that you will put in storage and what you will get out of storage -- you can’t pull individual stored files or folders out of a bag.
 - For example, a bag might be created for each of a library’s individual digital collections or projects.

- If you are handling content for lots of different submitters, you might want to keep each submitter's files in a separate bag.
- **Bag a copy of your primary files.** The Dell ECS system is a supplement to your local storage system, not a replacement for it. Don't remove your files from local storage in order to put them in shared storage - copy them, so they can live in both places. Remember LOCKSS - Lots of Copies Keep Stuff Safe. Preparing copies of the files also reduces the potential for introducing errors or modifications to the original primary files.
- **Have enough working space for preparing files.** If possible, the depositor should arrange for around 200 GB of free space on the administrator's local computer to serve as processing space. Each submitter's files can be removed from this space after they are bagged and moved into storage.
Tip: It is not recommended to use external hard drives for this purpose, as they have a short life and increase the risk of bit rot and other invisible damage to files.

Overview of File Preparation Steps:

- 1) Copy files onto your local workstation or network
- 2) Review file formats/types to include. Check for duplicates, invalid formats, etc.
- 3) Determine folder structure/organization
- 4) Generate item-level metadata file
- 5) Check for viruses

Folder Size and Structure

- Use standard directory organization with folders, sub-folders, sub-sub folders as appropriate.
- Bags should be about 100 GB, which means you should track the file size of each folder in your digital archival storage log so you know how many folders can be bagged together. If a single folder is larger than 100 GB, separate out into smaller folders.

Tip: You can bag much smaller folders individually, for instance if you want a single collection to be in a bag, but it's more time-consuming to upload folders that way - and will be more time consuming in the future to retrieve them. This decision will depend on the submitter, collection size, and future use recommendations.

Folder Naming

- Each folder should be named so that both the depositor and the submitter are identified. You may want to use a standard code or acronym, for example: SCLS_PDS or RW_DCHS. Add this standardized identifier as the prefix to all folders and sub-folders.

- Do not use spaces in folder or file names. The only punctuation can be underscore (_) or hyphen (-).
- File names within folders do not need the identifier.

Folders

- 📁 SCLS_ALB
- 📁 SCLS_BER
- 📁 SCLS_CIA
- 📁 SCLS_DCL

Sub-folders

- 📁 SCLS_ALB_AlbanysDistrictRecords
- 📁 SCLS_ALB_AlbanysSchoolYearbooksandAlumniListings
- 📁 SCLS_ALB_CemeteryListings
- 📁 SCLS_ALB_EarlySettlerandFamilyNarratives
- 📁 SCLS_ALB_HistoricalNarratives

[Advanced Renamer](#) is a good tool for bulk renaming folders on Windows computers. Mac computers have this capability [built in](#).

Prepared Directory Example

After completing the above steps, the directory should resemble the image below before proceeding to use DART. The SCLS_ALB folder includes eight sub-folders organized by project (with the identifier SCLS_ALB in every folder name) and two files (txt and xml) containing descriptive metadata for all items in each project.

Folders

- | Name |
|------------|
| 📁 SCLS_ALB |
| 📁 SCLS_ARP |
| 📁 SCLS_BER |
| 📁 SCLS_CIA |
| 📁 SCLS_COL |
| 📁 SCLS_CSP |
| 📁 SCLS_DCL |
| 📁 SCLS_MAR |
| 📁 SCLS_PDS |

Sub-Folders and metadata files

- 📁 SCLS_ALB_AlbanysDistrictRecords
- 📁 SCLS_ALB_AlbanysSchoolYearbooksandAl...
- 📁 SCLS_ALB_EarlySettlerandFamilyNarratives
- 📁 SCLS_ALB_HistoricalNarratives
- 📁 SCLS_ALB_LibraryHistory
- 📁 SCLS_ALB_LocalMemoirs
- 📁 SCLS_ALB_PhotographsandPostcards
- 📁 SCLS_ALB_WartimeNarrativesandMemori...
- 📄 CDM_Metadata_ALB_2022-09-20.txt
- 📄 CDM_Metadata_ALB_2022-09-20.xml

Virus Check

As a final step before creating bags, scan all folders for viruses. Use the anti-virus software your organization subscribes to, or use free anti-virus software such as ClamAV or AVG. If a virus is found, contact your IT department or follow your organization's protocols for quarantine or repair.

III. Package Files for Storage

Recommended Packaging Tool: DART

All files must be packaged into bags before they are placed in storage. The Digital Archivist's Resource Tool (DART), maintained by APTrust, is a free, simple open-source tool for creating bags.

Steps for getting started with DART:

- 1) Install DART on the computer where you will be creating bags and uploading them to the storage system. The current version, available for Windows, Mac, and Linux, is available here: <https://aptrust.github.io/dart-docs/download/>
 - a) Also save or bookmark the [DART documentation page](#) for future reference.
- 2) Configure a BagIt profile within DART. This generates some key information that will be applied to all of the bags you create. See [Appendix A: Configuring a BagIt Profile] for more information.

After you've installed DART and configured an institutional BagIt profile, you'll use DART to generate a bag for each of the folders you established in the file preparation step. Bags created using DART are temporarily saved on the computer where you're running DART before they are moved into the storage system. That means that you must have enough space available on your computer to hold the bag. We recommend keeping each bag around 100GB in size, so be sure to have at least 200GB of working space on the computer where you're creating the bags.

