# Permafrost Processing Workflow

Welcome to the Permafrost workflow! This general-use workflow is intended for Permafrost subscribers.

The workflow has been designed for groups of digital content accessible to you on your local computer or network. This content will be processed manually using Archivematica. It may be homogenous or heterogeneous in nature: the same or different content types, file formats, and metadata elements. It requires you to keep an eye on the process and make decisions at certain points.

The workflow does not include instructions for creating or processing disk images, using automation tools, using the backlog/appraisal functions (though you are welcome to use these!), or connecting to access repositories. Please contact us if you are interested in such things.

---

### ⓘ Tools you'll be using

First time? Please refer to the first-time tool installation and setup instructions.

- Exactly: A simple application that uses the BagIt file packaging standard to safely transfer digital materials. You might also want to try out alternative tools, such as Bagger or the Bagit Python tool.
- MakeCSV: A simple tool to create Archivematica-friendly CSV files for import.
- Archivematica: The main game!
- Swiftbrowser: A web interface for uploading and accessing transfer packages to the Ontario Library Research Cloud (OLRC). Note that Swiftbrowser works best with the Google Chrome browser.
- 7-Zip: A tool for extracting compressed AIPs.

---

### ⓘ Vocabulary

- **Archival package/AIP:** The final package of content that has been processed by Archivematica to make an AIP – the Archival Information Package.
- **Directory/ies:** Folders containing objects/files that make up your transfer.
- **Dissemination package/DIP:** A package containing access derivatives created by Archivematica - the Dissemination Information Package.
- **Object(s):** The individual files that make up your transfer.
- **Transfer:** The complete package that is passed through Archivematica. It is otherwise known as the SIP – the Submission Information Package.
- **Bag:** A type of structured data package that includes checksums and contextual metadata that can be used for transfer and validation of integrity across systems.

---

### ⓘ General Principles

- Work with Archivematica using a network connection - it does not function as efficiently over wifi.
- Use a browser like Chrome or Firefox. Internet Explorer/Edge and Safari are known to have issues.
- If something causes an error, do not despair! Let us know and we can try again.
- Generally only run one transfer at a time, though it is fine to let a transfer pause at a step and begin another one unless the transfer in question is very large.
- Descriptive or rights metadata CSV import is for item-level descriptions or rights data. If you only have transfer-level metadata to add, it's generally easier to add this via the Archivematica interface.
- Hitting the delete button ( 🔴 ) on a transfer will not remove that transfer from the workflow system. It only takes it out of the interface. If you want to cancel a transfer, always wait until presented with a 'reject' option.
    - That said, it's a good practice to remove your transfer from the dashboard after it has successfully stored and you are happy with the result. Generally aim to clear out transfer records on the dashboard on a regular basis.
- If you wish to delete a stored AIP, please notify us and give the name or UUID of the AIP. You can also put in a delete request via the interface (by heading to Archival storage, clicking on the specific AIP, and under 'Actions,' filling the delete tab form) but you still need to notify us.

## A. Prepare a Transfer

1. Organize the content you are interested in packaging for transfer. For example, gather digitized items belonging to a particular collection, fonds /series/file, or other logical grouping together. Place all archival objects and/or directories in a folder titled 'objects.' Please note that "objects" should be spelled lowercase.



---
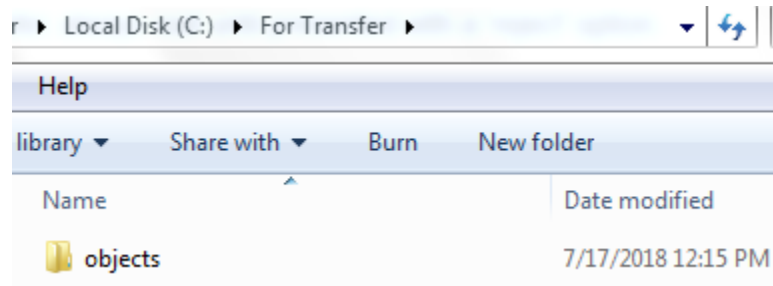
ⓘ **Transfer size guideline**

Note the following guideline for the size of your transfer:

- A Level 1 instance is most comfortably used for packages in the 0-50 GB range (unzipped) and consisting of fewer than 3,000 files.
- A Level 2 instance will be able to process larger-sized transfers over 50 GB and up to 250 GB (unzipped) and up to 10,000 files.

