5.25” FLOPPY DISK IMAGING MANUAL

Digital Preservation Unit

Department of Preservation and Conservation | University of Michigan Library

Updated by: Andrea Kang · Christina Min

Authors: Kayla Carucci · Leigh Gialanella · Alexa Hagen · Noa Kasman
Andrea Kang · Christina Min

Public Version 2.0

Last Revised | 23 August 2018

THE WORKFLOW

Part I: Preparation

1. Create Tracking Database
2. Prepare Main Directories
3. Prepare Media Directory
4. Photograph Media

Part II: Disk Imaging

5. Configure Hardware for Disk Imaging
6. Setup FC5025 (aka R2D2)
7. Create Disk Image and Generate Checksums

Part III: Brunnhilde

8. Navigating to the BitCurator Virtual Machine
9. Run Brunnhilde
10. Reports

Part IV: Bagging and Transfer to Storage

11. Package SIP using BagIt
12. Validate bag
13. Transfer to the Network Attached Storage (NAS)

PART I: PREPARATION

[STEP 1] Create Tracking Database

Create a working database on Google Sheets. This sheet will help keep track of the floppy disks that have been imaged and their results.

If the content is coming from Special Collections Research Center (SCRC), add another sheet to the Preservation Imaging and Logical Transfer Inventory found here. Content from SCRC should come with a finding aid. Incorporate any information you find necessary from the finding aid into the working database.

Input the following information into the working database:

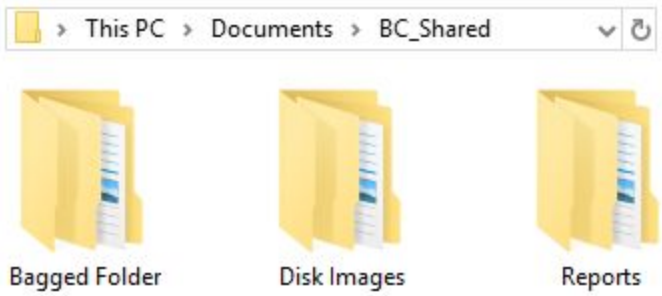
- Last handled: date the media/media image was last handled
- Collection: John Sayles
- Unique ID: copy from the inventory spreadsheet
- Barcode: barcode of media
- Label: label on the media
- Type: type of media
- Imaging Results: drop down
- Bad Sectors: note any bad sectors from the imaging results
- Mount Results: drop down (based on Brunnhilde results)
- Virus Check Results: drop down (based on Brunnhilde results)
- Bag Check: write "Y" or "N" for whether package was bagged or not
- Reports Created: "Y" or "N" for whether reports directory for SCRC and the Digital Preservation Librarian have been copied from the *metadata* folder
- Transfer Check: write "Y" or "N" for whether bag was transferred or not to DeepBlue
- Notes: any information regarding the disk (e.g. bent slider) or tools (e.g. Brunnhilde froze)

Reconcile the finding aid with items sent to Buhr and notify the Digital Preservation Librarian of any discrepancies. Barcode and physically inspect all media, examine disks for mold and/or physical damage.

- Mold on disks looks like spilled milk and smells of mildew. If mold is present, notify the Digital Preservation Librarian. Refer to this document for more information on handling mold.
- Record any damages (e.g. bent plastic jacket) of the media on the working database
- If the media has a storage case or container, the media item and storage case or container should be given the same barcode and added to the working database.

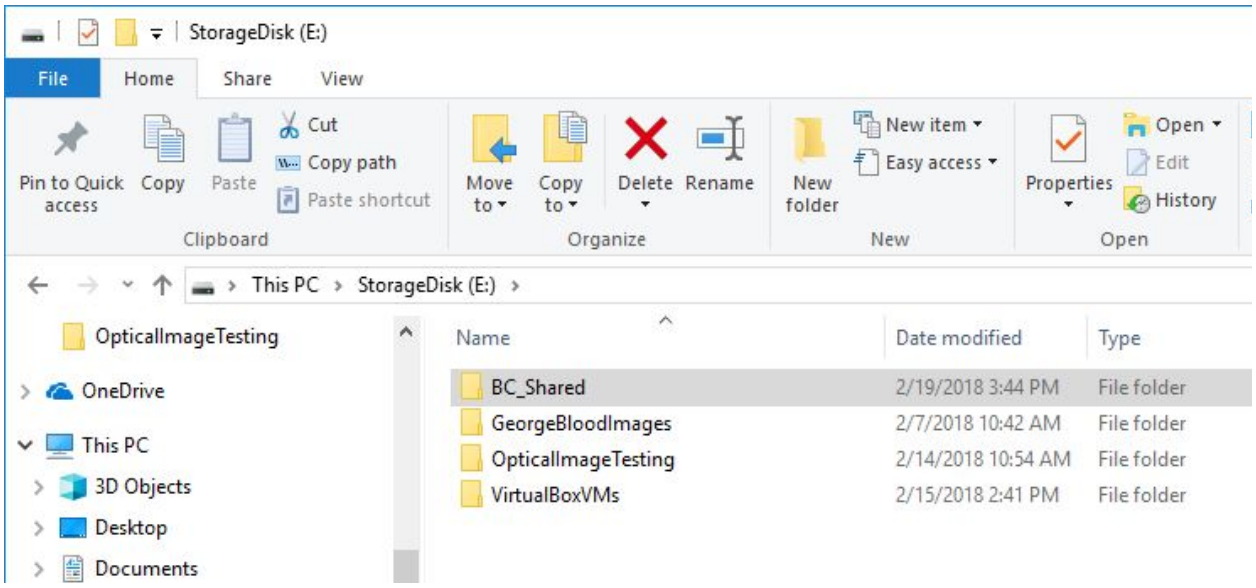
[STEP 2] Prepare Main Directories

- Prepare three directories named *Disk Images*, *Reports*, and *Bagged* in the *BC_Shared* folder, which is the folder shared by the BitCurator Virtual Machine (BC VM) and the host computer. It is *BC_Shared* on the host computer and *sf_BC_Shared* on the BC VM.



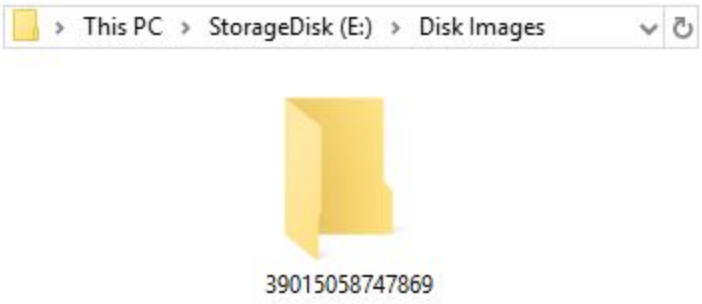
- These directories distinguish the folder you generate when you create a disk image, the final bagged folder created through Bagger, and the reports that will be sent off to Special Collections and the Digital Preservation Librarian.

- The *sf_BC_Shared* folder can be found on the StorageDisk (E:) hard disk on the host computer (below) or the BC desktop (above).

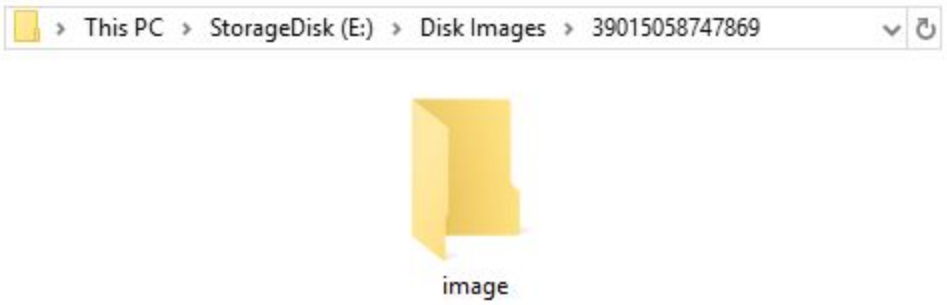


[STEP 3] Prepare Media Directory

- Create a new folder inside the *Disk Images* folder. Use the barcode scanner to scan the barcode number on the media as the file name.

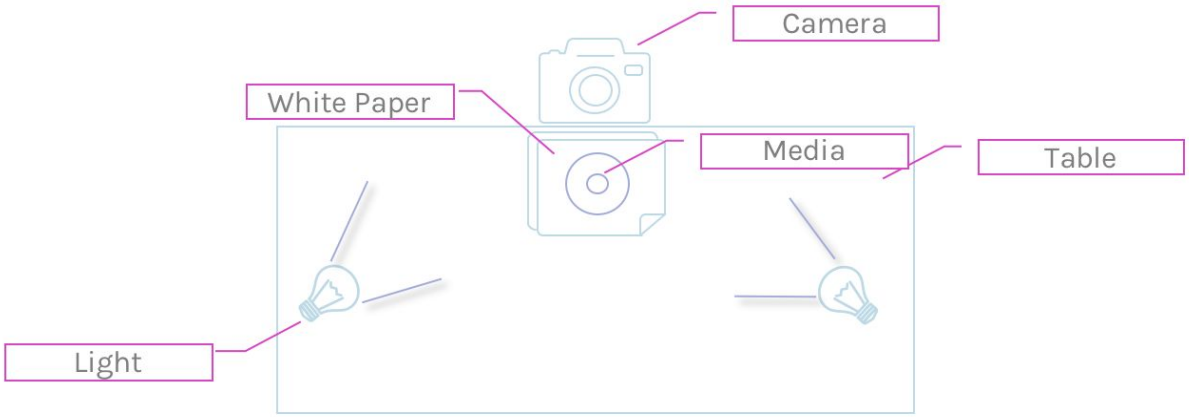


- Within that new folder, create a new folder called *image*. Other folders will be generated using tools on the BC VM.



