# Setting up DART tool to deposit bags to APTrust demo and repo using Curation Scripts

[Setting up DART tool to deposit bags to APTrust demo and repo using Curation Scripts 1](#_heading=h.baofawomskhr)

[Downloading DART: 2](#_heading=h.9gka94zgk9i6)

[Section 1 2](#_heading=)

[Setting up Storage services on DART: 2](#_heading=h.o5uelcuhf1ut)

[Section 2 4](#_heading=)

[Setting up Bagit profile on DART: 4](#_heading=h.6odi5sittxcg)

[Setting up Workflows on DART: 4](#_heading=h.j0yngrmos8c8)

[Section 3 6](#_heading=)

[Running a workflow on DART app manually: 6](#_heading=h.txbzpk22ipup)

[Section 4 7](#_heading=)

[Running a workflow on DART app through VTDR workflow scripts: 7](#_heading=h.30j0zll)

[Section 5 8](#_heading=)

[Depositing bags in tar format manually to APTrust storage system and accessing deposited bags in tar format from APTrust: 8](#_heading=h.n6534boakx5s)

[Section 7 9](#_heading=h.lqp6g1eh4ary)

[VT libraries S3 setup 9](#_heading=h.vjofheqm403e)

# 

# Downloading DART:

-Download DART app from [Download - DART Documentation (aptrust.github.io)](https://aptrust.github.io/dart-docs/download/) and do the DART setup

# 

## Setting up Storage services on DART:

**The following description shows how to create storage settings. We will be creating settings for 3 storage services:**

1. Go to “Settings”, pick “Storage Service”, fill in following tags:

Name: APTrust Demo Receiving Bucket

Description: Receiving bucket for ingest to the APTrust Demo repository

Protocol: s3

Host: s3.amazonaws.com

Port: 0

Bucket: aptrust.receiving.test.vt.edu

Allows Upload: Yes

Allows Download: Yes

Login: (enter credentials)

Password: (enter credentials)

Click “Save” and save this storage setting.

1. Create another storage service, click “New” under “Storage Service” and fill in the following tags:

Name: APTrust Production Receiving Bucket

Description: Receiving bucket for ingest to the APTrust Production repository

Protocol: s3

Host: s3.amazonaws.com

Port: 0

Bucket: aptrust.receiving.vt.edu

Allows Upload: Yes

Allows Download: Yes

Login: (enter credentials)

Password: (enter credentials)

Click “Save” and save this storage setting.

1. click “New” under “Storage Service” and fill in the following tags:

Name: Virginia Tech Library S3 bucket

Description: Virginia Tech Library S3 bucket ESRI storage

Protocol: s3

Host: s3.amazonaws.com

Port: 0

Bucket: gisfiles.vtlibcloud.com

Allows Upload: Yes

Allows Download: Yes

Login: (enter credentials)

Password: (enter credentials)

Click “Save” and save this profile.

# 

## Setting up Bagit profile on DART:

**The following description shows how to create a “Bagit” profile. We will be creating settings for 1 “Bagit” profile:**

-Open the DART app, go to

1. Settings->Bagit Profiles
2. Click “New”
3. Select “Base Profile”: APTrust 2.2 default Bagit profile
4. Name\*: APTrust with default values for Virginia Tech
5. Description: Customized version of APTrust for Virginia Tech
6. Click “Save”
7. Go to the profile created under ‘BagIt Profiles’
8. Go to ‘Serialization’, under ‘Serialization’, make sure the option selected is ‘optional’
9. Go to ‘Tag Files’, click ‘Title’, under ‘Required’ make sure ‘Yes’ is selected, if this field has ‘No’ then change to ‘Yes’
10. Go to ‘bag-info.txt’ , click ‘Bag-Group-Identifier’, under ‘Required’ make sure ‘Yes’ is selected, if this field has ‘No’ then change to ‘Yes’
11. Go to ‘bag-info.txt’ , click ‘Source-Organization’, under ‘Required’ make sure ‘Yes’ is selected, if this field has ‘No’ then change to ‘Yes’
12. Go to ‘aptrust-info.txt’ , click ‘Access’, under ‘Required’ make sure ‘Yes’ is selected, if this field has ‘No’ then change to ‘Yes’
13. Go to ‘aptrust-info.txt’ , click ‘Storage-Option’, under ‘Required’ make sure ‘Yes’ is selected, if this field has ‘No’ then change to ‘Yes’.

## Setting up Workflows on DART:

**The following description shows how to create a workflow for depositing bags to APTrust. We will be creating settings for 3 workflows corresponding to the 3 storage settings:**

1. Go to the “Workflows” tab, click “New”. Fill in the following tags:

Name\*: Workflow for depositing bag to APTrust-Demo

Description: Upload bag to the APTrust-Demo system

Package Format: Bagit

Upload To(Select this option): APTrust Demo Receiving Bucket

Click “save”.