Note that these values do not take normalization into account as there is no hard-and-fast rule to calculate how large a transfer will be after normalization. See the community webinar on this subject for more information.

---

2. You can choose to upload the objects folder as-is (and proceed to step E), add descriptive metadata to be added into the final METS file (step B), or bag the transfer (step C & D ).

## B. Create a metadata.csv file (optional)

1. Ensure the MakeCSV tool is installed using the instructions.

2. Retrieve or compile descriptive metadata as necessary and have it on hand.

3. Run MakeCSV by right-clicking any file inside the objects folder and selecting 'makeCSV' from the drop-down menu. This will create a metadata folder and a metadata csv file at the level of the objects folder.

4. Open the metadata.csv file and enter Dublin Core headers and metadata, or other metadata as desired. For example:



blue arrow = transfer-level metadata for top-level 'objects' directory in transfer

orange arrow = item metadata for a file in 'objects' folder

red arrow = subdirectory-level metadata for a folder within 'objects'

green arrow = item-level metadata for items in a subfolder of the 'objects' directory

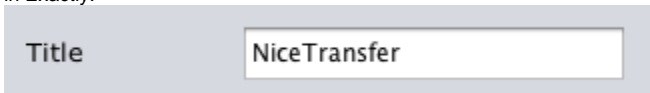It's up to you what metadata to enter. You can find a good guide to Dublin Core in the AtoM documentation.

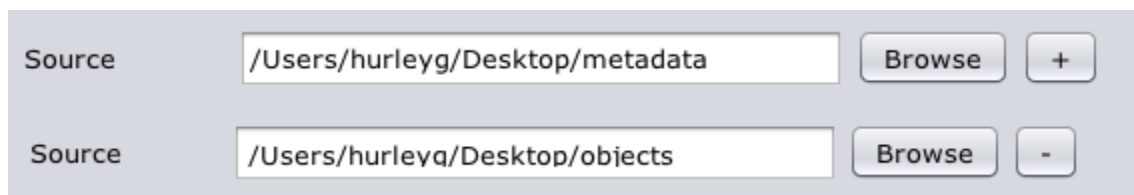5. Once you're done, here's how your directories should generally look:



There are a number of other import options available, including including adding submission documentation (such as a donor agreement or other files about the transfer), rights metadata or access copies – please visit the Archivematica transfer documentation for details on these features.

## C. Bagging Yourself a Transfer (optional, but recommended)

1. Open Exactly or a similar bag-creating tool as you wish. The instructions below pertain to Exactly.

2. Give your transfer a title. This title will be reflected in the final package name. Please note that there is a 250 character limit to titles for transfers in Exactly.



3. Click 'Browse.' A file directory browser will pop up.

4. Navigate to the location of your 'objects' and 'metadata' folders as seen above in step B.5, where applicable.

5. Select all folders for transfer as applicable. You can add multiple folders to the transfer by selecting more than one folder in the browser, or hitting the little + sign button.



6. Check the box 'zip files?'

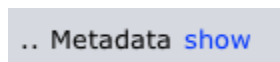7. Select a destination on your computer for the transfer package.



Note: The bagging process may take longer for larger groups of files.

## D. Adding Metadata to Your Bag (Optional)

Metadata for your bag transfer can be added at this stage. It's optional, but a good idea would be to add a few fields about the bag's creator and source (you). It also allows to link sequences of bags if you are dividing a collection into several transfers. It doesn't have to be detailed.

You can find examples of bag metadata here.

1. Under the 'Metadata' heading as pictured below, there should be a little blue button that says 'show.' Click on this.



2. If you click the tick box "Show all reserved fields," Exactly will expand to show the metadata fields that make up part of the BagIt specification. The standard provides a description and example of the fields. Note that some of them (Bag date, bag size, payload oxum) are auto-generated.



Archivematica can parse this metadata into the final METS file, but it needs to be formatted correctly: field names cannot have spaces. Default field names are formatted correctly; ensure any custom ones you might wish to enter are also formatted correctly.