[STEP 4] Photograph Media

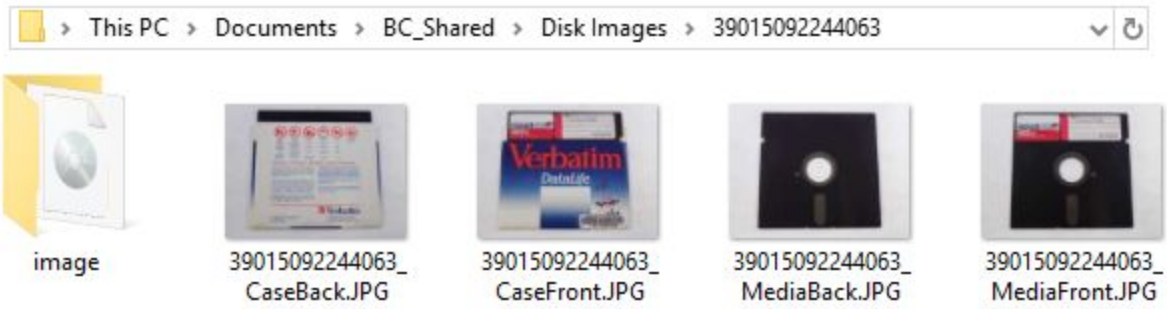
- Next, set up your photographing station by placing a white piece of paper under the cloth to make the backdrop as white as possible.
    - Make sure to put the paper as close to edge of the table as possible.
    - If there is lint or other debris on the white backdrop, use tape to make it as clean as possible.
    - An iron is available for removing wrinkles. Keep it on the synthetic setting and use towels to protect the material from full heat.
- Place the media you are photographing about an inch or two away from the edge of the table.
- Point the lights at a 45 degree angle towards the media, but not so directly that the reflection of the light is too strong.
    - You can also adjust the strength of the lights by pulling them closer or further away from the source.
    - The dimmer switch in the room can also be used for lighting purposes.



- Attach the camera to the tripod and angle the camera lens so that it can capture a flat image of the media.
    - You may not be able to get it completely flat without photographing outside of the white backdrop. Get it as flat as possible.

- Zoom into the media to ensure that there is only the white backdrop in the background.
- Photograph the front and back sides of the media.

- Photograph the front and back sides of the storage case for the media. This is especially important if the case contains relevant writing or a barcode. You do not need to take a photo of the case if you gave it a new storage case during inventory.
    - Take 3 sets of photos (resulting in a total of 6 photos per media, not including the media case) that are set to different exposures.
        - Turn the camera on.
        - Click the button with the +/- signs that is situated on the right hand side of the camera.
        - Take photos of the front and back of the media at exposures 0.0, +0.3, and +0.7.
- When finished, connect the camera to the computer using the USB cable.
- Select the photos with the most similar lighting of the highest quality and move them to the appropriate barcode folder. Once the photos have been moved, delete all the photos on the camera and unmount.
- Rename the photos with the barcode, underscore, and description of the image.
    - Some descriptions you can use are:
        - MediaFront
        - MediaBack
        - CaseFront
        - CaseBack
        - Notes01
    - An example for a floppy disk would 1234567890_MediaFront



- Save the photographs you have taken of the media in the barcoded folder for the disk image. These will later be dragged into the *metadata* folder that Brunnhilde generates.

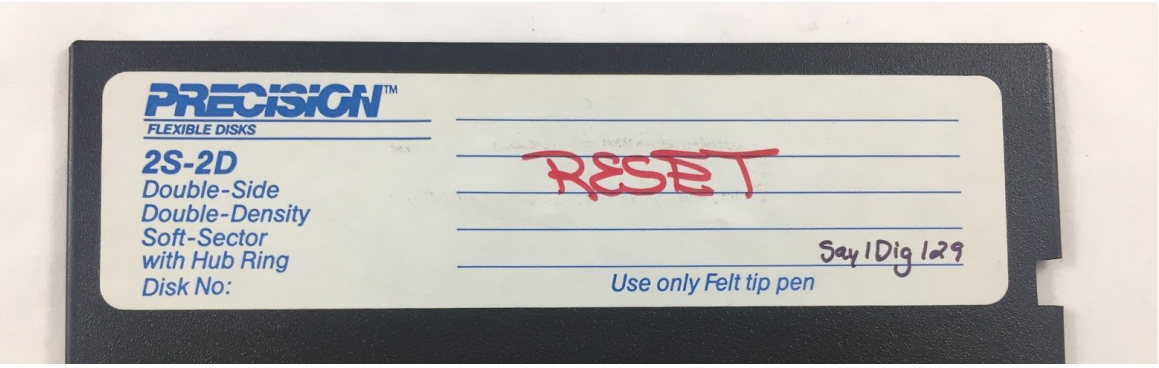
PART II: DISK IMAGING

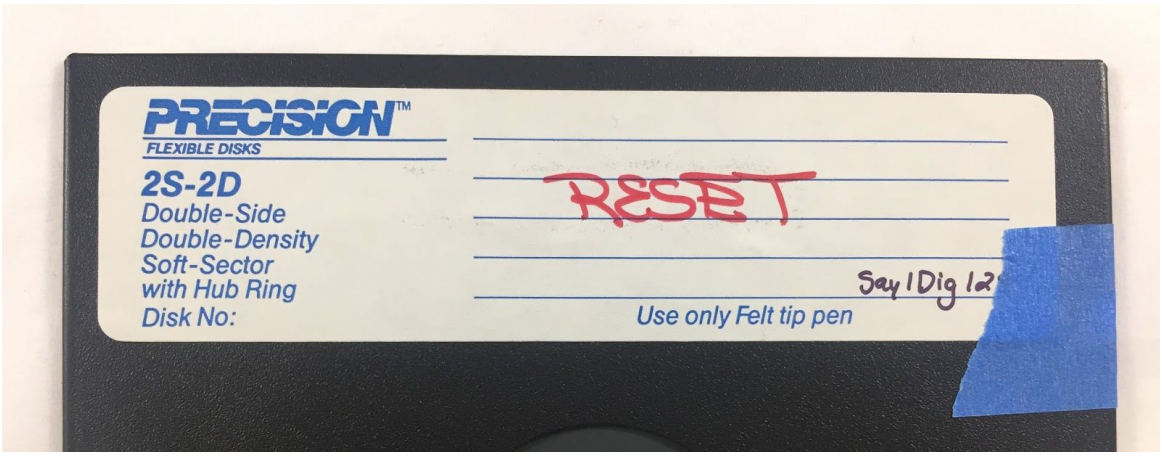
[STEP 5] Configure Hardware for Disk Imaging

5.25” FLOPPY DISK

Select the appropriate 5.25in floppy disk drive, locate the FC5025 and the appropriate cables.

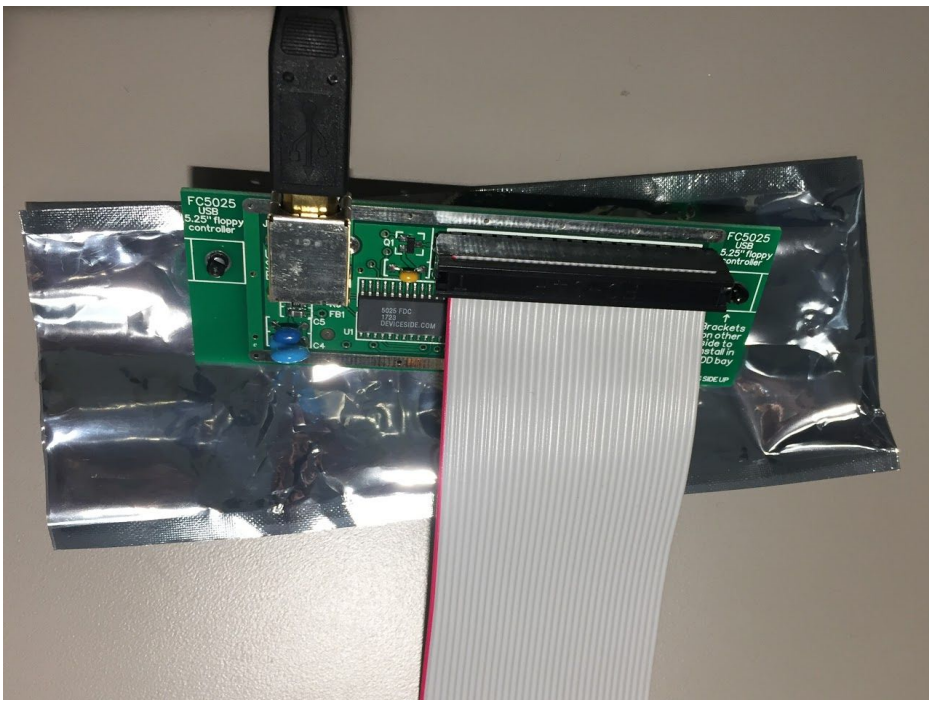
Cover the write protect notch on the right side of the disk by covering it with a piece of painter’s tape or sticky note.