*Tip: **Do not** configure DART to place bags on an external hard drive, as this will increase the potential for corruption or errors.*

Tip: When you are using DART, it may be helpful to temporarily disable any local software that runs automatic backups or syncing of files (like Carbonite or Dropbox). The software may initiate creating a back-up and you will not be able to delete the bag until the backup is completed. Not disabling the software may slow down the workflow.

Create a Bag

To create a bag in DART:

- Click **Jobs >> New**.
- Drag and drop the items you want to package. Include folder(s) plus associated metadata files.
- Click Next.

Files

Drag and drop the items you want to package.



File Path	Directories	Files	Total Size	
D:\BackUpProject\SCLS_PDS\Postcards	21	40	720.83 MB	✘
D:\BackUpProject\SCLS_PDS\CDM_Metadata_PDS_2022-09-08.txt	0	1	1.94 MB	✘
D:\BackUpProject\SCLS_PDS\CDM_Metadata_PDS_2022-09-09.xml	0	1	2.83 MB	✘
Totals	21	42	725.61 MB	

Delete

Next >>

- In **Jobs >> Packaging**, set up the following configuration:
 - Package Format = BagIt
 - BagIt Profile = Name of the BagIt profile you set up earlier
 - Example: *SCLS Digitized Resources*
 - Serialization = Leave blank
 - Package name = Your chosen file prefix plus file or folder name. This will become the name assigned to your bag.
 - Example: *SCLS_PDS_Postcards*
 - Output Path = This will automatically populate with the drive location set in your profile, plus the package name. This is where your bag will be placed after it is created.
 - Example:
C:\Users\scatramski\Documents\ Dart\SCLS_PDS_Postcards
- Click Next.

Packaging

Note: Although this page displays a number of serialization formats, DART currently supports only the "None" and ".tar" formats. Additional formats will be coming soon.

Package Format

BagIt Profile [?](#)

Serialization [?](#)

Package Name

Output Path

<< Back

Next >>

- In **Jobs >> Bag Metadata**, use the following configuration:
 - External-Identifier = Should be identical to the Package Name on the previous screen.
 - Example: *SCLS_PDS_Postcards*
 - Library Name and Address = Add the official name, address, or any other basic details that will help identify the individual library or other submitting institution(s) associated with the bag.
 - Example: *PDS=Ruth Culver Community Library (Prairie du Sac)*
 Tip: If you'll be creating multiple bags for each submitter, create a list of library names to copy/paste into this field in order to save time and improve consistency.
 - Click Next.

Bag Metadata

Add New Tag Show All Tags

[bag-info.txt](#) *Tags with default values are not showing.*

External-Identifier *

Library Name

[bagit.txt](#) *Tags with default values are not showing.*

<< Back
Next >>

Alternatively, if you did not create a BagIt profile, or you need to change any information in the profile, you can input more detailed metadata for the bag at this step.

- To input more detailed metadata or to edit metadata, In **Jobs >> Bag Metadata**, use the following configuration:
 - In the upper right, click Show All Tags

- Bag Size, Bagging Date, and Bag Software = Leave blank. These will automatically populate later in the process.
- Contact Email, Contact Name, and Contact Phone = Enter the submitter's name (library or content partner staff), email address, and phone number. This is in case someone needs to retrieve contact information for this particular bag in the future.
- Bag-Count = Leave blank
- Bag-Group-Identifier = Leave blank
- External description = Should automatically populate.
- External-Identifier = Use the package name here that you set up earlier.
 - Example: *SCLS_PDS_Postcards*
- Library Name and Address = Add the official name, address, or any other basic details that will help identify the individual library or other submitting institution(s) associated with the bag.
- BagIt Version = 1.0
- Tag-File-Character-Encoding = UTF-8
- Click Next.

Bag Metadata

Add New Tag

Hide Default Tags

bag-info.txt

Showing all tags.

Bag-Size * ?

Bagging-Date * ?

Bagging-Software ?

Contact-Email *

Contact-Name *

Contact-Phone *

External-Description *

External-Identifier *

Library Name

Organization-Address *

Payload-Oxum * ?

Source-Organization *

bagit.txt Showing all tags.

BagIt-Version * ?

Tag-File-Character-Encoding * ?

<< Back

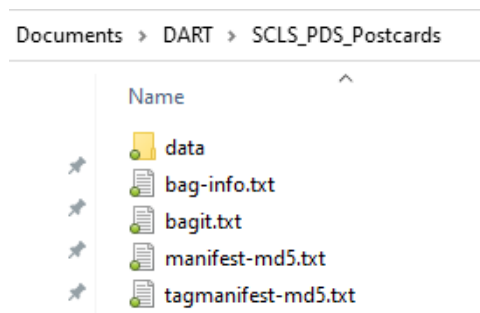
Next >>

- In **Upload Targets**, click Next.
- Click Run Job.