Take a look at the example below - the final field name is not formatted correctly.

| Label | Source-Organization | - |
| --- | --- | --- |
| Value | An organization | |

| Label | Organization-Address | - |
| --- | --- | --- |
| Value | An address | |

| Label | Custom-Field | - |
| --- | --- | --- |
| Value | Something else | |

| Label | Source Organization | - |
| --- | --- | --- |
| Value | | |

3. Make sure to remove any fields you do not use by clicking the '-' button beside the fields.

4. Ensure there are no line breaks in your metadata - this will cause your bag to fail later on in Archivematica.

5. Click 'Save' when you're done. This schema will be saved so that you don't need to re-add it during the next transfer. But of course, you can choose to change these settings as needed.

6. You're almost done this part! Click the big 'transfer' button on the right side of the pane! The transfer will create a .zip file in the location you specified.



Transfer

Validating Bag Info
Metadata saved successfully.
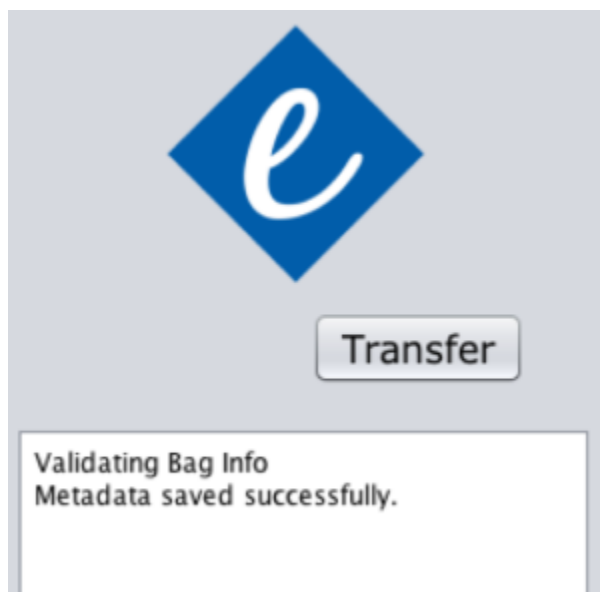
## E. Uploading your transfer to the OLRC

The OLRC Swiftbrowser interface has a drag-and-drop functionality. This should work for all browsers with bagged .zip files. However, if you choose not to zip files and are processing a standard transfer or an unzipped bag directory, **you must be using the Chrome browser for the drag-and-drop functionality to work**.
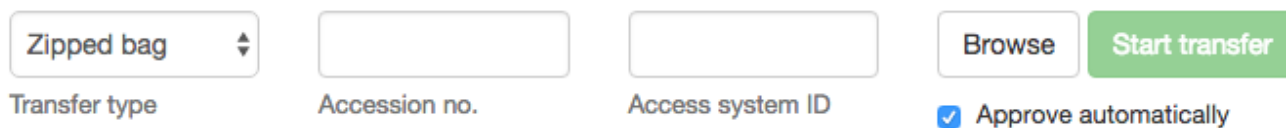
1. Go to https://olrc.scholarsportal.info/swiftbrowser/ and log in using the credentials on your main page.

2. Click on the 'TS' directory.

3. Drag and drop your .zip file or objects directory into the window. Or, click the blue 'Upload files' button at the top right. If your upload is larger than 5GB, click on the blue "large upload" button on the top right. Note that if you need to interrupt a large upload you can do so, but to resume it, you have to refresh your browser window and/or restart the upload.