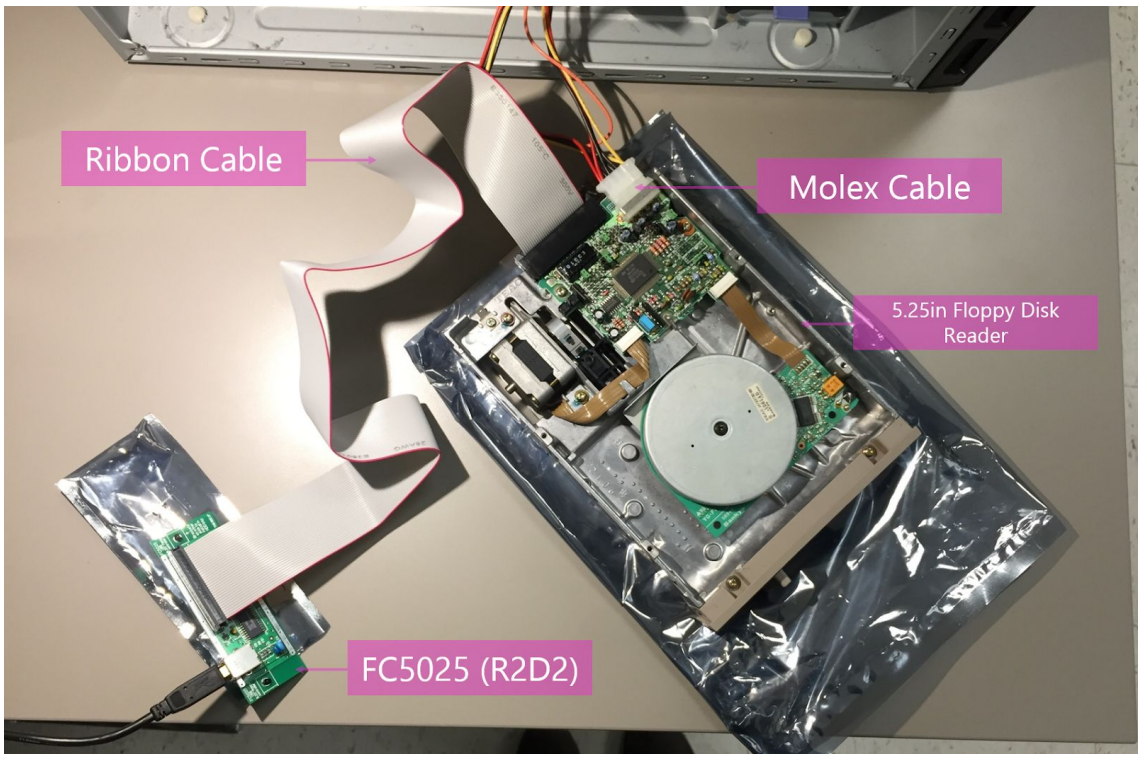


[STEP 6] Setup FC5025 (aka R2D2)

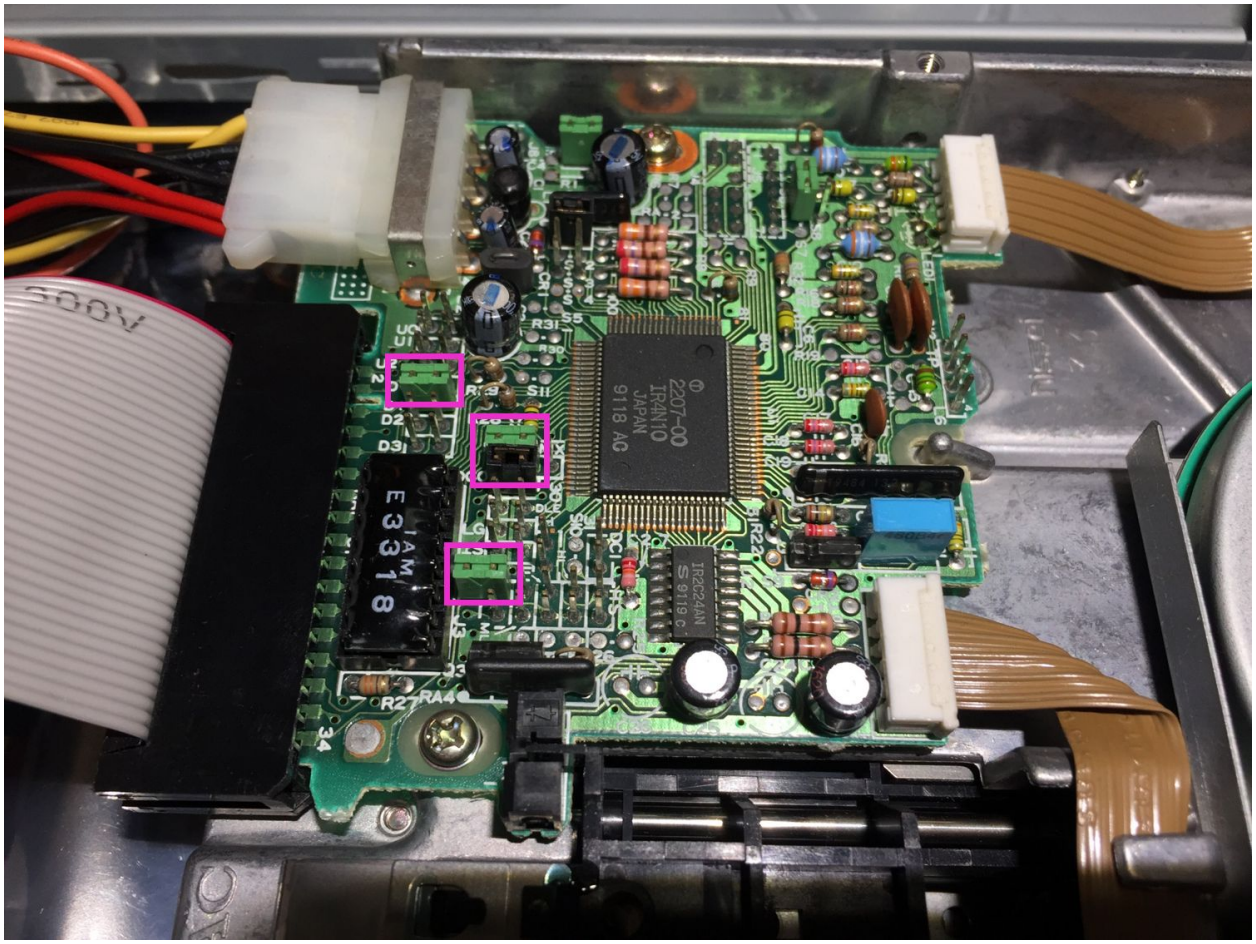
- R2D2 is a floppy disk controller that you will need to connect the 5.25in floppy disk reader to a modern computer.



- Connect R2D2 to the 5.25in floppy disk reader with the grey ribbon cable like the picture below. Make sure it is connected in this way.
    - Ensure that the red stripe goes on the same side that says “Pin 1 (Red Stripe”).
    - Note that the molex cable is used to connect to a power source.



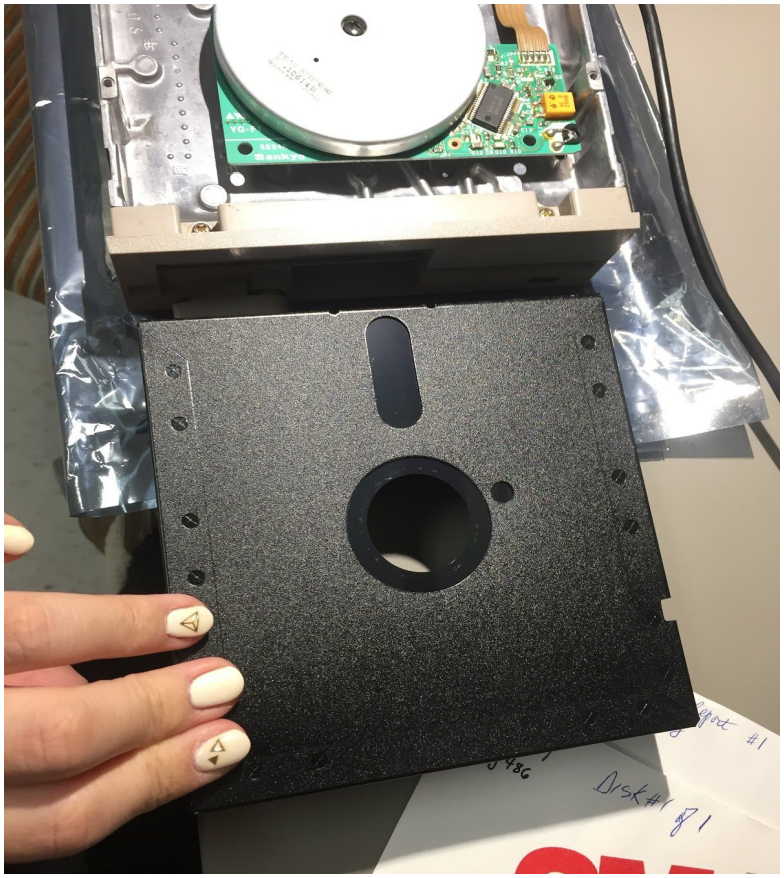
- Make sure that the jumpers are in the correct place as pictured below.
  - Visit the Maryland Institute for Technology in the Humanities page for more information about jumpers and R2D2 in general.

