Islandora Westvault automates the preservation of your objects in COPPUL's WestVault LOCKSS network.

Getting Started

To use Islandora Westvault, you first need a WestVault account. If you're a BC ELN member, contact the Admin Centre to help you get started. If you're a COPPUL member, talk to COPPUL. This module also works if you have any other OwnCloud-based preservation tools.

Once you have an account, you will need to enable the Islandora Westvault module. It also depends on a few other modules:
- Islandora Bagit
- Islandora Bagit Complex
- Islandora Westvault Extras

When everything is enabled, go to the Permissions page to make sure that you have permission to use Islandora Westvault, along with any other roles that you want to grant this permission to. Keep the "administer" privilege very restricted.

Finally, you're ready to get started. Open the Islandora Westvault configuration page under Islandora Utility Modules -> Islandora Westvault Preservation.

In a separate tab, log in to your Westvault account at https://deposit.coppul.ca.

There are a few things you need to set up in Westvault first:
- Click the Files header at the top, and then click WestVault
- This is where you'll see all of the files that have been sent to the LOCKSS network.
- Click the Settings link, and there's just one setting you need to change…
    - Make sure that the "Remove completed deposits" box IS CHECKED.
    - This way, once your files have been sent to WestVault, they will be cleared out of our local server.
    - This will NOT remove your objects from Islandora; it just clears out the temporary storage where files sit waiting to be sent for preservation.
- Now go back to Files, and take a look at the directory list.
- There are two folders here that matter to you: lockss-preserved, and lockss-restored.
    - The lockss-preserved folder is synchronized to a folder in our server's filesystem, where preservation files are stored temporarily on their way to LOCKSS.
    - The lockss-restored folder is where your objects go if you need to restore them from preservation in the case of data loss or what have you.
    - If everything goes well, you won't need to actually interact with either of these.

Let's go back to the module configuration, and figure out what goes into each field…
- OwnCloud Local Path:

- This is the directory on our server that will sync with your WestVault lockss-preserved folder. You don't need to worry about this; I will set it once everything else is done. (So please don't change it after the fact.)
- OwnCloud WebDAV URL:
  - This URL tells the module which folder in your WestVault account to send the files to.
  - By default, it targets the lockss-preserved folder - which is the folder that actually syncs with the LOCKSS network. So please don't change this, ever… if you do, your objects will not end up in the LOCKSS network.
  - (In this test environment, I have it instead targeting a test directory -- this means that any objects I've marked for preservation end up in OwnCloud without getting synced with WestVault - because I don't want to fill up our WestVault space with test objects.)
- Temporary File Location:
  - Don't worry about this; its default setting is fine.
- Maximum filesize in Westvault:
  - This setting was created because of a bug in WestVault's implementation of OwnCloud, which is the software that interfaces between our server and the LOCKSS network.
  - Basically, it cannot send files larger than 2 GB to the LOCKSS network -- at least, not right now.
  - COPPUL are working on fixing this limitation, but for now, we're stuck with it.
  - So this setting tells the module not to bother creating the preservation file, and to log the problem in our Drupal logs. The object will still be flagged for preservation in case the problem is fixed in the future.
- OwnCloud user ID and password:
  - These are the credentials for your WestVault account. The system needs them in order to sync the local directory with the preservation directory in your WestVault account.
  - The password field will appear blank after you fill it in and save it -- this is deliberately done, for security. The bolded text below will tell you whether or not a password is saved.

Once Islandora Westvault is configured, you need to configure one more module to get things working: Islandora Bag-It.

Under Islandora Utility Modules, go to Islandora Bagit.

This module creates Bags, which are basically ZIP files containing your objects' datastreams and metadata. Bags are a standard format for preservation, and are what will be sent to WestVault to preserve your objects.

The main settings you need to concern yourself with are:
- Collection batch type:
  - This determines what kind of bag is generated if you create a Bag for a Collection.

- - You must choose "The Collection object only; children will not be bagged".
    - The other options either zip up the entire collection with all the objects contained therein, or all of the child objects individually with no record of the collection itself.
    - If you choose any other setting, Islandora Westvault will refuse to operate and tell you to fix it.
  - Object plugins:
    - Bag-It requires you select at least one plugin in order to work.
    - The most important one is plugin_object_ds_basic, which ensures that all the object's datastreams are included in the Bag.
    - You can pick other ones as well if you care about them, but make sure ds_basic is selected.
  - Bag Metadata:
    - It's a good idea to fill this out, so that pertinent information about your organization and its relationship to the objects is not lost.
  - Complex Objects:
    - Islandora Westvault expects specific configuration for complex objects (that is, objects with children - books, newspapers, compound objects, newspaper issues).
    - This section determines whether or not the children of these complex objects also get Bagged.
    - Islandora Westvault requires that Books and Newspaper Issues are checked, while Newspapers and Compound Objects are not.
    - This is so that Pages of Book and Newspaper Issue objects are bagged when their parent object is bagged.
    - For Compound Objects and Newspapers, Islandora Westvault already deals with them by marking each of their children for preservation as independent objects. So we don't need to worry about them.
    - So check off Books and Newspaper Issues; leave the others unchecked.
    - If your settings are any different, Islandora Westvault will refuse to run and complain at you to change your settings.
  - Add WestVault Tags:
    - Since WestVault has specific tags in the object metadata that their system requires, you should make sure this is checked.