4. You'll see a window like this pop up:

## Add Files

| | NiceTransferTest.zip | 791.30 KB | Start | Cancel |
|---|---|---|---|---|

**+ Add files...**   **Add Directories...**   **↑ Start upload**   **⊗ Cancel upload**   **Close**
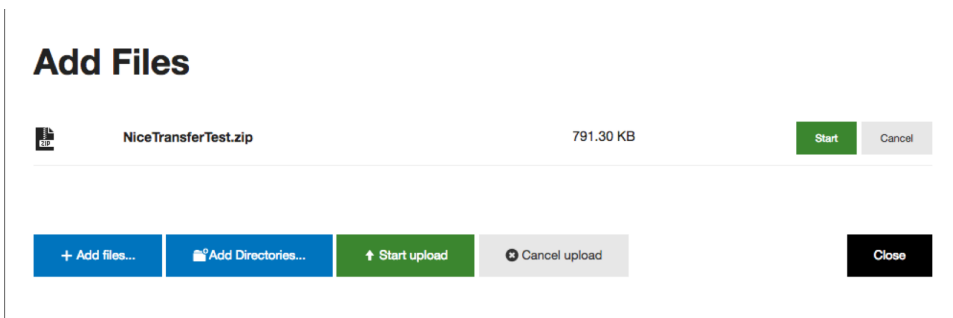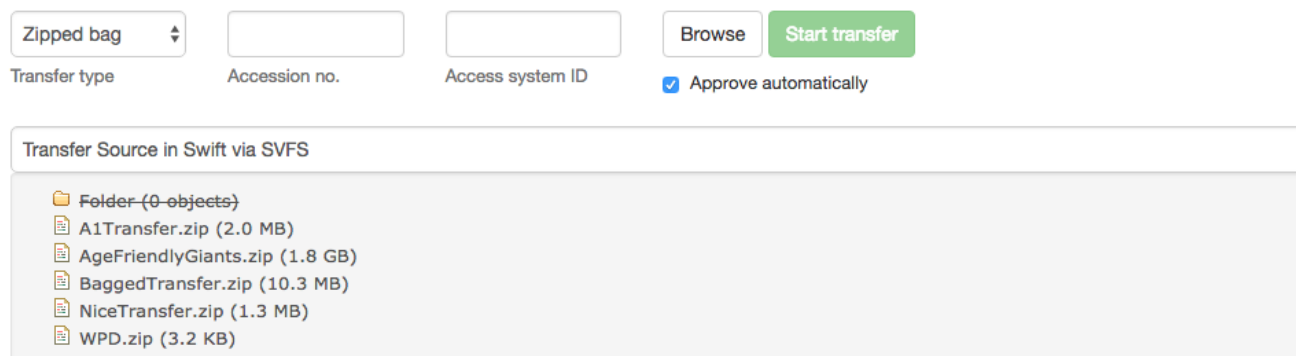
5. Click on the green 'Start upload' button.

6. Your file will upload. You should see a success message in green at the top of the frame.
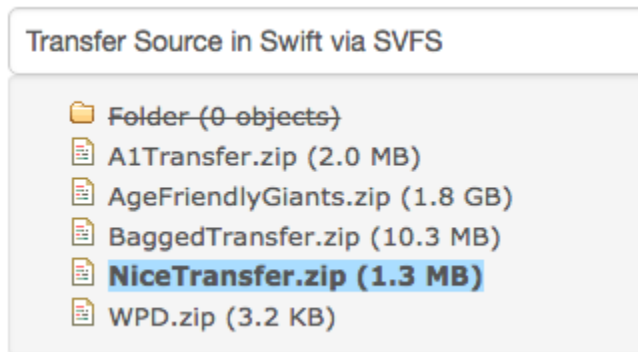
7. You're done!

## F. Prepare the Transfer Package Using Archivematica

1. Log into Archivematica at the URL and with the credentials provided.
2. Near the top of the page, you'll see a transfer initiation pane as below.

Zipped bag ⬍ | | | Browse | **Start transfer**
Transfer type | Accession no. | Access system ID | ☑ Approve automatically

3. Under 'Transfer type' select the appropriate transfer type, e.g. 'zipped bag' for bags created in step C & D, or 'standard transfer' for basic directories.

4. You can enter an accession number if you'd like. For a non-zipped standard or bagged transfer you must enter a transfer name.

5. Hit the 'Browse' button.

6. A window will pop up showing the available applicable transfers in the transfer source. If you are processing an unzipped folder, you should select the folder. Don't worry if the files in that folder are crossed out in the window. Note that it may take a few seconds for your transfer to show up if you recently uploaded it to the OLRC. You may have to refresh your page on Archivematica as well.

Zipped bag ⬍ | | | Browse | **Start transfer**
Transfer type | Accession no. | Access system ID | ☑ Approve automatically

Transfer Source in Swift via SVFS ⬍

📁 Folder (0 objects)
📄 A1Transfer.zip (2.0 MB)
📄 AgeFriendlyGiants.zip (1.8 GB)
📄 BaggedTransfer.zip (10.3 MB)
📄 NiceTransfer.zip (1.3 MB)
📄 WPD.zip (3.2 KB)

7. Find the file/directory you just loaded, select it, and click the blue 'Add' button. The transfer will be added to the top of the pane. If you add additional transfers at this stage, they will be processed separately.