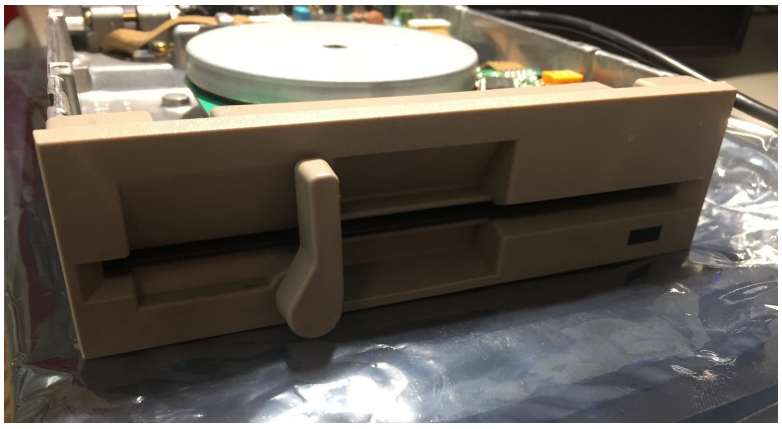


[STEP 7] Create Disk Image and Generate Checksums

FC5025

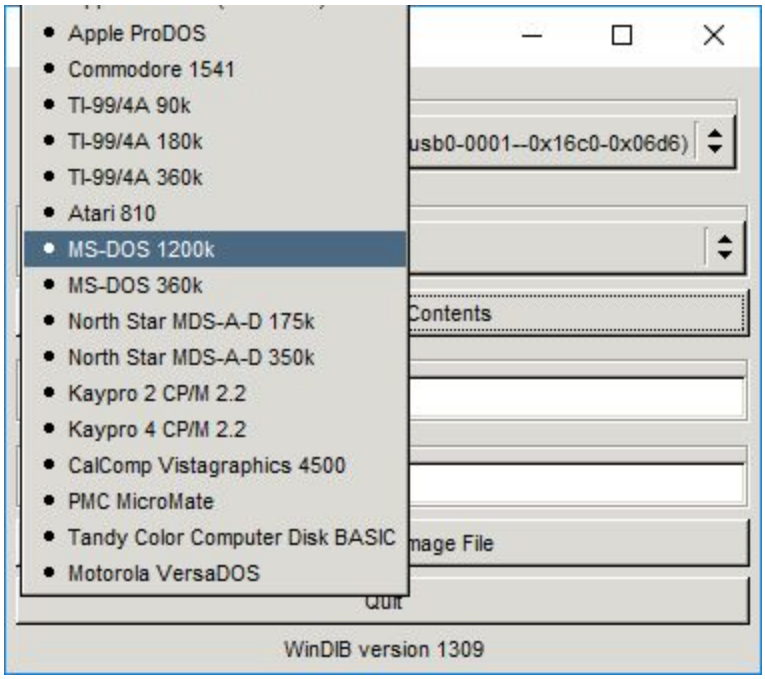
- The floppy disk driver is used upside down and the floppy disks will need to be inserted into the reader upside down. Once it is in the reader, flip the switch in an upright position to secure it in place.


Disk #1

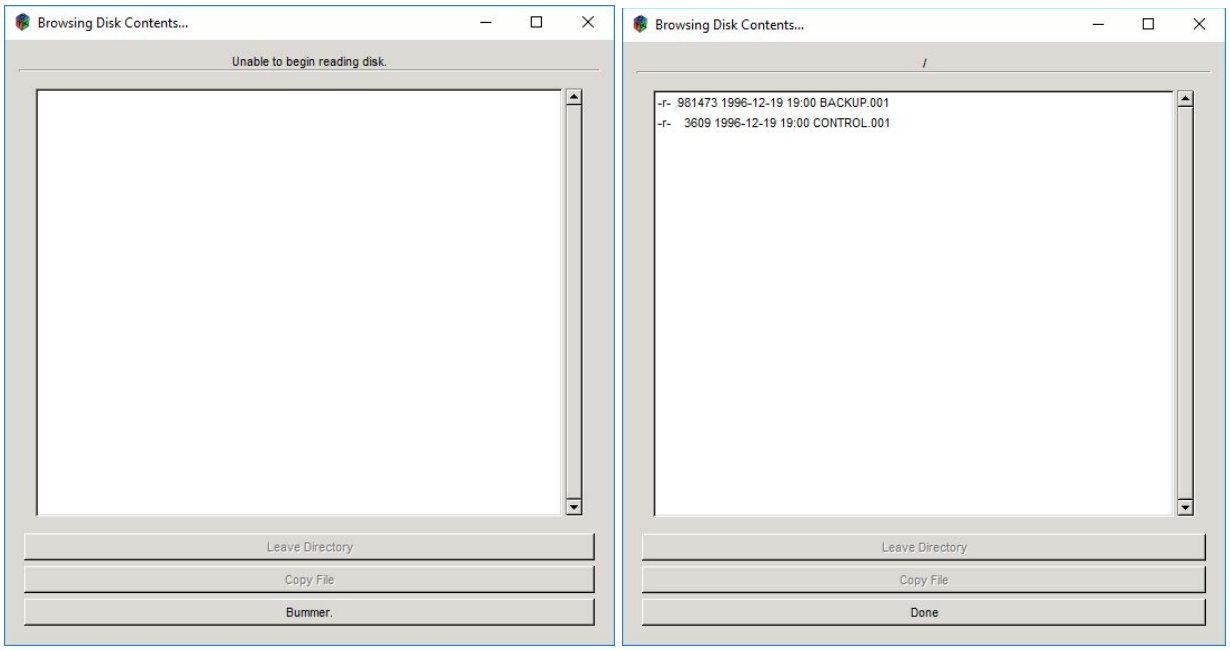


- Open R2D2 on the host computer by clicking the “Disk Image and Browse” icon.

- A “Disk Image and Browse” window will pop up.
- Under “Disk Type”, you must select the type of disk that you have inserted.
    - You must do so by clicking through each option to see whether or not you can browse the disk contents.
    - When you select a potentially correct disk type, the “Browse Disk Contents” button will change to a selectable button. Click this button to see if you can see a list of files on the disk.

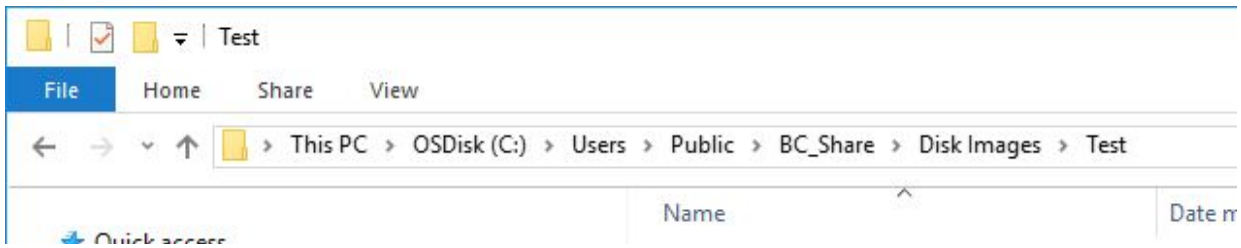


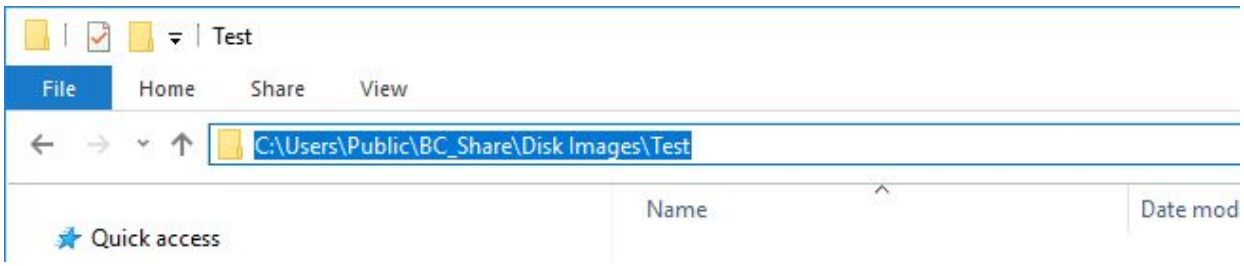
    - If it is not the correct disk type, nothing will appear except “Unable to begin reading disk” at the top. If this happens, click “Bummer.” and select another disk type.



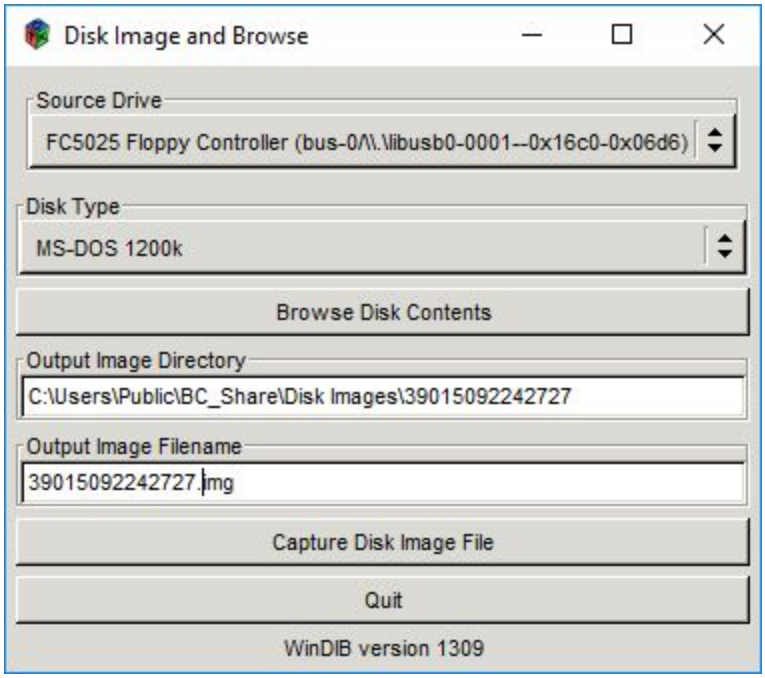
- For the most part, the 5.25 in floppy disks in our collection have been MS-DOS formatted.