Verify the Bag

- Locate the saved bag on your system and open it.
- Do a visual check to confirm that the following folders and files were created:
 - data folder - this is where your files for storage are gathered
 - bag-info.txt
 - bagit.txt
 - manifest-md5.txt
 - tagmanifest-md5.txt
- Open the data folder and visually verify that it contains what you intended to bag. There is no need here to do a close examination of each file, just glance in the folder to confirm the number of files and folders, the file types, file names, etc. appear correct.
- Open the bag-info.txt document and verify the following:
 - Packet Size: In MB or GB, should match the total size of the files you bagged - check against the DART log in your project spreadsheet
 - BagIt-Version: 1.0
 - Tag-File-Character-Encoding: UTF-8

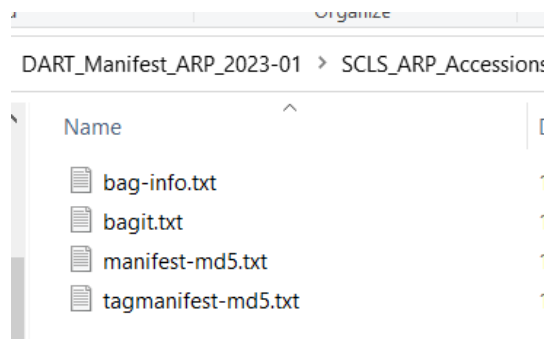
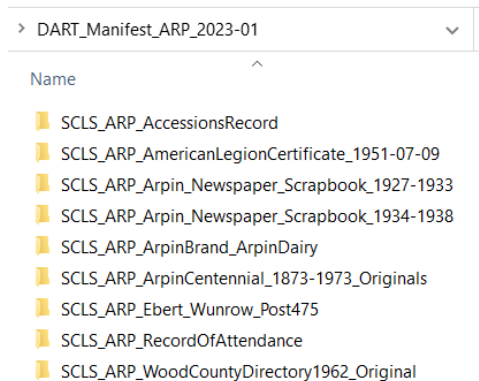
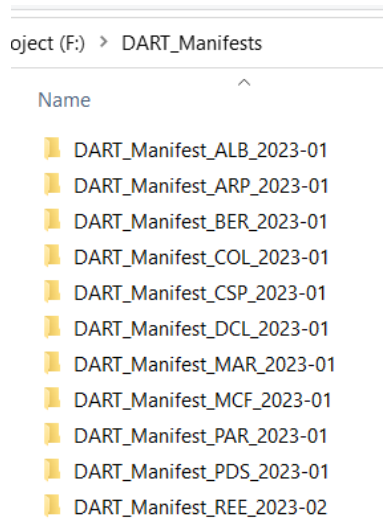
Example image: The data folder and four txt files together are the bag.



Log the Bag and Save the Manifests

- Record in the DART Log tab of your digital archival storage log that this bag has been created in DART, indicating that it is ready to be moved into storage.
- Open the manifest-md5.txt document and make a note of the file types (jpg, tif, etc.) contained in the bag. This data may help identify future migration needs.
- Copy the four text files (bag-info.txt, bagit.txt, manifest-md5.txt and tagmanifest-md5.txt) to the unique folder on your local system dedicated to manifests for this project. This will allow you to easily access the information contained in the manifests if you need it in the future, without having to access the entire bag in the storage system. **Do not rename or modify these files.**

- Create a unique folder on your local system dedicated to manifests. There can be one top-level folder (example: DART_Manifests) with a sub-folder for each submitter. Folders within each sub folder should be manually renamed to match the name of the bag.



Tip: The data folder and the four text files will always have the same names. **Do not modify these names; they are important for future validation.** Because they do not have unique names, it is important that the folder where you save them locally *does* have a unique name.

Troubleshooting DART

If there's a validation failure (bag does not validate in DART):

- Check the amount of memory available on your computer. If there is not enough memory, bagging may fail.
- Close DART and retry bagging.

If bagging or validation goes on for a longer than normal period of time:

- Check the folder size. Large folder sizes (greater than 100 GB) may cause this to happen. Break the folder into several smaller folders.

IV. Move Files into Storage

After you have created and validated your bags, you can begin uploading them into your institution's assigned bucket within the storage system. These data transfers will be made by connecting to the Dell ECS system over a VPN and uploading through an S3 browser.

SCLS Technology staff will provide access credentials and separate documentation for connecting to the Dell ECS system. You will need to install a VPN client such as [FortiClient](#) (free) and the [S3 Browser](#) (freeware version available, but the \$40 pro version is highly recommended because it will significantly increase data transfer speeds).

Moving data into the Dell ECS system can be time consuming, so we recommend that you only start an upload when the computer can be left on indefinitely until the upload is completed. It is likely that all other activity on the computer will be slower during the upload process.

Tips:

- If possible, using a direct Internet connection is recommended.
- Uploading more than one bag at a time is an option. However, doing so is not recommended. It can get messy tracking the progress status when multiple bags are being uploaded at one time.

Uploading a Bag

- Authenticate through the VPN.
- Open the S3 browser.
- Open Dell ECS platform.
 - Select your bucket (Example: SCLS_Archive)
 - Important: Highlight the save (level/location) where the folder should be added.
 - Save location can be a sub-folder of an existing folder.
 - Select Upload, Select Folder (the bag that was created by DART)
 - The upload will start immediately after selecting the folder.
 - Verify Failed = 0

The bag is initially uploaded to the SCLS data center in Madison. It will be replicated to the LEAN data center in Eau Claire within 24 hours. ECS administrators will get alert emails if a replication fails and will share any alerts about any failing replications.