8. Click the green "Start transfer" button and you're off to the races! Note: if the "Approve automatically" checkbox is clicked under the "Browse" button as pictured above, your transfer will begin running up to the file identification step. If the box is not checked, you will have to approve the transfer to initiate it.

Note that the bigger the initial package, the longer it will take to download from the OLRC to the server where Archivematica is installed. You may have to wait anywhere from a few seconds to several minutes depending on the size of your transfer package.

## G. Process the Transfer

The transfer steps are determined based on the standard configuration for Permafrost, and there are a number of options along the way. It also does not make use of the backlog/appraisal functions, but you are welcome to do so. Consult the appropriate documentation to do so here.

At any time until the normalization step is reached, you can add descriptive and rights metadata. You do so by clicking on the little  icon at the top of the transfer bar.

Note that descriptive metadata can only be added from the interface at the level of the transfer, rather than for individual files or directories. Otherwise, you must import this metadata.

1. **Approve transfer**: If the "Approve automatically" checkbox is clicked under the "Browse" button as pictured under step 6 above, your transfer will begin running. If the box is not checked, you will have to approve the transfer to initiate it. You can choose approve or reject (you can reject if you want to start over for some reason or another). Please note that the  button will only hide the transfer from view - it will not cancel the transfer.



2. A number of services will run, including file format identification, characterization, and validation. Some of these processes may experience failures; if you run into issues, contact us. At the end, you **have the option of creating a single SIP and continuing processing**. The general case is to select 'create single SIP.' If you want to use the Appraisal tab, select "Send to backlog." For information on this function, please consult Archivematica's documentation here.

3. **The SIP will move to the Ingest page**. You have to click on the Ingest tab to continue. Under 'Ingest' a number of services will run.
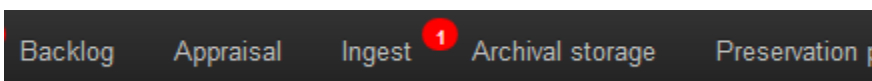


4. **It will pause at normalization**. Select 'normalize for preservation' or 'do not normalize' to create just an AIP. If you want to create access copies (i.e. a DIP), you can select 'normalize for preservation and access' or 'normalize for access.' The question of whether to normalize or not is bigger than can be addressed here - feel free to reach out to discuss it with us! You may also consult the Permafrost community webinar on the subject.

5. After normalization, you can **review and approve normalization** by clicking on the little report icon: ![icon] This takes you to a separate tab where you can see the results of the normalization process. If you click the white "Review" button next to the micro-service task, you can also see the list of files in the AIP-to-be at this point, including those resulting from normalization.

| File name | File format | Preservation normalization attempted | Preservation normalization failed | Already in preservation format | Access normalization attempted | Access normalization failed | Already in access format |
|---|---|---|---|---|---|---|---|
| image2.jpeg | JPEG | Yes | No | No | Yes | No | Yes |
| image1.jpg | JPEG | Yes | No | No | Yes | No | Yes |

6. Back on the main transfer page, if you click the white "Review" button, you can click through the files created as part of the normalization process.
7. Assuming normalization was successful, select "approve."

8. A final set of functions will run, including ingest of the descriptive metadata into the METS file, and processing of any submission documentation.
9. If you chose to normalize for access, the **Store DIP** option will come up first, followed by the **Store AIP** option. It's the best practice to deal with the AIP first, so wait for this option to arrive and store the AIP before the DIP. The rationale is that if there's some error in the AIP, you don't want to replicate it in the DIP.

10. You'll have the option to **store or reject the AIP**. The normal case is to store, but it's possible you might want to pause at this point or start over. After selecting "Store AIP," there will be a few more automatic steps, such as verifying and indexing the AIP. When the green checkmark appears after "Job: Remove the processing directory," the AIP will be stored - by default it will be on the OLRC. You can now search for and download it from the Archival Storage tab. See Accessing AIPs below.

| Submission Information Package | UUID | Ingest start time | |
|---|---|---|---|
| FieldNotes | 02fadc76-a3ab-4ec5-8ebc-c1f4d59562bf | 2019-02-25 16:12 | |
| ▶ Microservice: Store AIP | | | |
| Job: Store AIP  Review  [?] | | Awaiting decision | ✓ Actions |
| Job: Move to the store AIP approval directory | | Completed successfully | - Reject AIP |
| ▶ Microservice: Prepare AIP | | | - Store AIP |
| ▶ Microservice: Add README file | | | |
| ▶ Microservice: Generate AIP METS | | | |
| ▶ Microservice: Bind PIDs | | | |
| ▶ Microservice: Process metadata directory | | | |

11. **For the DIP**, you will be prompted with the option to store the DIP. When the option to **Store DIP** is available, select "Store DIP" or reject it by selecting "Do not store," if you want. By default, the DIP will be stored on the OLRC. It will be accessible there - not through the Access tab in Archivematica, which controls only DIPs uploaded to a connected access system like AtoM. See the instructions for Accessing DIPs below.