Once everything is configured

When all your settings are correct, contact the Admin Centre (that is, me). I will do the rest of the setup.

The rest of this segment explains how things work in the back end -- a peek under the hood that you will never have to touch. Feel free to zone out; if all you want to do is tell Islandora to preserve certain objects, you will not need this information.

I will set your OwnCloud directory to one that is unique for your site. Then there are two extra bits that I have to set up in the server back-end: Cron jobs.

A cron job is a task that the server performs automatically at some specified regular interval. For Islandora Westvault, there are two Drush commands that need to be run in order to extract your objects and sync them to WestVault: westvault-bagit, and westvault-sync.

"westvault-bagit" does a number of things:
- Generates a list of all objects that have been marked for preservation
- Eliminates any objects that have already been preserved and sent to WestVault
- Creates a Bag for each object, and then:
  - Analyzes the Bag to make sure it contains everything it's supposed to
  - If the Bag validates all checks, it gets sent to the OwnCloud directory to await syncing, and the object is marked "Preserved"
  - If there's a problem with the Bag that is produced, the Bag is deleted, the problem is recorded in the logs, and the object remains in queue for preservation.

The "westvault-sync" command synchronizes the OwnCloud directory with your remote WestVault "lockss-preserved" directory. Any files therein will be sent off to OwnCloud. There they will sit in the "lockss-preserved" folder until WestVault's own cron job sends them off to the LOCKSS network.

Once they're sent off, the files are removed from the "lockss-preserved" folder. And the next time "westvault-sync" is run, the Bags that have already been sent to WestVault will also be removed from our server.

How to use it

This section covers how to mark objects for preservation.

To preserve an individual object

- Navigate to that object
- Manage the object
- Click the Preservation tab
- Check the "Preserve this object" box

Islandora Westvault creates a new datastream on the object, called "PRESERVATION". It contains metadata about the object's preservation status. Initially, it has a "date preserved" timestamp, indicating when precisely you marked it and created this datastream.

When the automated westvault-bagit command runs, it will find this object, along with any others that have the PRESERVATION datastream, and create Bags for them.

Once the object has been bagged, the PRESERVATION datastream is updated with a "confirmPreserved" tag. This hides it from any subsequent westvault-bagit commands, ensuring that it will not be preserved more than once.

To cancel preserving an object

If you marked an object for preservation by accident, you can undo it… Just go to the PRESERVATION tab, and un-check the box. This will delete the PRESERVATION datastream.

If the object has not already been preserved, doing this will stop it from happening.

If it **has** already been preserved, doing this will erase that information -- so if it's ever marked for preservation again, it will be Bagged and sent to WestVault again. You may want to do this if you have made substantial changes to the object… Just know that it means you will end up with two copies of it in WestVault.

Preserving multiple objects at once

You can also preserve objects that are parents of other objects, like Books, Newspapers, Collections, and Compound Objects. There are a few neat tricks with these.

Collections

To preserve a Collection, manage it and go to the Preservation tab, as usual. But on Collections, you'll see a new option: "Apply preservation policy to…"

You can choose to make a collection behave in various ways:
- Preserve this Collection object and any new children (not the current children)
- Preserve this Collection object, all of its children, and any new children that come in
- Preserve this Collection object and its current children (and not any new children that get added)

That choice gets written to the preservationMethod element, which gets checked any time a new object is added to the collection. If you selected an option that preserves current children, a new PRESERVATION datastream is created for every child object that doesn't already have one. (If it does already have one, it gets skipped.) And if you selected one that applies to new children, any new child of this collection will get a PRESERVATION datastream if it doesn't already have one. This also applies to objects that are shared with or migrated to this collection.

If you un-check the "Preserve this object" box from a collection, the collection's PRESERVATION datastream is deleted -- but all of its children are left untouched. This is because we cannot predict how you'll want to deal with child objects after you un-preserve a collection; you'll need to figure that out on your own.

Newspapers

You can choose to preserve individual Newspaper Issues, or you can preserve whole Newspapers.

To preserve a Newspaper, do the usual: manage, Preservation tab, check the box. A PRESERVATION datastream will be created for the Newspaper object, as well as for all of the Newspaper Issues associated with it.

If your Newspaper object is marked for preservation, any new Newspaper Issues that are added to it will also get a PRESERVATION datastream.

Compound Objects

If you mark a Compound Object (that is, the parent object) for preservation, so will all of its child objects, and any new objects that are later made its children. As with collections and newspapers, un-checking the "Preserve this object" box will delete the Compound Object's PRESERVATION datastream, but will leave the children's PRESERVATION status unchanged.

Books and Newspaper Issues

Preservation of the pages of Books and Newspaper Issues is taken care of by the Islandora Bagit module, rather than Islandora Westvault. Treat them just like any other objects. However, it's wise if possible to make sure they are complete before you tag them to be preserved, since any new pages added after you preserve them will have to be preserved separately.