After Upload is Completed

- Check the Properties tab in the S3 browser for details about the upload. The information here can be copy/pasted into your Digital Archival Storage Log. Information from the DART manifest files (see section 5. Bagging) can be compared to information in these fields.
 - Recommended fields to log include:
 - File Types
 - Total Size
 - Total objects/files/folders/

Property	Value
Creation date	11/28/2022 10:57:41 AM
Location	Default Region (us-east-1)
Total objects	47220
Total files	43975
Total folders	3245
Total size	833.00 GB (894 426 829 103 bytes)
Logging	failed - NotImplemented: The requested functionality is not implemented.
Versioning	Disabled
Cross-region replication	failed - Forbidden: The remote server returned an error: (403) Forbidden.
Transfer Acceleration	Not supported
Default Storage Class	STANDARD
Server-side encryption	Disabled
Requester pays	failed - NotImplemented: The requested functionality is not implemented.
File types	Text Document, XML File, TIF File, TIFF File, Adobe Acrobat Document, JPG File, Shortcut, Wave Sound, MPEG-4...
Server-side modified	12/7/2022 10:23:00 AM - 2/7/2023 12:15:44 PM
Owners	sclstr (sclstr)
Storage classes	STANDARD

Note: You will see several fields (Logging, Cross-region replication, and Requester pays) marked as “failed” in the Properties tab. You can ignore these! They are the result of some features that are available in the Amazon S3 environment (which the S3 browser was developed to navigate) but not available in the Dell ECS environment.

- After a bag has been uploaded, the job can be deleted from the DART software.
 - Select List from the Jobs tab.
 - Click on a job name.
 - Select Delete.

The screenshot shows the DART software interface. At the top, there is a navigation bar with 'Dashboard', 'Settings', 'Jobs', 'Workflows', and 'Help'. The 'Jobs' tab is active, and a dropdown menu is open, showing 'List' (circled in red) and 'New'. Below the dropdown is a 'Validate a Bag' button and a blue 'New' button. The main area displays a table of jobs:

Name	Created	Packaged	Validated	Uploaded
SCLS_VER_Directories_1950-1961	2023-01-25T18:11:05.439Z	1/25/23 12:38 PM	1/25/23 2:16 PM	-
SCLS_VER_Directories_1916-1949	2023-01-25T16:24:18.425Z	1/25/23 10:42 AM	1/25/23 11:51 AM	-

- If you have not done so already, save a copy of the text files from the bag folder (see section 3 of Bagging). At this point, you can also delete the bag and the files that had been saved locally for the bag from your working computer.

Tip: Remember to empty the recycle bin after deleting a bag, since it takes up so much space!

V. Check Files in Storage

The functions of the [Dell ECS system](#) include automatic validation and healing. In addition to this automatic monitoring, it's a good idea to run occasional manual checks on your files to confirm the data has remained unchanged. Each depositor is responsible for running these checks on their own schedule. We recommend spot-checking a selection of your stored files (or, if time and resources allow, a full check of all files) every 12 months.

Checks within the ECS System

A very quick, minimal way you can check your files is to review bag properties within the storage system.

- Authenticate through the VPN, open the S3 browser, then open Dell ECS platform.
- Check the Properties tab in the S3 browser to view details about the bag. The information here can be compared to the information from the original upload in the digital archival storage log.
 - Recommended fields to compare include:
 - File types
 - Total size
 - Total objects/files/folders

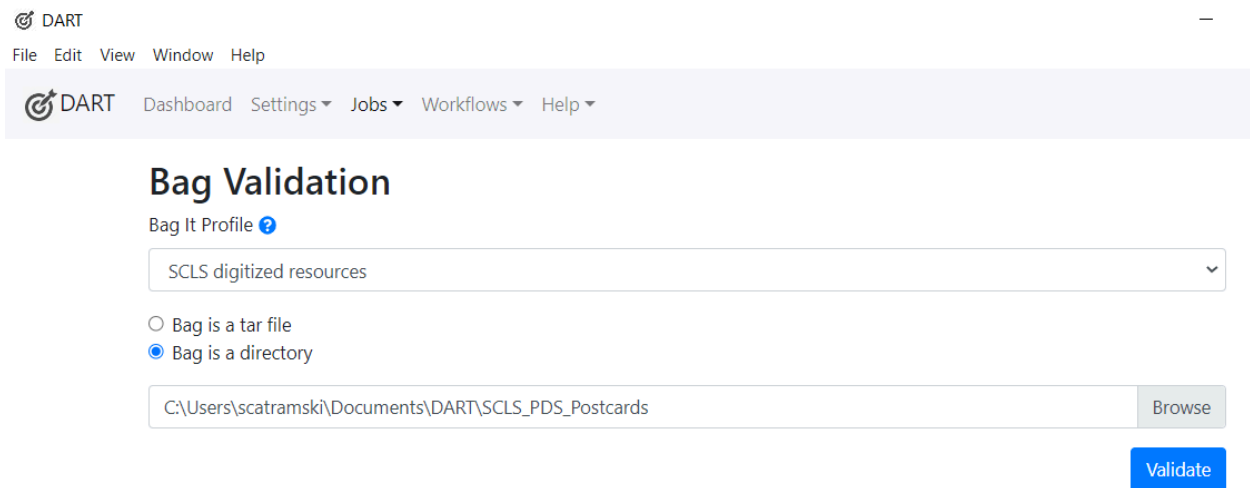
Using DART to Audit Files

A more thorough and accurate way to check your files is by using DART or another checksum validator to audit the integrity of your files. For more information on file fixity and integrity checks, see the [Store and Maintain section of the Recollection Wisconsin Digital Readiness Toolkit](#).