- Under the “Output Image Directory”, input the file path of the location you want to save the image.
  - You can either type this out or you can go to the actual location, click on the drop down bar at the top, and copy and paste it into the “Output Image Directory”.

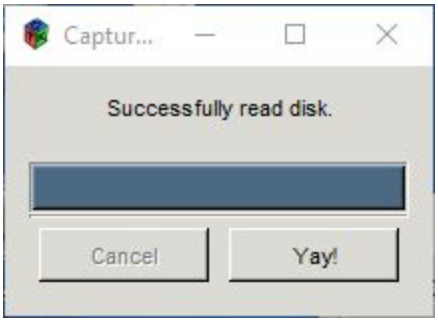




- R2D2 will create a .img file. Insert the barcode of the media under “Output Image Filename”. It should look something like the following:
  - 39015092242727.img



- Click “Capture Disk Image File”.
- Once the imaging is finished, it will either say that it was successful or that it failed in capturing the image. If it failed, click “Bummer.” and try again. If not, click “Yay!” and move onto the next step.

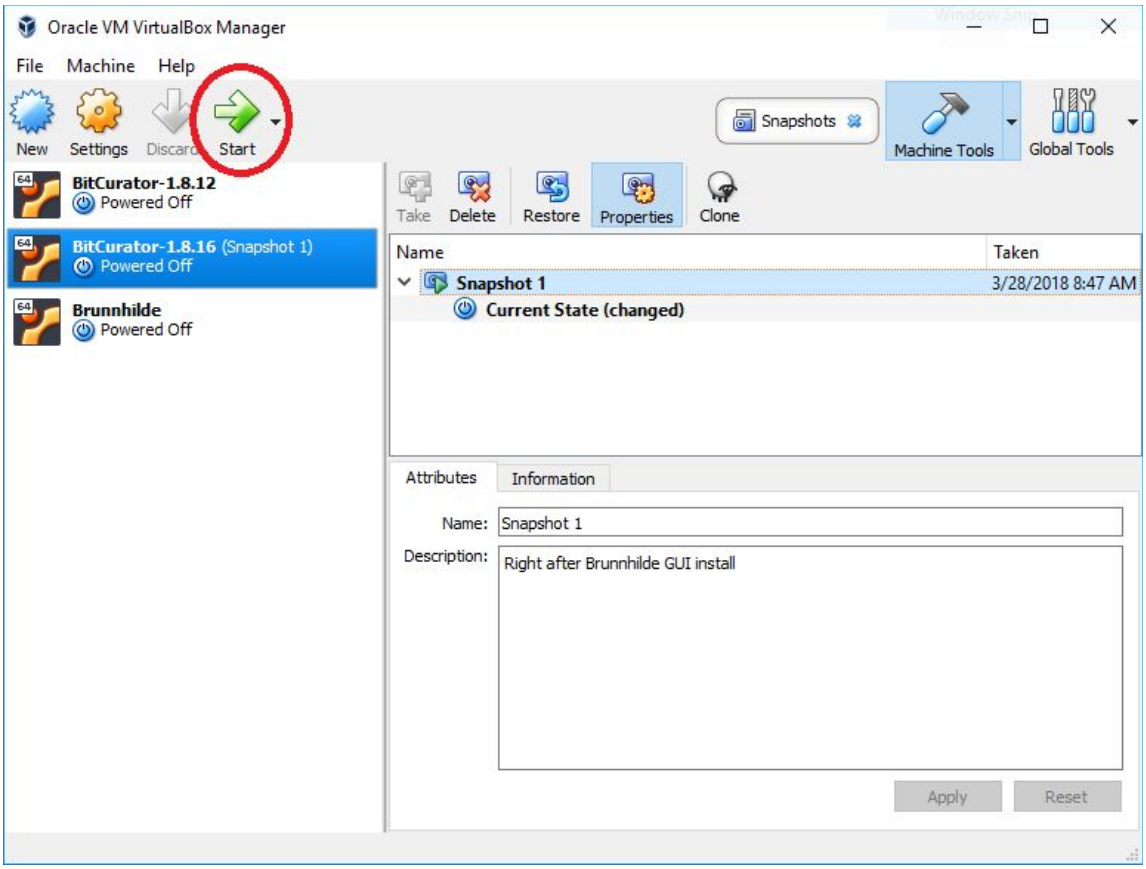


PART III: BRUNNHILDE

[STEP 8] Navigating to the BitCurator Virtual Machine

- In order to run Brunnhilde, you must access the BitCurator (BC) environment using a virtual machine (VM). The BC VM has already been set up by Digital Preservation Unit staff.
- Double click the “Oracle VM VirtualBox” on the host computer’s desktop.

- Then, select the appropriate VM and click “Start”.
    - If you have already used this, it will automatically be highlighted when you launch the VirtualBox and all you will have to do is click “Start”.



- Once the BC VM loads, you can use all the tools that come in the BC package and access the *BC_Share* folder, which is located on the BC VM desktop.
- When you are ready to shut everything down, just click on the power button at the top right hand corner of the BC VM screen, scroll all the way down on the dropdown menu, and select “Shut Down…”.

BEFORE YOU START

- Confirm that BitCurator is in WRITABLE mode.
- The drive icon at the top right hand corner of the screen will be red if writable mode is activated.
- We are setting BitCurator to writable due to a mounting related quirk that the read-only mode creates.

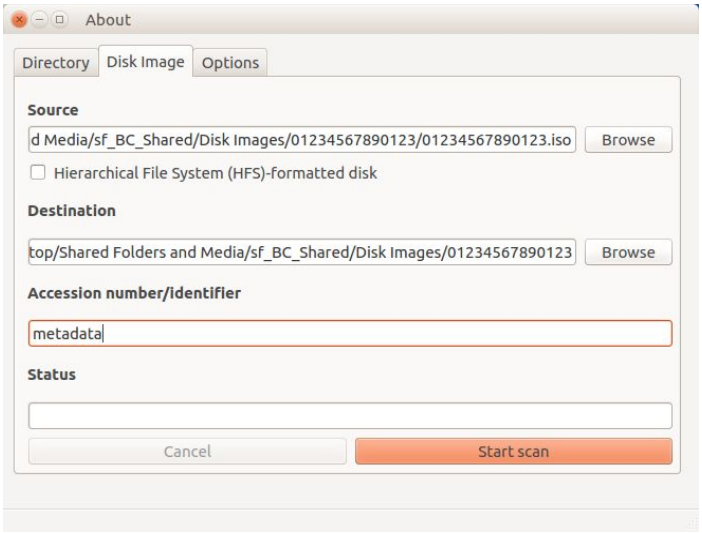
[STEP 9] Run Brunnhilde

Brunnhilde is the program used to run reports from the image. It is available in two forms as a Graphical User Interface (GUI) and command line. Instructions for the GUI is below. Command line instructions comes after.

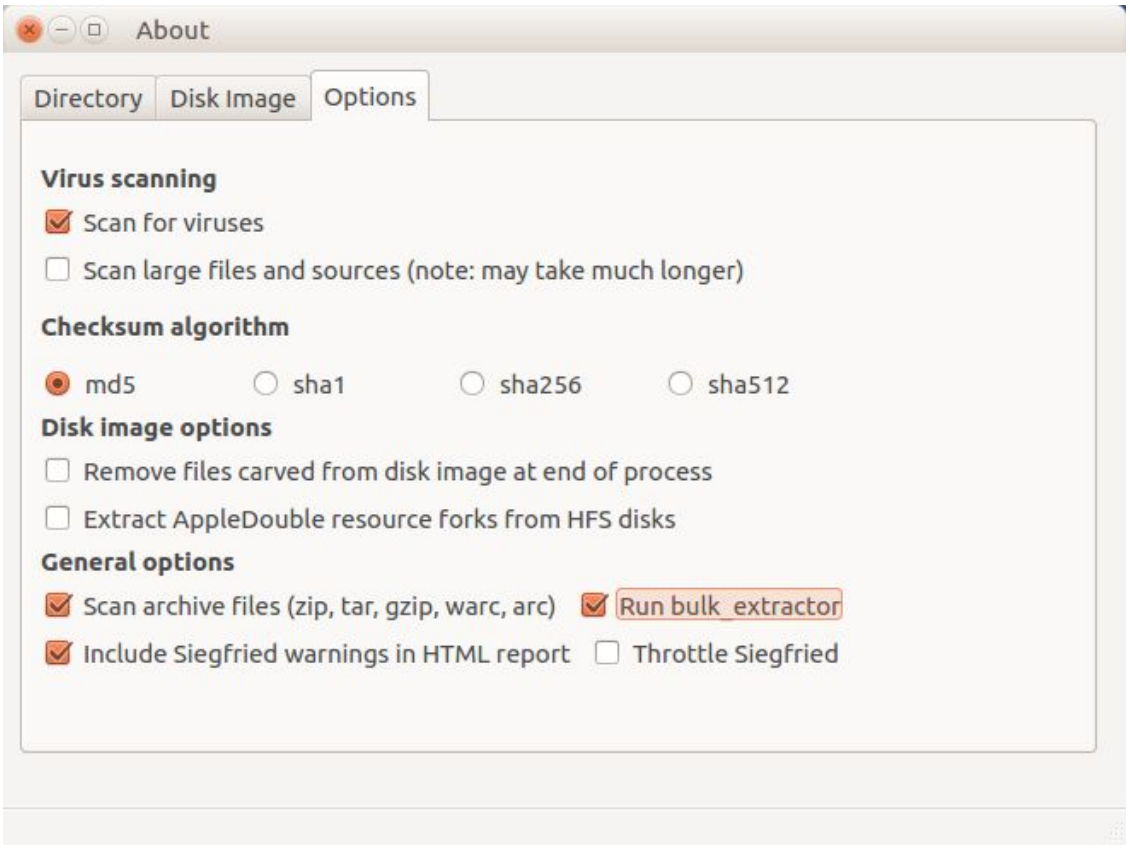
BRUNNHILDE GUI (only available on Darth)

- Open the Brunnhilde GUI in the *Forensics and Reporting* folder on the BC VM Desktop.