Now open “Workflow for depositing bag to APTrust-Demo” (opening this after saving enables us to pick a Bagit profile), under “Bag It Profile Id” pick “APTrust with default values for Virginia Tech”. Click Save.

1. Go to the “Workflows” tab, click “New”. Fill in the following tags:

Name\*: Workflow for depositing bag to APTrust-Repo and VT library S3 bucket

Description: Upload bag to the APTrust production system (Repo) and VT library S3 bucket

Package Format: Bagit

Upload To(Select this option): APTrust Production Receiving Bucket

Upload To(Also select this option): Virginia Tech Library S3 bucket

Click “save”.

Now open “Workflow for depositing bag to APTrust-Repo and VT library S3 bucket” (opening this after saving enables us to pick a Bagit profile), under “Bag It Profile Id” pick “APTrust with default values for Virginia Tech”

1. Go to the “Workflows” tab, click “New”. Fill in the following tags:

Name\*: Workflow for depositing bag to VT library S3 bucket

Description: Upload bag to the Virginia Tech library S3 bucket

Package Format: Bagit

Upload To(Select this option): Virginia Tech Library S3 bucket

Click “save”.

Now open “Workflow for depositing bag to VT library S3 bucket” (opening this after saving enables us to pick a Bagit profile), under “Bag It Profile Id” pick “APTrust with default values for Virginia Tech”

1. Go to the “Workflows” tab, click “New”. Fill in the following tags:

Name\*: Workflow for depositing bag to APTrust-Repo

Description: Workflow for depositing bag to APTrust-Repo

Package Format: Bagit

Upload To(Select this option): APTrust Production Receiving Bucket

Click “save”.

Now open “Workflow for depositing bag to APTrust-Repo” (opening this after saving enables us to pick a Bagit profile), under “Bag It Profile Id” pick “APTrust with default values for Virginia Tech”

# 

## Running a workflow on DART app manually:

**The following section shows how to run a DART workflow to create a bag (in tar format) manually using default values for Virginia Tech**.

Note: In most cases, ingest and publication bags are created by calling DART app through [VTDR workflow scripts](https://github.com/PadmaCarstens/VTechDataRepo), but in some cases when the dataset is large (in the order of 100’s of GB), it is usually broken up and bagged up into different parts. DART is run manually in such a case to create a bag for each of these parts.

1. Go to “Jobs”, click “New”, drag the files that need to be bagged.

*For example*: If XYZ submits a 100GB dataset, this dataset would be downloaded in parts following the naming convention: VTDR\_I00NNN\_XYZ\_XYZ\_v01\_xof8\_date, where NNN is the ingest number. Go inside this folder and drag the contents into the “Files” box (Note: The files inside this folder are to be dragged and dropped and **not** the folder by itself)

1. Click “Next”, fill in the following tags:

**Package Format:** Bagit

**Bagit Profile:** APTrust

**Serialization:** .tar

**Package Name:** VTDR\_I00NNN\_XYZ\_XYZ\_v01\_xof8\_date (or) VTDR\_P00Num\_I00NNN\_DOI\_ArticleID\_CorresAuthor\_v0X\_date

(Leave Output Path to default, note that the default path C:\Users\username\.dart can be changed under Application Settings)

**Output Path:** C:\Users\username\.dart\bags\VTDR\_TestBag**.tar**

***Note:*** Make sure that the extension .tar appears at the end of the output path, if it does not appear at the end of the bag name then add it before clicking “Next”

**Title:** VTDR\_I00NNN\_XYZ\_XYZ\_v01\_xof8\_date (or) VTDR\_P00Num\_I00NNN\_DOI\_ArticleID\_CorresAuthor\_v0X\_date

**Bag-Group-Identifier:** VTDR\_I00NNN (or) VTDR\_P00Num

**Source-Organization\*:** Virginia Tech

1. Click “Next”, select the storage system where you would like the bag to be deposited.
2. Click “Next”. Click “Run Job”. Note that a copy of the bag that will be uploaded to the selected storage system can also be found on your local computer at the “Output Path” displayed on this DART page.

