

Arca Admins Meeting Notes

Date: Tuesday, March 28, 2023, 10 - 11:30 AM (Pacific)

Note-taker: Sasha Gaylie

This meeting is an informal chat about the upcoming migration to get a sense of questions that administrators currently have and to share things to keep in mind as we go forward.

Migration progress update

- Brandon has been having conversations with DGI about our technical requirements to make sure we're set up to move ahead.
- The plan is to get started with initial migration tasks in April and involve admins as much as possible.
- Sites may have heard about the recent announcement of funding from the Ministry, which will cover 2 years of the migration process.

Migration Q&A

Is Islandora 2 able to support Azure SSO?

- Most likely, yes. Because Islandora 2 is much more strictly Drupal-based, we can use a lot of the modules made available by Drupal.
- There is an [Azure SSO module](#) that we can leverage, and there will be plenty of other ways to manage users and logins after migration.

Will Islandora 2 help with Google Scholar indexing?

- Brandon has been discussing Google Scholar with folks at MRU for the last little while.
- The MRU team has been hard at work to improve Scholar indexing, efforts that should be beneficial to all Islandora members wanting to optimize indexing.
- Years ago, a representative from Google presented to Arca some ways to make objects more available to Scholar, but it looks like some things have evolved (e.g., needing direct links to PDF datastreams in the XML sitemap).
- Islandora 2 will likely improve the overall indexing situation.
- Follow-up question in the chat: **Will Islandora 2 help with Google Search indexing?**
 - It should be relatively straightforward.
 - The way Google Search works has changed a lot over the years; there is now the Console for reviewing problems that come up.
 - There were a few sites recently that had issues with getting indexed and it turned out their sitemaps weren't being properly viewed by Google.
 - If you notice specific issues around this, please reach out to Brandon.

What metadata cleanup will be needed for migration? Can we get started now?

- The metadata cleanup process began last year with the [XPATHS document](#).
- Brandon will be sending out information about specific metadata tasks shortly.

- It would be helpful for admins to get started on cleanup within their collections immediately, if there is capacity.
- Some things to look out for:
 - controlled terms are being used for *genre* – check that these are using correct authorities and formats (all lowercase, plural when meant to be, etc.)
 - date formats are correct (numerical only: YYYY-MM-DD or YYYY-MM or YYYY)
 - all *role* terms match the authoritative list
 - ensure your internal subject terms are consistent as they'll be mapped to a vocabulary in the new system

Will current display setups for forms migrate?

- Because metadata and object structure are handled differently in Islandora 2, your specific display will not be the same in the new environment.
- We're not looking at MODS or SOLR fields, everything will be powered through Drupal fields.
- Metadata will be more consistent throughout the whole repository and there will not be unique ingest forms for different content types (all content is just a repository item), meaning that displays will not have to be quite so customized since there will not be unique SOLR fields.
- Specifically regarding custom labels based on content type (e.g., the use of "credits" instead of "contributors" on video objects), it's not clear how much of this will be customizable in Islandora 2.
- Follow-up question in the chat: **How do we recreate role terms?**
 - In Islandora 2, all of your names (e.g, contributors) and role terms are going to be stored in taxonomies.
 - You will not lose role terms but as previously mentioned, custom labels based on particular contexts (e.g., content types) may not be available.
- Follow-up question in the chat: **What are custom datastreams?**
 - Most sites won't have them, so not a great concern.
 - In Islandora Legacy, you have the option to add a file as a custom datastream to the list of ingested and derived datastreams (these custom datastreams don't necessarily display on your object page but are stored).
 - If you do have custom datastreams, we need to talk about how best to migrate and keep them (e.g., if they need to be migrated on the side).

Will older metadata for MODS or DC be migrated or only current one?

- This question is related to versioning because in our current environment, older versions of datastreams are saved.
- Only the current versions of the metadata will be migrated – new Islandora doesn't use MODS or DC so we won't be keeping older copies of that metadata.
- You may keep copies for your own institutional records, but they won't be moved into the repository so you will need to export them independently.