To perform a spot-check using DART, choose several bags in storage to download and re-validate.

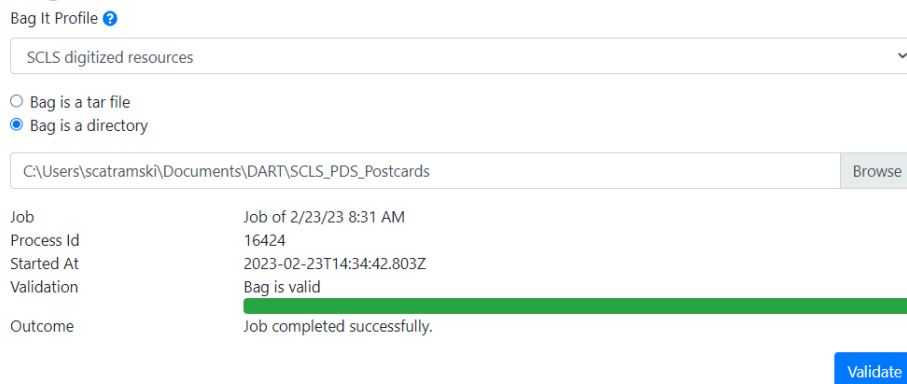
- See download instructions in Section VI.

- Validate the downloaded bag using DART.
 - Open DART
 - Select **Jobs >> Validate a Bag**.
 - Ideally, all bags would be periodically downloaded and verified. You can choose to do all bags at the same time, or do several per day or week over a period of time. This practice should be repeated annually.
 - BagIt Profile: Select your profile
 - Select: **Bag is a directory**
 - Browse: Select the folder that was downloaded and **Upload**.
 - Select: **Validate**



- DART should tell you that the bag is valid; the Outcome screen should say that the job was successfully completed.

Bag Validation



If the validation fails, first re-download the bag from the Dell ECS platform and attempt to re-validate.

If the bag still does not validate properly, continue checking other bags in the Dell ECS platform. You may need to re-upload the original files if there has been a failure.

Log the spot-checked files and results in the Quality Control tab of your digital archival storage log.

VI. Retrieve, Remove, or Update Stored Files

Retrieving Files

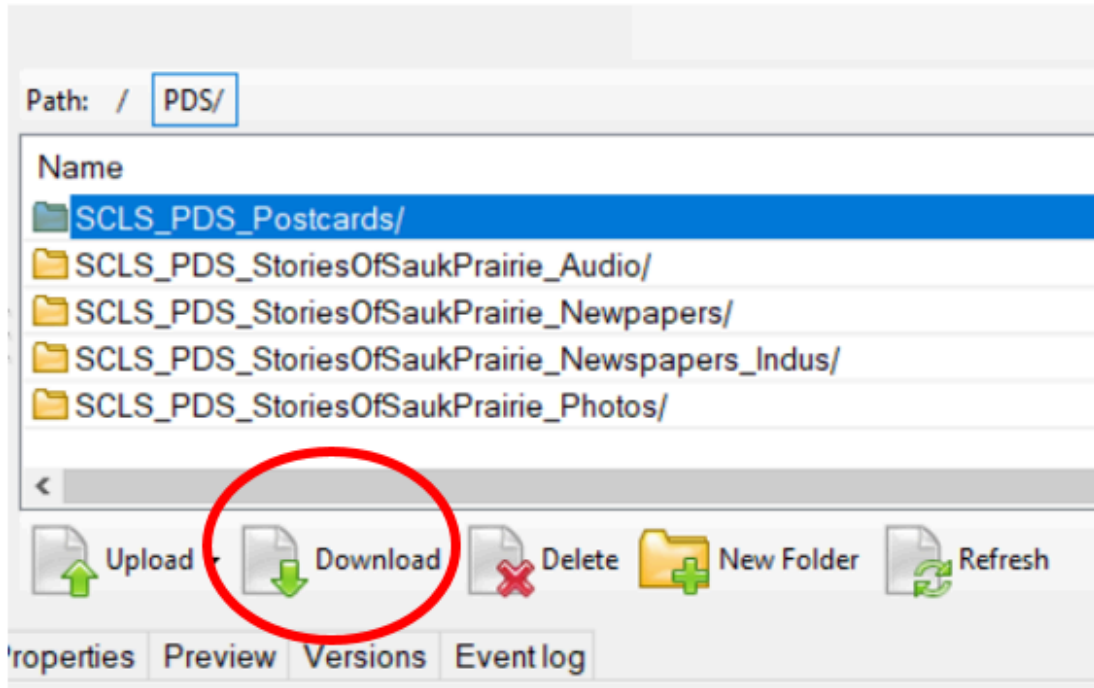
In the event of a disaster or other local data loss, retrieving bags from storage may be necessary in order to replace a submitter's local files.