12. You're done!
    a. The default is to compress AIPs and DIPs. You'll need a program capable of extracting 7z files to open it. I use The Unarchiver. 7-Zip is another common method that works well in Windows. However, you can choose to turn off compression in consultation with SP. For more information and instructions, see Accessing AIPs below.

    b. As mentioned elsewhere, please send a note to permafrost@scholarsportal.info if you wish to delete stored AIPs or anything in the Backlog.

## H. Accessing AIPs

You can search and download AIPs via the Archivematica interface.

1. Click on the "Archival storage" tab.

2. From here you can search for datasets using the search field at the top.

3. To access a dataset, click on its name or UUID.

![Archivematica archival storage browse screenshot]

4. To download an AIP, click on the "Download" button (circled in purple).

## NiceTransfer Archival Information Package

| | |
|---|---|
| **UUID** | 1519cd2d–5d29–4fb8–b3fe–cf8ac39a2011 |
| **Size** | 7.44 MB |
| **Date stored** | 2018-08-02 08:46 |
| **Status** | Stored |
| **Encrypted** | False |
| **Location** | Download   [...]/NiceTransfer–1519cd2d–5d29–4fb8–b3fe–cf8ac39a2011.7z |
| **Pointer file** | View |

## Actions

| Upload DIP | Re-ingest | Delete |
|---|---|---|

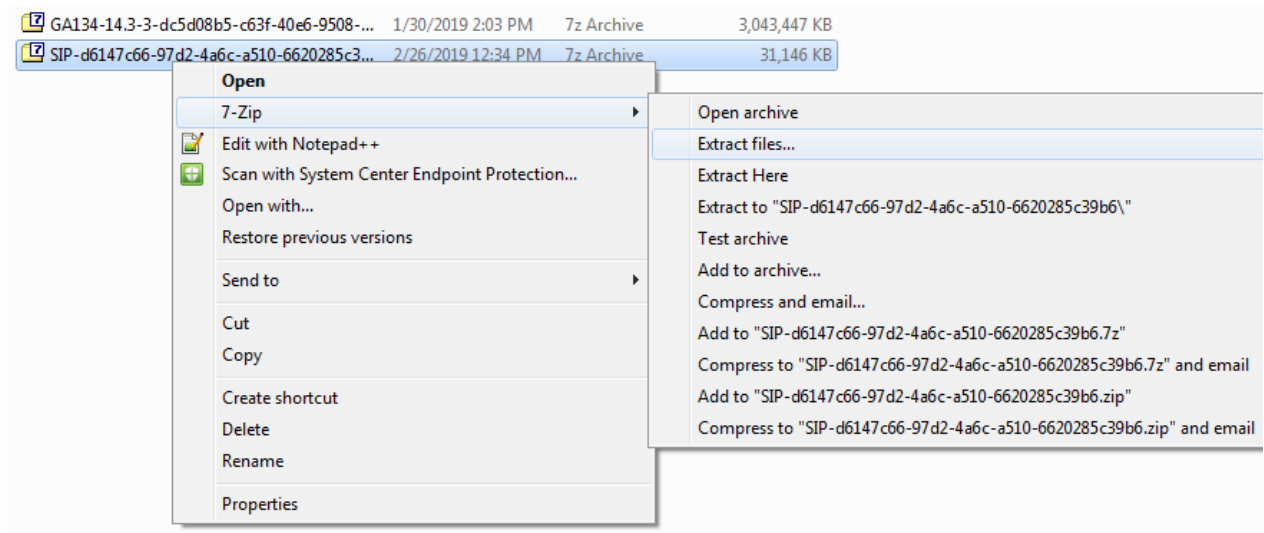## Metadata-only DIP upload to AtoM

**Insert slug \***
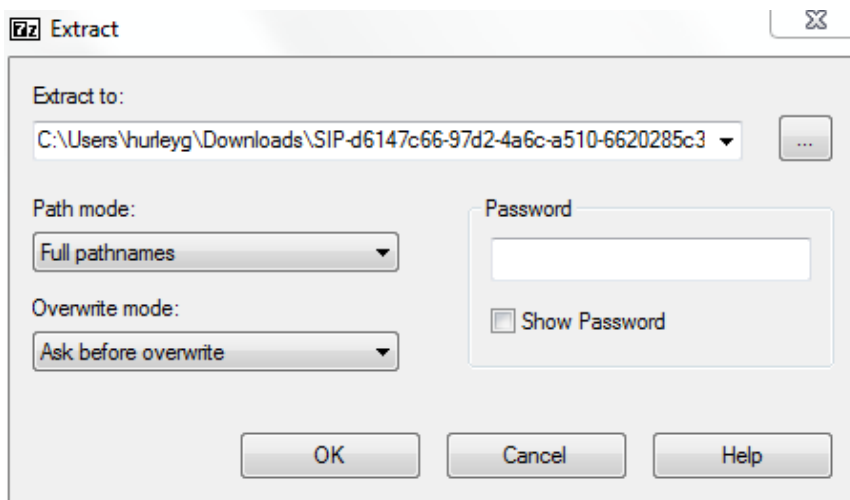
Only compatible with AtoM 2.4 or newer.

**Upload**

5. Additional actions, such as reingest and deletion are available under "Actions." If you wish to delete a stored AIP, please notify us.

6. Archivematica by default compresses AIPs as 7z files, an open source type of zip file. This is configurable - contact us if you wish to disable this configuration. See Permafrost First Time Tool Setup - 7-Zip for instructions on installing 7-Zip in Windows. If using OSX, try out The Unarchiver. Here's how to open 7z files in Windows once you've installed 7-Zip:

  A. Right-click on the file. Under 7-Zip, select "Extract files."

| | | | | |