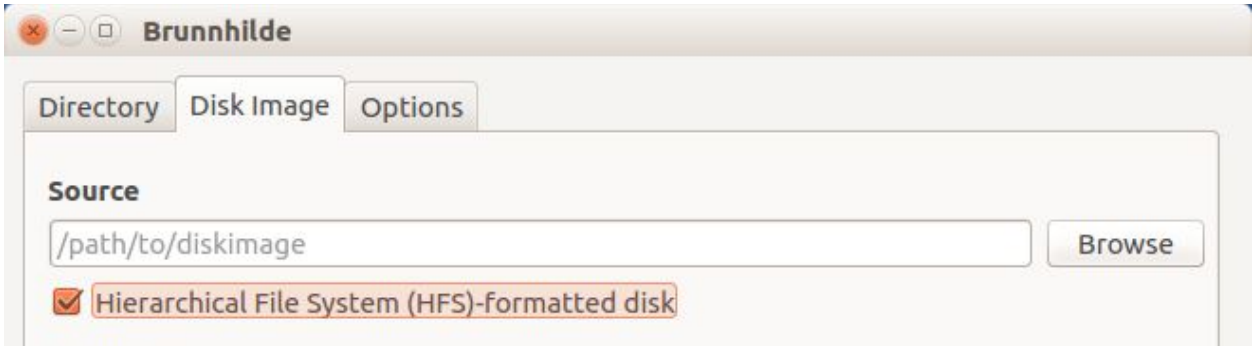
- You can choose to run a directory or a disk image. For our purposes, choose disk image.
    - In the "Source" section, browse for the IMG image file and select it.
    - In the "Destination" section, select the barcoded folder of the media.
    - In the "Accession number/Identifier" section, input the following output file name:
        - "metadata"



- In the options tab, make sure the following boxes are checked/selected:
  - Scan for viruses
  - md5
  - Scan archives files (zip, tar, gzip, warc, arc)
  - Include Siegfried warnings in HTML report
  - Run bulk-extractor



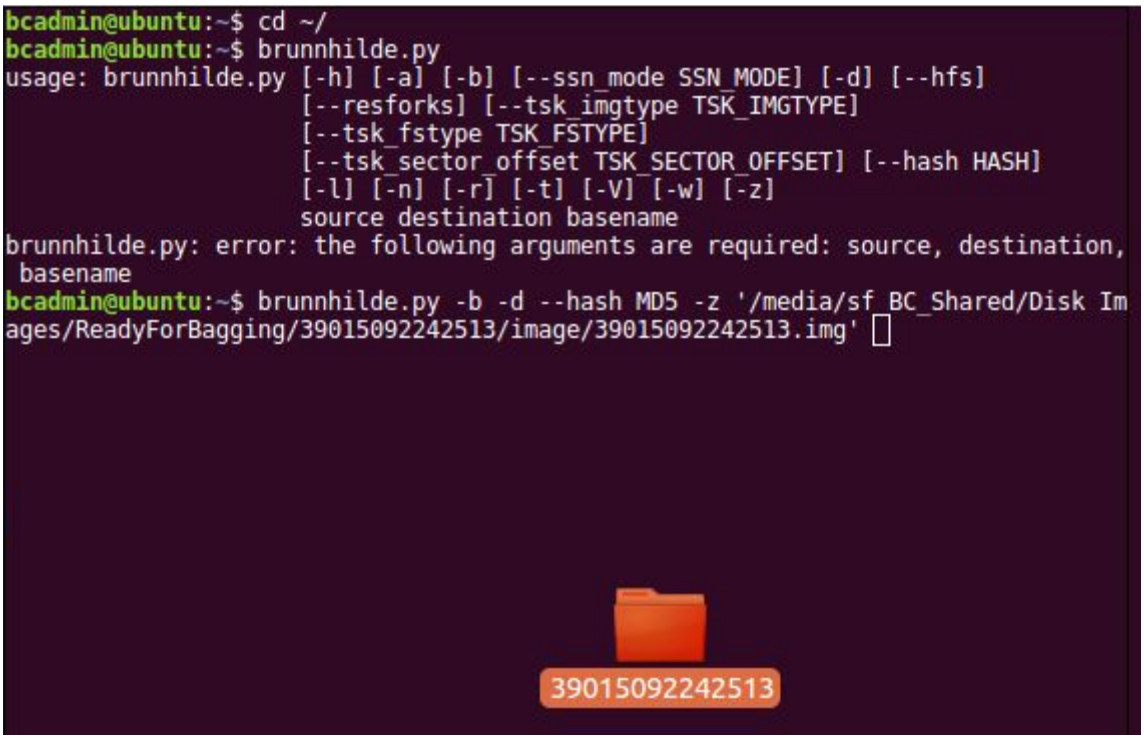
- Go back to the disk image tab and click “Start scan”.
  - It will typically take less than a second for Brunnhilde to run on the floppy disk image because of its small size. However, for larger files, it will take longer.
- Once the scan is completed, click “OK” to close the box that appears saying “Finished Brunnhilde scan complete.”
  - Sometimes, the scan will not work even though it creates the *metadata* folder you wanted it to create. If this happens, delete the newly created *metadata* folder and click the box that says “Hierarchical File System (HFS)-formatted disk” and start the scan again. Note it in the Transfer Inventory spreadsheet.



BRUNNHILDE COMMAND LINE

- Open Brunnhilde in the *Forensics and Reporting* folder. The icon looks like this:

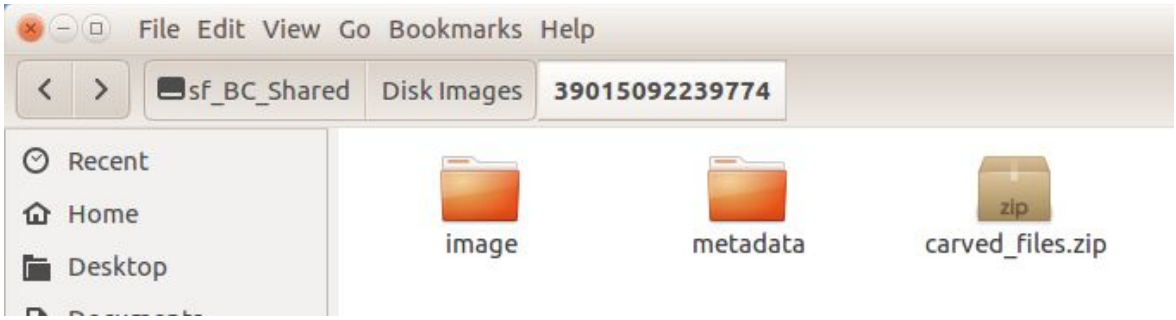
- Information on command line options are listed on Brunnhilde's Github page under "Usage" and on the "DPL Software Information" page in the shared *Digital Preservation Lab Documentation* Folder on Google Drive.
- Typical command line for one image file is:
  - brunnhilde.py -b -d --hash MD5 -z [image file file path] [barcode directory file path] metadata
- It is possible to drag and drop the image file or directory into the command line interface for file paths. Do not type it out.



- Hit “enter” on the keyboard. Any errors and reports (virus scan) will be shown on the command line.

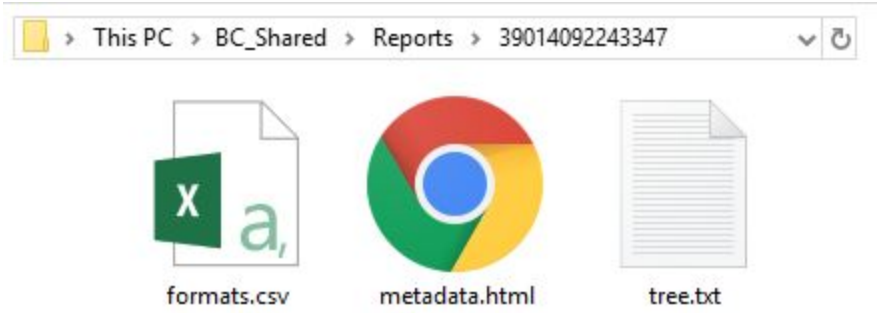
DIRECTORY CLEAN-UP

- Open File Explorer, navigate to the barcode folder and then the *bulk_extractor* folder. Click on the “view items as a list” icon in the upper righthand corner to change the display of the files. Filter by “size” and delete all files measuring 0 bytes.
    - These clutter up the *bulk_extractor* folder and do not provide us with any useful information about the disk image.
- Drag the photos taken earlier that are in the barcoded file into the newly created *metadata* folder.
- In the *metadata* folder, drag the *carved_files* directory into the same level as *metadata* and *image*.
- Right-click on *carved_files* and select “Compress…”. In the new window, keep the filename as *carved_files* and select “.zip” as the extension. Hit “Create” and delete the unzipped *carved_files* directory.