- Authenticate through the VPN
- Open the S3 browser
- Select the bag to download.

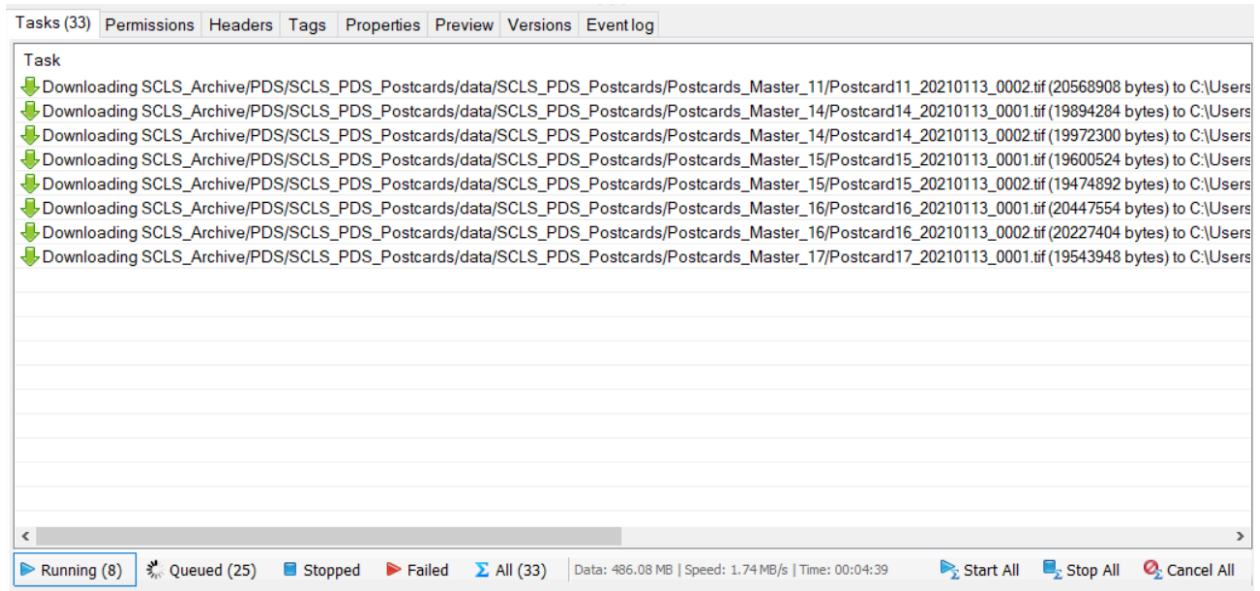
Tip: Check the total size of the bag by selecting the Properties tab to determine how much memory will be used by the download. Verify there is that much memory available on the save-to location.

Tip: **Downloading does not remove the bag from the Dell ECS system.** The bag remains in place and a copy is downloaded.

- In the pop-up window: Select Folder: Determine save location
 - The save location should be your local workstation. After re-validating using DART and virus-checking (see steps below), the bag's contents can be delivered to the submitter using your preferred file transfer method (i.e. FTP, external hard drive, etc.).
- After clicking Select Folder, the download will begin.

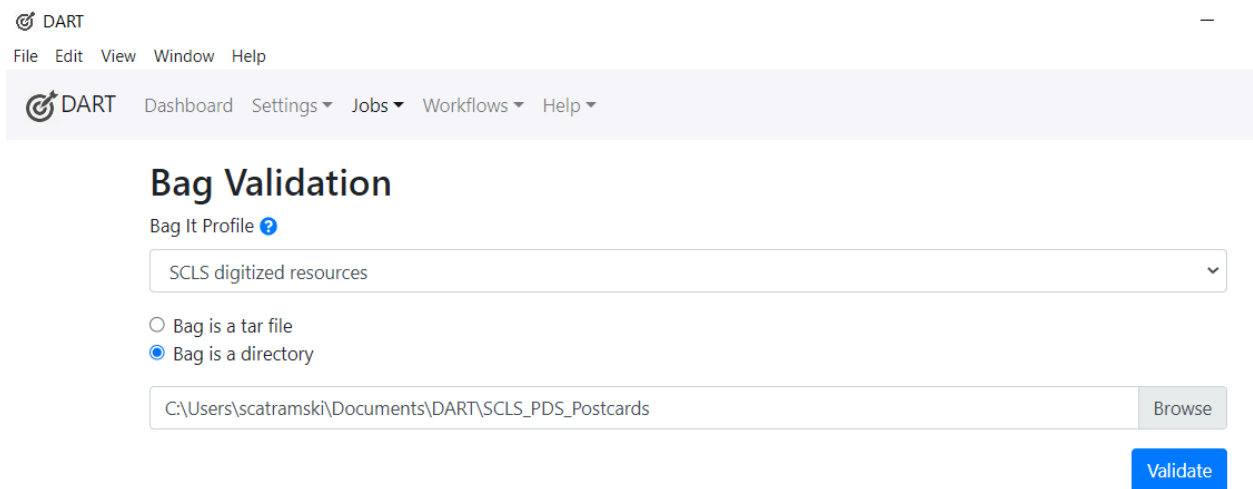


- The download can be monitored by selecting the Tasks button.
- When the download is complete, the Queued number will be 0.
- Check that Failed=0.
A possible reason for a failure is if there is not enough memory on your save to location.
- The download speed will vary depending on your Internet speed. Even small folder sizes can take an hour or more to download.
- To stop or cancel a download, the buttons are on the lower right-side of the screen.
If a download is stopped, it can be resumed by selecting Start All. It is recommended that a download not be stopped AND the computer shut down, as the download may result in failures after reconnecting to the server.