|---|---|---|---|---|
| GA134-14.3-3-dc5d08b5-c63f-40e6-9508-... | 1/30/2019 2:03 PM | 7z Archive | 3,043,447 KB | |
| SIP-d6147c66-97d2-4a6c-a510-6620285c3... | 2/26/2019 12:34 PM | 7z Archive | 31,146 KB | |

| | |
|---|---|
| **Open** | |
| 7-Zip ▶ | Open archive |
| Edit with Notepad++ | Extract files... |
| Scan with System Center Endpoint Protection... | Extract Here |
| Open with... | Extract to "SIP-d6147c66-97d2-4a6c-a510-6620285c39b6\" |
| Restore previous versions | Test archive |
| Send to ▶ | Add to archive... |
| | Compress and email... |
| Cut | Add to "SIP-d6147c66-97d2-4a6c-a510-6620285c39b6.7z" |
| Copy | Compress to "SIP-d6147c66-97d2-4a6c-a510-6620285c39b6.7z" and email |
| Create shortcut | Add to "SIP-d6147c66-97d2-4a6c-a510-6620285c39b6.zip" |
| Delete | Compress to "SIP-d6147c66-97d2-4a6c-a510-6620285c39b6.zip" and email |
| Rename | |
| Properties | |

  B. Another window will pop up. Select OK.

C. Navigate to your file folder and check out your AIP. You can open your METS file with a text editor like Notepad++ or Sublime Text, or upload it to METSFlask (or run METSFlask on your own system if you want to keep the files private).

D. Want more info on looking at your AIP? Check out You've Just Made an AIP: A Checklist.

## I. Accessing DIPs

Accessing stored DIPs is slightly less intuitive. You must access and download your stored DIPs via the OLRC. The DIPs are nested by Archivematica within a UUID-based folder structure that is not the same as the AIP UUID. Users can retrieve this UUID path by following the steps below:

1. On the ingest pane, click on the gear icon next to the "Store DIP" micro-service entry.

2. In the "STDOUT" window that opens up, you'll see the DIP path recorded in the area that is highlighted below starting with "Checking if."

3. Users then can log into the OLRC and click through the folders within the DIP directory that match this UUID.



Permafrost users also receive lists of AIPs and their associated DIPs in storage on a monthly basis - visit your specific institution's service page to access these.