[STEP 10] Reports

- Go to the *Reports* folder created earlier and create a new directory with the barcode as the filename. Make a copy of the "tree.txt", "formats.csv", and "metadata.html" from the *metadata* folder.
    - The "tree.txt" and "metadata.html" files will be in the *metadata* folder while the "formats.csv" file will be in the *csv_reports* folder in the *metadata* folder.
    - This is the package of reports that will be sent to Special Collections.



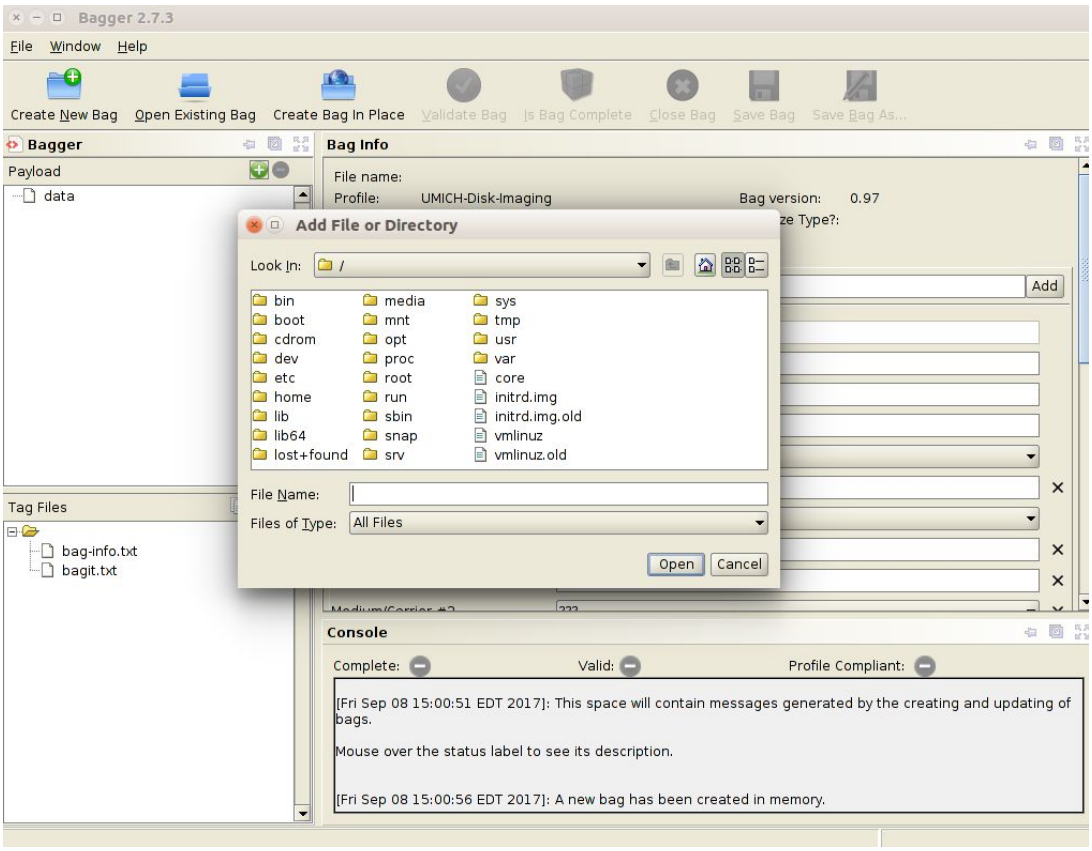
PART IV: BAG AND TRANSFER TO STORAGE

[STEP 11] Package SIP Using BagIt

****NOTE: It is important to follow these steps in order to ensure that the metadata saves properly.*

- Now, switch from the BC VM to the host computer's desktop, and open the "bagger.bat-Shortcut".

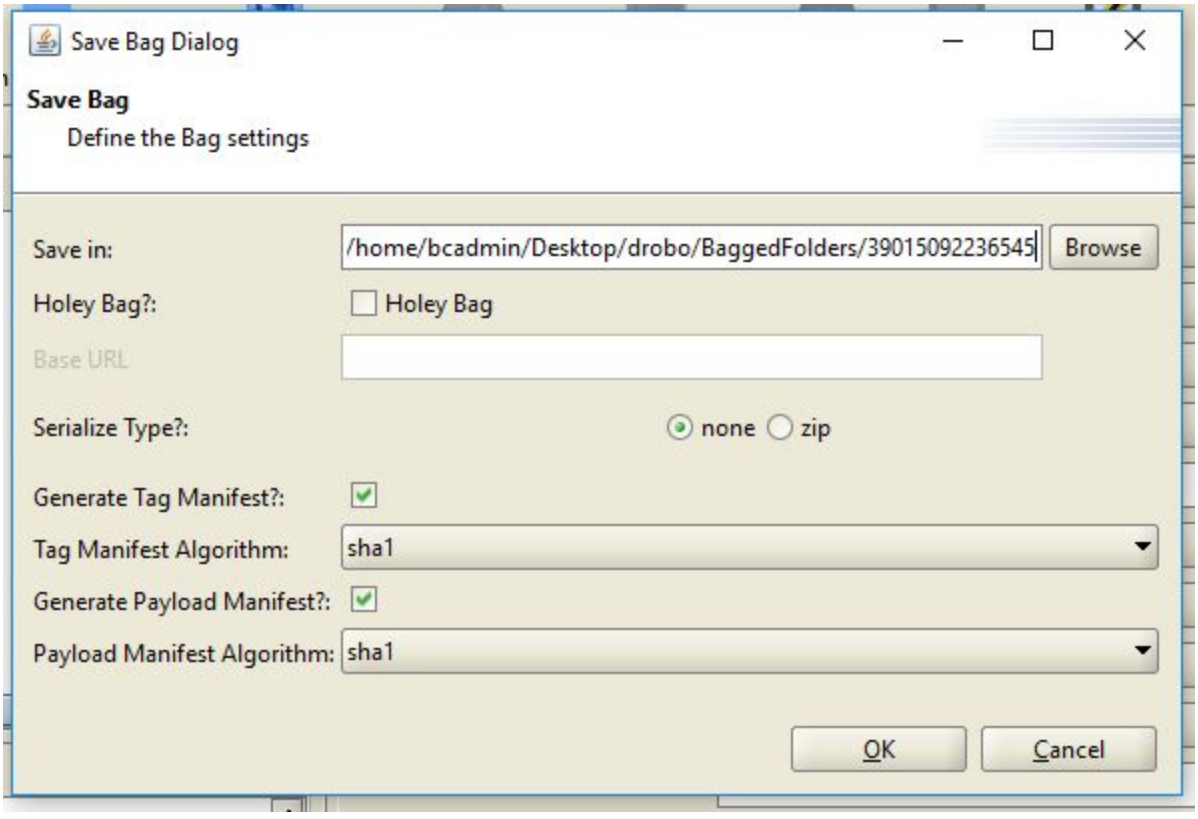
- On the upper lefthand corner, click “Create New Bag”.
- Select the “UMICH-DISK-Imaging Profile”.
- Click on the small green plus sign.
    - You will be prompted to “Add File or Directory”.
    - Select the parent folder containing the disk image and metadata (do not double click).
    - Click “Open”.



- Under the “Payload” section on the left, the folder you selected will show. Select (only click once) the folder prior to saving your bag.

- It is recommended that you do this in the beginning in case you forget to do it at the end.

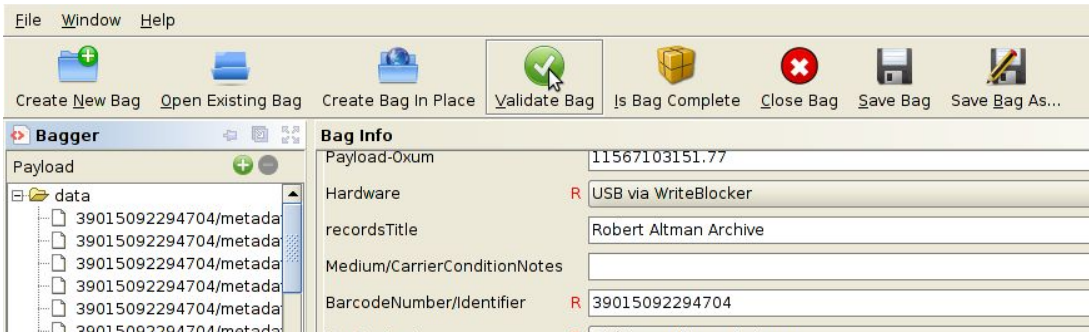
- Fill out the profile with metadata by consulting the Imaging Metadata Report.
- Once you are finished selecting the parent folder and filling out the metadata, click “Save Bag As...”
- In the “Save Bag Dialog”:
  - Navigate to the designated directory and create a file with the barcode as the title inside the parent folder.
  - Make sure to leave “Holey Bag” unchecked.
  - Select the following:
    - None
    - Generate Tag Manifest?:
    - Generate Payload Manifest?:
    - SHA-1 (for “Tag Manifest Algorithm:” and “Payload Manifest Algorithm:”)

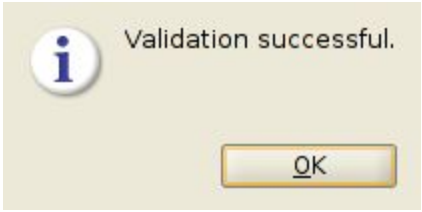


- Once all of the above are selected, click “OK”
- The Bag Info will then be generated. The results will appear in Bagger and the Bag will be created in the specified directory.