After downloading, the bag should be re-validated using DART.

- Open DART
- Select **Jobs >> Validate a Bag**.
- Bag It Profile: Select your profile
- Select: **Bag is a directory**
- Browse: Select the folder that was downloaded and **Upload**.
- Select: **Validate**



- DART should tell you that the bag is valid; the Outcome screen should say that the job was successfully completed.

Bag Validation

Bag It Profile [?](#)

SCLS digitized resources

- Bag is a tar file
 Bag is a directory

C:\Users\scatramski\Documents\DART\SCLS_PDS_Postcards

Browse

Job	Job of 2/23/23 8:31 AM
Process Id	16424
Started At	2023-02-23T14:34:42.803Z
Validation	Bag is valid
Outcome	Job completed successfully.

Validate

If the validation fails, first re-download the bag from the Dell ECS platform and attempt to re-validate.

After the bag has been validated, run a virus check on the bag using your tool of choice **before** delivering the requested files to the submitter.

Log the retrieval request and details about fulfilling the request in the digital archival storage log.

Removing Files

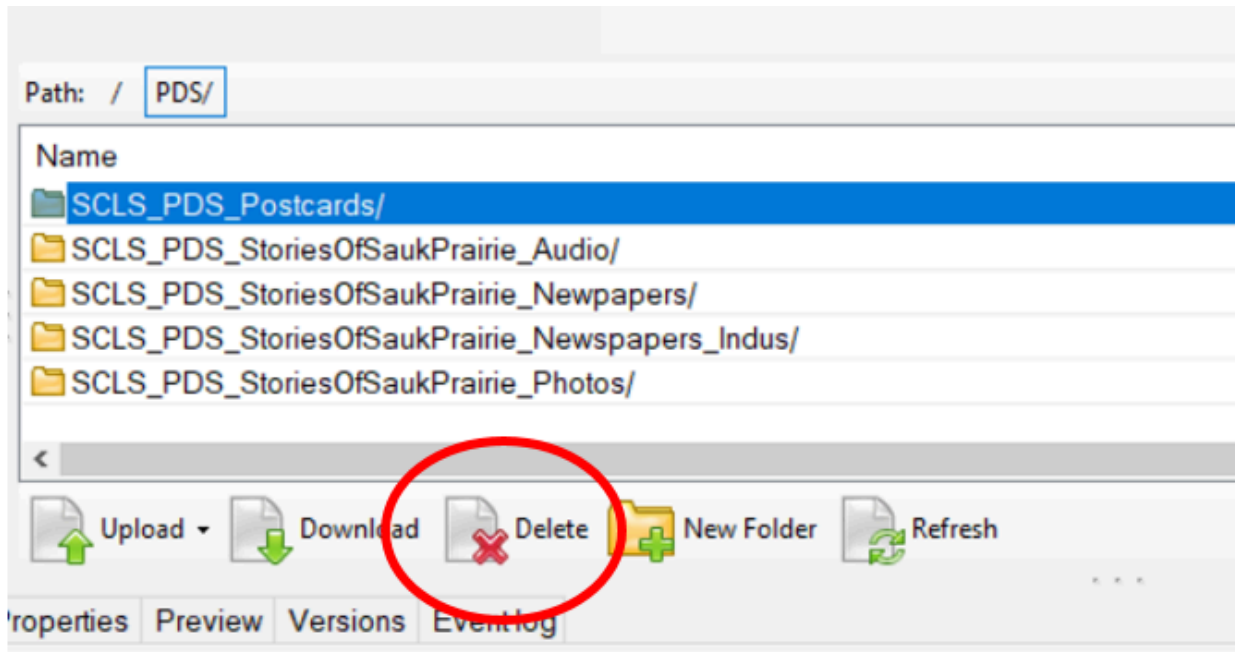
Removing individual files from a bag is not possible. All files that were bagged together have to be deleted together. If a single file needs to be deleted, the bagged folder must be deleted. Then the materials can be re-bagged and re-uploaded without the deleted file.

To delete a bag:

- Authenticate through the VPN, open the S3 browser, then open Dell ECS platform.
- Select the folder you want to delete.
- Select Delete.
- A window will open: Confirm File Delete
- Select Yes or No.

Note: The folder will be immediately deleted when Yes is selected. The file does not go to a recycle bin and there is not an option for undo or recovery after

selecting Yes. No other administrator confirms the deletion of these files, so be very sure you want to delete this bag before selecting Yes.



Log deleted bags and reasons for deleting on the digital archival storage log.

