Archivematica Demo Sandbox

Welcome!

Thank you for stopping by! We're glad you're interested in trying out the Dataverse-Archivematica demonstration sandbox hosted by Scholars Portal. This page contains information on how to access the sandbox, notes on its limitations, and a description of the workflow to use it.

Scholars Portal sponsored Artefactual Systems Inc. to develop the ability for the preservation processing tool Archivematica to receive packages from connected Dataverse instances. The integration was released as part of Archivematica 1.8 in 2018. The sandbox has been updated to Archivematica v. 1.13.2.

You can read more about the project at the Dataverse-Archivematica wiki page as well as in Meghan Goodchild and Grant Hurley's 2019 iPres paper and presentation slides here: https://osf.io/wqbvy.

There are a handful of documented issues with the integration available in the Archivematica issues repository in GitHub. These impact especially the treatment of tabular derivative files and how they are represented in the AIP's METS file.

Please note that individuals at institutions that are members of the Ontario Council of University Libraries (OCUL) can access the sandbox as-is with the credentials below. If you are from an institution outside of OCUL, please email dataverse@scholarsportal.info to request access.

We are **seeking feedback** on the integration to identify areas for future development. Please send your feedback to dataverse@scholarsportal.info.

If you have any questions, or are experiencing technical issues? Send them to dataverse@scholarsportal.info too!

Accessing the Sandbox

The sandbox is available at: https://archocul.scholarsportal.info/

Username: **test**
Password: **testtest**

Members of OCUL can access the sandbox without further setup - you're good to go!

All users outside of OCUL schools must submit a request to gain access to the instance. Please email **grant@scholarsportal.info**.

Notes on the Sandbox

- The sandbox is connected to a test Dataverse repository in Scholars Portal's demonstration Dataverse. The test Dataverse contains three sample datasets for testing purposes.

    - If you'd like to submit your own data to test, you may do so but **you will need to create an account on the Demo Dataverse** instance **first**. Your dataset will be reviewed by Scholars Portal staff and you will be informed whether it is suitable to open for testing. To submit a dataset for review, navigate to the Archivematica Test Dataverse and click the "Add Data" button on the right side. If you need instructions on how to add datasets, see this guide starting from step 3. Please also note that if you submit your own test data, it will be **available to anyone who has access to the sandbox**.

- The sandbox refreshes nightly. If you wish to keep any stored data, please download it immediately.

- If you use the sandbox, we would appreciate **your feedback** to identify areas for future development. Please send your feedback to dataverse@scholarsportal.info or complete this Google form, which you can submit anonymously.

Want to Learn More?

Visit the Dataverse page on Archivematica's wiki, as well as the Dataverse documentation for Archivematica and the Archivematica storage service for lots more documentation.

Workflow

This workflow is specific to the Dataverse integration. OCUL users can also request instructions on testing other kinds of transfers by e-mailing permafrost @scholarsportal.info.

Need a quick intro to Archivematica? Check out the Overview guide in Archivematica's documentation.

Made an AIP and not sure what the heck it is? Check out the Archivematica Documentation's page on this subject.

ⓘ **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. This is the package that gets delivered from Dataverse.
- **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**

- Use a browser like Chrome or Firefox. Internet Explorer/Edge and Safari are known to have issues with Archivematica.
- Generally only run one transfer at a time, though it is fine to let a transfer pause at a step and begin another one.
- 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.
- Please treat your fellow participants respectfully - if you happen see them processing a transfer (since everything is visible to each user) don't interfere with it!
- Please do not change any settings under the Administration tab.
- Have questions? Something failing? Let us know at dataverse@scholarsportal.info

---

A. Starting a Transfer

1. Log into Archivematica at the URL and with the credentials provided above.

2. Near the top of the page, you'll see a transfer initiation pane as below.

3. Under 'Transfer type' select "Dataverse" as pictured above.

4. Enter a transfer name. You can leave "Accession no." and "Access system ID" blank.

5. Hit the 'Browse' button.

6. A window will pop up showing the available applicable transfers in the transfer source. Click on the dropdown menu that shows "Transfer Source in Swift via SVFS" and select instead "Archivematica Test on Demo Dataverse." The three sample datasets will appear as pictured below.

7. Select one of these transfers by clicking on it.

8. 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.

9. Click the green "Start transfer" button and you're off to the races! You may have to wait a few hot seconds until the transfer begins processing, so please be patient. 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.

B. Processing a Transfer

The transfer steps are determined based on a standard configuration with some option-based stops 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 use these functions here.

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 up to the file identification step (#2 below). 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. 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 (a little action number will appear!) to continue.. Under 'Ingest' a number of services will already be running.
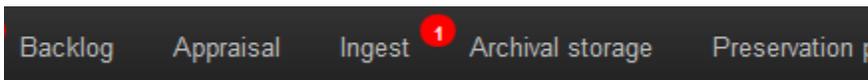
4. **The processing will pause at Normalization**. Normalization means that Archivematica will identify files in the transfer and convert a copy of the original file into a preservation-friendly format, based on its default policies. Select "Normalize for preservation" to create an AIP only. If you want to create additional access copies (i.e., a DIP), you can select "Normalize for preservation and access." You can also choose not to normalize by selecting "Do not normalize."

5. After normalization, you can **review and approve normalization** by clicking on the little report icon: ▨ This takes you to a separate tab where you can see the results of the normalization process.

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, it will display the files created as part of the normalization process.