[STEP 12] Validate Bag

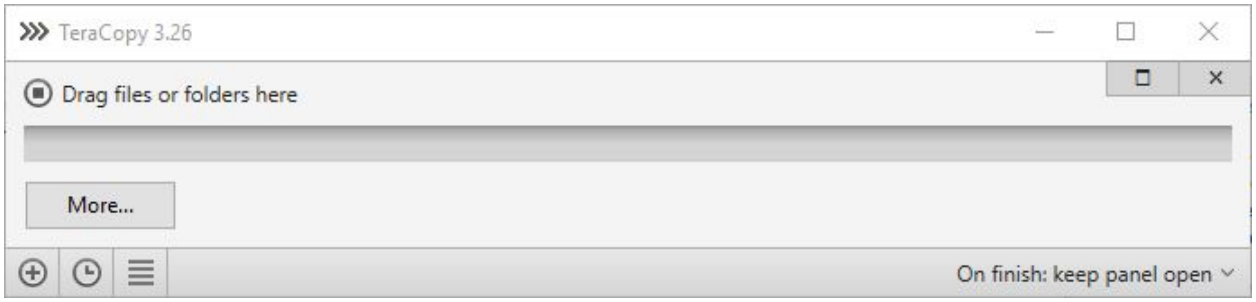
- After the bag is generated, click “Validate Bag” to validate the bag you just created.





[STEP 13] Transfer to the Network Attached Storage (NAS)

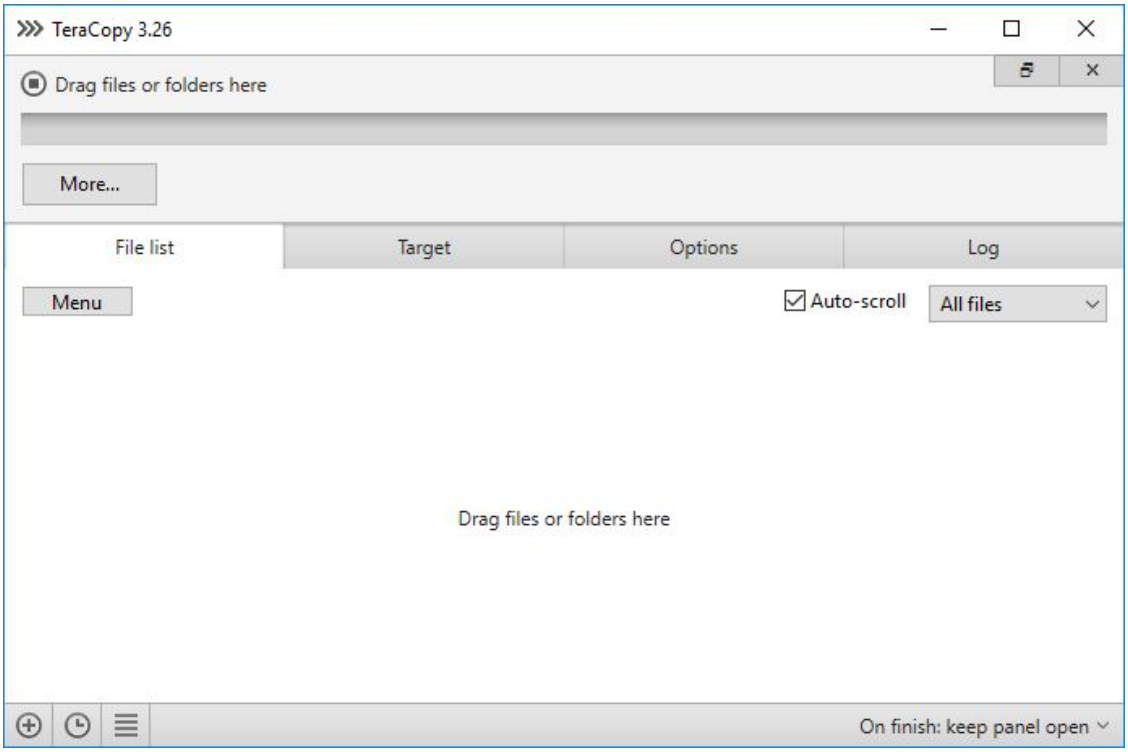
- At then end of the week, copy all files into public (\\192.168.255.10) (Z:) drive. This is the step prior to moving bags into Deep Blue.
- Search for “Tera” in Windows menu bar and select “Teracopy”. A window like this will appear:



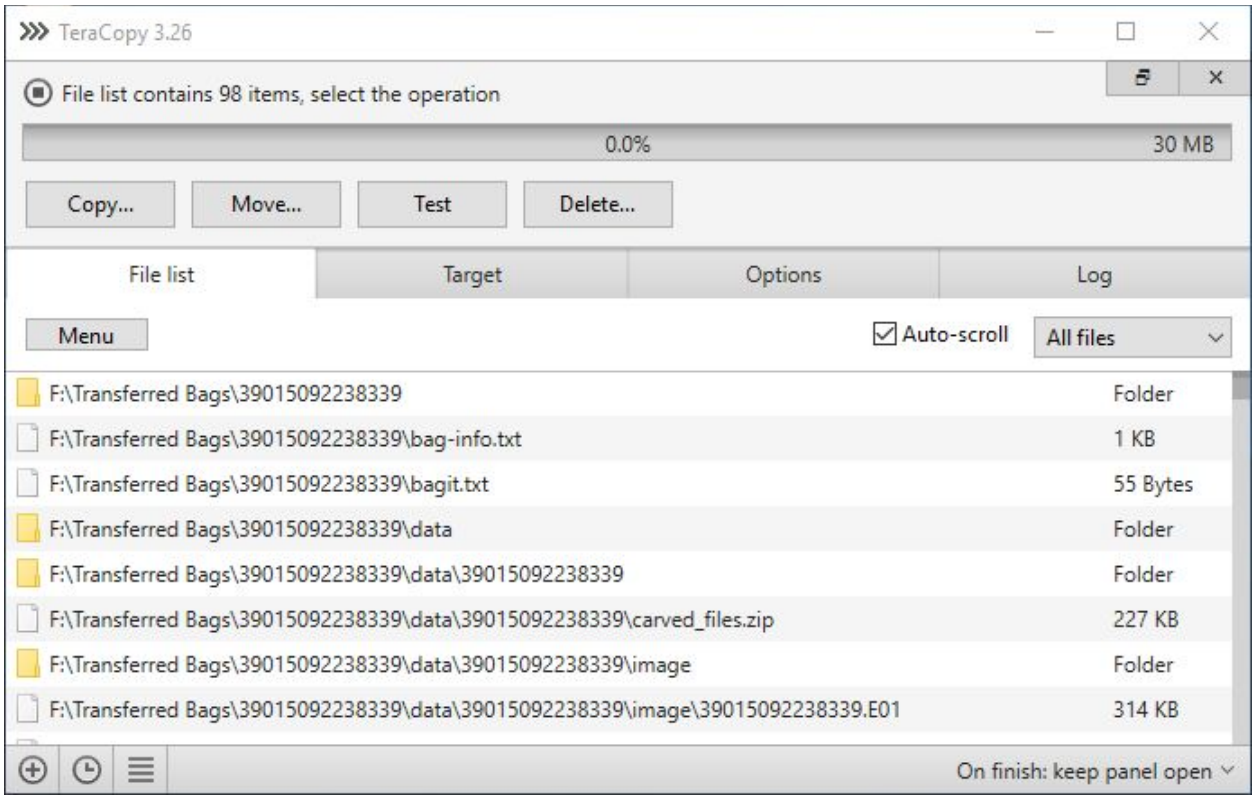
- Information from the previous transfer may load. If it does, click the on the

bottom left corner for a fresh window and the little “x” on the top right to exit out of the old window.

- Press the little window button on the top right to extend the window. The new area is where bags from the *Bagged Folder* are dragged. The screen should now look like this:



- Select files from *Bagged Folder* that are ready to be transferred. Drag them into TeraCopy. Files will appear like this:



- Under the “Target” tab, select “Browse” → public(//192.158.255.10)(Z:) → “TransferQueue”.
- The “Options” tab should stay the same with “Prompt on filename collision” and “MD5” as default.
- Press “Copy…” near the top of the screen. Selecting “Move…” will remove bags from *Bagged Folder* and into NAS.
- The “Log” tab reflects directories that have successfully/failed copying.
- When TeraCopy is finished, check the “TransferQueue” directory to see if files have been copied.
- Create a new Directory in *Bagged* named *UploadedtoDarkBlue* if it has not been created already. Move the bags you transferred using Teracopy into this new directory so you don’t copy it twice to the NAS.