Updating Files

To add newly digitized files, updated files, updated metadata, or to include any other changes to stored files, you will:

- Create a new bag (using DART) with all of the previously uploaded files PLUS the newly created or updated files and/or metadata. ALL of the files from this collection will then be housed in this new bag.
- Move the new bag into the Dell ECS platform (see Section IV).
- Delete the original bag from the Dell ECS platform.

Appendix A: Configuring a BagIt Profile

You should only need to set up this profile one time at the beginning of your project. You'll download a profile template and customize it for your own organization.

Set save location for bags in this profile:

In DART, go to **Settings >> Application Settings**

Use the default location or create new a new folder, for example:

C:\Users\scatramski\Documents\DART

Import profile template

Go to **Settings >> BagIt Profile**.

- Import profile
- Import from JSON (Cut and Paste)
- Paste the JSON code
- Import

Customize profile

Select the newly imported profile and set up the following configuration:

- About
 - Name - for your system's profile, i.e. *SCLS Digitized Resources*
 - Description - same as the profile name, a description of what the profile will be used for
 - Save

BagIt Profile

About Info General Manifests Serialization Tag Files ▾

Name *

SCLS digitized resources

Description

Imported from <https://www.scls.info/documents/bagit-profile-scls-v0.1.json>

Delete Profile Export Profile Cancel Save

- Info (see screenshot below for SCLS example data)
 - Info Identifier - the profile identifier - that is, the permanent URI where the profile will live online.
 - Info Contact Email
 - Info Contact Name
 - Info External Description = Brief description of the function of the profile.
 - Info Source Organization = Full name of institution.
 - Info Version
 - To add more detail to your profile specifications, see [this guide](#).

BagIt Profile

[About](#) **Info** [General](#) [Manifests](#) [Serialization](#) [Tag](#)

Info Identifier

Info Contact Email

Info Contact Name

Info External Description

Info Source Organization

Info Version

- General
 - Accept Bag It Version = 1.0
 - Allow Fetch Txt = No

BagIt Profile

About Info **General** Manifests

Accept Bag It Version *

0.97
1.0

Allow Fetch Txt

No

Delete Profile Export Profile

- Manifests
 - Select md5 (and only md5) for all
 - Select Save

About Info General **Manifests** Serialization Tag Files

Manifests Allowed * ?

md5
sha1
sha224
sha256

Manifests Required

md5
sha1
sha224
sha256

Tag Manifests Allowed ?

md5
sha1
sha224
sha256

Tag Manifests Required

md5
sha1
sha224
sha256

Delete Profile Export Profile Cancel Save

- Serialization
 - Serialization = forbidden
 - Accept Serialization = select (only) application/zip and application/gzip
 - Tar Dir Must Match Name = No

BagIt Profile

[About](#) [Info](#) [General](#) [Manifests](#) [Serialization](#) [Tag Files ▾](#)

Serialization

forbidden

Accept Serialization

application/x-7z-compressed
application/tar
application/zip
application/gzip

Tar Dir Must Match Name

No

Delete Profile

Export Profile

- Tag Files
 - **Bag-info.txt** (Leave as default, or click on a Tag Name to edit it, or add New Tag) See below for SCLS settings
 - Add Library Name identifier tag for each separate library
 - Select bag-info.txt
 - Select New Tag
 - Tag Name = Library Name
 - Required = No
 - Values = Leave blank
 - Default Value = Leave blank
 - Save
 - The following tags can be deleted:
 - Internal-Sender-Description
 - Internal-Sender-Identifier
 - Bag-Count

■ Bag-Group-Identifier

BagIt Profile

[About](#) [Info](#) [General](#) [Manifests](#) [Serialization](#) [Tag Files ▾](#)

New Tag

Tag Name	Default Value	
Bag-Size	[Set automatically by bagger]	✗
Bagging-Date	[Set automatically by bagger]	✗
Bagging-Software	[Set automatically by bagger]	✗
Contact-Email	tramski@scls.info	✗
Contact-Name	Tamara Ramski	✗
Contact-Phone	+1 (608) 242-4866	✗
External-Description	Bagit Profile for SCLS member library digitized resources	✗
External-Identifier		✗
Library Name		✗
Organization-Address	4610 South Biltmore Lane, Madison, WI 53718	✗
Payload-Oxum	[Set automatically by bagger]	✗
Source-Organization	South Central Library System	✗

- **Bagit.txt** (Leave as default or click on a Tag Name to edit it)
Example: See below for SCLS settings

BagIt Profile

[About](#) [Info](#) [General](#) [Manifests](#) [Serialization](#) [Tag Files ▾](#)

New Tag

Tag Name	Default Value	
🔗 BagIt-Version	1.0	✗
🔗 Tag-File-Character-Encoding	UTF-8	✗

Delete Profile

Export Profile

Cancel

Save

Other DART settings in the **Settings** menu: Leave unchanged.

Once your profile is updated to your specifications, export the JSON code and publish it online in a permanent location.

- DART >> Settings >> BagIt Profiles
- About tab
- Click “Export Profile”
- Copy the resulting code.
- Publish publicly online to a webserver with a URL that will be unchanged so that bag validators can retrieve it and validate the profile.