7. Once you've decided that normalization was successful, choose to **approve** (or reject or redo if you're not happy).

8. Some more functions will run.

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 a few more automatic steps, the AIP will be stored - by default it will be on the Ontario Library Research Cloud (OLRC), Scholars Portal's storage cloud. You can search for and download the AIP from the Archival Storage tab in Archivematica.

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.

C. 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 AIPs using the search field at the top.

3. To access a stored AIP, click on its name or UUID (universally unique identifier).

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 re-ingest and deletion are available under "Actions." Note that re-ingest does not function with Dataverse packages, and stored packages are automatically deleted every evening, so there is no need to submit a delete request.

6. Archivematica by default compresses AIPs as 7z files, an open source type of zip file.

(i) **Installing 7-Zip**

If you are in Windows, download 7-Zip to extract 7z files. In OSX, try out The Unarchiver.

By default, Archivematica compresses AIPs as 7z files, which is an open source type of zip archive. When you download a 7z-compressed AIP, you will need the 7Zip extraction software to extract the files. 7z files can be opened in Windows with the 7-Zip utility. OSX users can use the Unarchiver. For 7-Zip in Windows:

1. Navigate to https://www.7-zip.org/

2. Download the 32- or 64-bit version depending on your installation of Windows. Not sure which one you've got? Consult this guide.

3. Double-click the .exe file and install as required by your system. You should be able to run 7-Zip without administrator privileges, but you may need to consult your IT folks if you do not have the appropriate permissions.

4. Consult the documentation below for instructions on opening 7z-type AIPs.

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."
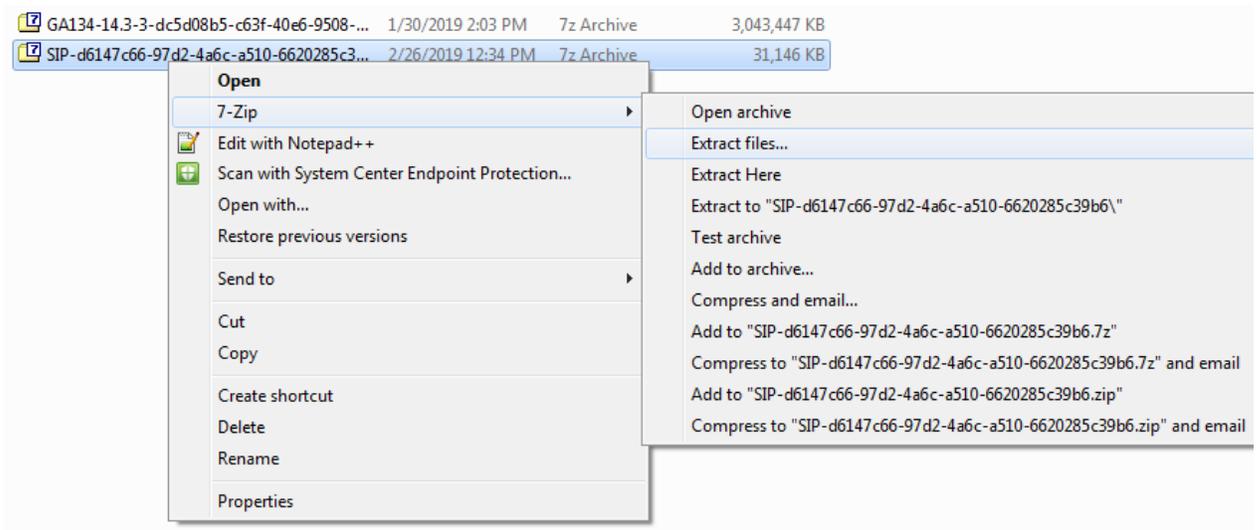
B. Another window will pop up. Select OK.
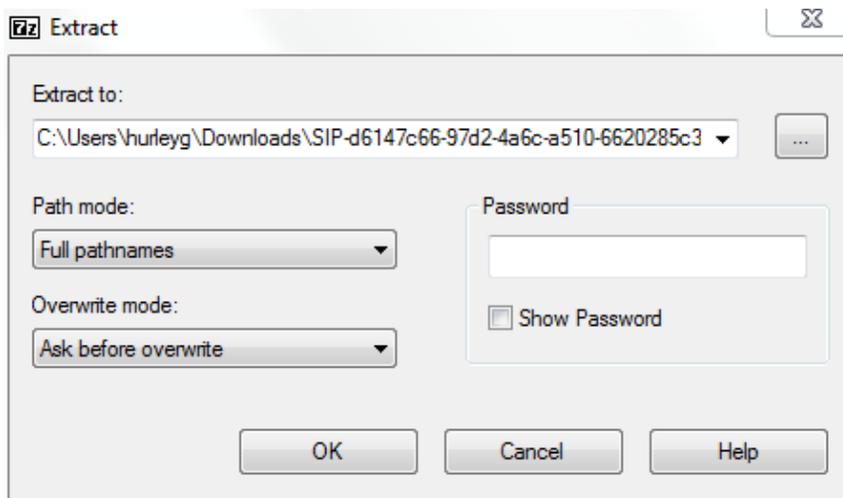
C. Navigate to your file folder and check out your AIP.

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). Not sure what a METS file is? What to know more about the structure of an AIP? Check out the Archivematica documentation.

D. Accessing DIPs

Accessing stored DIPs is not offered as part of the sandbox, as doing so requires navigating directly to storage. Please contact us if you wish to access a DIP.

Thank you!

We are currently **seeking feedback** on the integration to identify areas for future development. Please send your feedback to dataverse@scholarsportal.info.