To check the contents of the bag in tar format created at the output path, go to the Start->Command Prompt and type:

>cd <Output Path> (*For Example*: if bag is created at an output path C:\Users\username\Documents\DART then change directory to

cd C:\Users\username\Documents\DART)

>tar -xvf VTDR\_I00NNN\_XYZ\_XYZ\_v01\_xof8\_date.tar

This will extract the contents of the bag above in the same directory as the bag in tar format

# 

## Running a workflow on DART app through [VTDR workflow scripts](https://github.com/PadmaCarstens/VTechDataRepo):

**Following section describes how to select the workflows created above through the** [**VTDR workflow scripts**](https://github.com/PadmaCarstens/VTechDataRepo)**.**

The 3 workflows created in section 3 can be enabled/disabled through the ingest dataset download and deposit script or publication dataset download and deposit script by commenting/uncommenting certain lines of codes in the script as described below:

1. For Ingest dataset download and deposit workflow:

Open IngFolder\_Download\_TransferBagAPTrust.py from the cloned [**VTDR workflow scripts**](https://github.com/PadmaCarstens/VTechDataRepo) in the curation folder (See ScriptsSetupAndExecution\_CurationWorkflow\_Windows/ScriptsSetupAndExecution\_CurationWorkflow\_Mac document for instructions to set up the environment and run the [**VTDR workflow scripts**](https://github.com/PadmaCarstens/VTechDataRepo)) and enable/disable different workflows as described in step 3 below.

1. For Publication dataset download and bag deposit:

Open PubBagDART\_TransferAPTrust.py from the cloned [**VTDR workflow scripts**](https://github.com/PadmaCarstens/VTechDataRepo) in the curation folder (See ScriptsSetupAndExecution\_CurationWorkflow\_Windows/ScriptsSetupAndExecution\_CurationWorkflow\_Mac document for instructions to set up the environment and run the [**VTDR workflow scripts**](https://github.com/PadmaCarstens/VTechDataRepo)) and enable/disable different workflows as described in step 3 below.

1. Pick a workflow for depositing bag to APTrust demo/repo/VT S3 (type 1, 2 or 3 from the options below):

Option 1: type “1” (without the quotations) to deposit the ingest bag to APTrust-Demo storage, this option is for depositing test bags to APTrust-Demo storage system. Bags deposited here will be deleted after a certain amount of time

Option 2: type “2” to deposit the ingest bag to APTrust-Repo and VT library S3 storage systems (this is the standard option). APTrust-Repository does periodic integrity checks on the bags deposited. Depositing bags to the VT library S3 storage system will allow us to access the bags if needed. Bags deposited to VT library S3 storage will be deleted every few months.

Option 3: type “3” to deposit the ingest bag to VT library S3 storage system only

Option 4: enter “4” to deposit the ingest bag to APTrust-Repo

## Depositing bags in tar format manually to APTrust storage system and accessing deposited bags in tar format from APTrust:

[APTrust Partner Tools](https://aptrust.github.io/userguide/partner_tools/) tools enable us to upload the bags manually to APTrust without using DART app. In order to do this, download the [APTrust Partner Tools](https://aptrust.github.io/userguide/partner_tools/) found at this page, set up a configuration file as described in [APTrust Partner Tools](https://aptrust.github.io/userguide/partner_tools/) and save it as .aptrust\_partner.conf in your homepath (e.g. C:\Users\username), the downloaded tools (e.g. apt\_upload.exe, apt\_list.exe etc.) should be in the same folder as the configuration file created so that these can interact with each other

To upload/ingest content to APTrust-Demo or APTrust-Repo directly:

Open the command prompt, and change the directory to the APTrust partner tools directory and execute apt\_upload. In the example below, VTDR\_P00179\_I00209\_DOI\_19709143\_EmoriS\_v01\_20220509.tar is the publication bag, bagged in accordance with APTrust bagging conventions, “C:/Users/username/.dart/bags/” is the path where the bag is stored:

>cd C:\Users\username

>apt\_upload --key="VTDR\_P00179\_I00209\_DOI\_19709143\_EmoriS\_v01\_20220509.tar" C:/Users/padma/.dart/bags/VTDR\_P00179\_I00209\_DOI\_19709143\_EmoriS\_v01\_20220509.tar

Other tools for displaying the list of recent bags uploaded, checking ingest status etc. can also be implemented following the instructions found at [APTrust Partner Tools](https://aptrust.github.io/userguide/partner_tools/)

The download/delete tools are not available for users listed under institutional user roles. These capabilities are only accessible for users under institutional admin roles.

**Other tips**: If you would like to store the bags that DART creates in a different directory (default is C:\Users\Username\.dart\bags), go to Settings->Application Settings and change the “Value” to the path where you want to store the bags.

# Section 7

## VT libraries S3 setup

Download WebDrive [Download WebDrive WebDAV Client, Cloud Client, FTP Client](https://webdrive.com/download/) for windows/Mac system. Setup WebDrive to access VT libraries S3 bucket following instructions in Accessing\_Bags\_S3 document. VT libraries S3 bucket can also be set up for access from cyberduck, download cyberduck for windows/Mac system: [Download (cyberduck.io)](https://cyberduck.io/download/) and follow the instructions in the Accessing\_Bags\_S3 document.

*Note:* Manual deletion of bags periodically can be done using cyberduck software, deletion cannot be done if using WebDrive