- Brandon is happy to help you figure this process out, whether you're using WestVault or another storage system, so please reach out if this is a concern for your site.

When exporting metadata, what scheme and format will be used?

- Drupal 9/10 uses RDF but metadata should be exportable to various formats.
- Clarification: **For example, in the current version, you can download metadata as an XML file. Is there a line-by-line structure like this? Possibility to map to an existing schema, like Dublin Core?**
- There will be lots of ways to map and export metadata, so it depends on what admins will be using it for and how they would want to do it.
- There will be a question shortly about OAI-PMH, which is one way to render the metadata in any format you like.
- If you reach out to Brandon with your specific use case, he can provide a more precise answer.

Will relationships migrate as currently established?

- Yes! The idea of migration is to move the objects and their relationships so you will not lose that structure.

Do we have a general timeline for the migration process?

- A rough project charter has been drafted with timelines sketched out, but it is extremely loose and very likely to change.
- Right now, the Admin Centre is working on finalizing technical requirements and ensuring that we can set up all the virtual machines as we need them.
- We hope to start thorough metadata evaluation in April and our goal is to complete this task by June. At that time, we'll also be determining the specifics of our site configurations. We will then work together on fixing up metadata from May to the end of July.
- We can begin our first test migrations as early as mid-August or September. We will be migrating sites individually, starting with small ones to ensure everything goes smoothly before moving on to larger and more complex sites.
- Hoping to have everyone moved by Fall 2024, which would be a little under two years.
- Again, this tentative timeline depends heavily on figuring everything out on the technical side.

Will Views migrate?

- Views cannot be migrated.
- Islandora Legacy stores objects in Fedora, Drupal can see them using SOLR fields, and Islandora SOLR Views can read those SOLR fields as data and display them.
- In Islandora 2, everything is in straight Drupal, meaning that Views are much more powerful and we can do a lot more with them.
- We will be moving all objects but sites will need to recreate custom displays in their own environments.

- The Admin Centre recommends planning this out now to envision how you would like to display things.
- Follow-up question: **We're using a View URL in our research department; can we keep that URL?**
 - Not a problem – you can keep your custom path to your new view.
 - Paths to individual objects will likely change, though this will need to be explored a little more during migration.

On our repository homepage, will logos be able to link back to our institution's main website instead of the repository?

- This has not yet been investigated but likely doable; flagged for follow-up.

Is Islandora 2 OpenAire compliant?

- It should be!
- Islandora 2 uses the Drupal REST OAI-PMH module, which is highly customizable.
- The module can be used to look at RDF or straight views that you create of your data, but we will need to look more deeply into the specifics of its configuration.

Will there be an opportunity to export metadata for batch remediation?

- Yes, you can do that now.
- Islandora Datastreams I/O module is used for batch exports (see [Arca Minutes video](#) on this topic).

Do Compound Objects exist in Islandora 2?

- Yes, they exist and it shouldn't be a problem to migrate existing compound objects, assuming nothing strange has happened with the relationships.

Will institutional repositories have new URLs after migration?

- Websites will only have new URLs if you want them to; will try our best to keep repository URLs the same.

Will existing forms stay the same?

- No, ingest forms exist for writing metadata to XML for the repository.
- In Islandora 2, you will be filling out a form consistent across all of your objects.

Will the simple workflow function be the same?

- No, that is a good question; flagged for follow-up.
- We can look up what workflows we like for drafting and approving objects.

Are we going to use any kind of project management platforms or software?

- The Arca Admin Centre will definitely be using some kind of project management software but in terms of communicating within our community, it would be good to pull together some Slack channels to manage these things.

- We will also be pulling together task groups for different work related to migration, so having a dedicated communication channel would be helpful.

Is it possible to export data as spreadsheets?

- When Brandon talked to DGI about metadata remediation, they did mention something about CSVs; flagged for follow-up.

Will submission processes stay the same? We have self-submission set up.

- There will definitely be ways to do that but it will be different in Islandora 2.
- Flagged for follow-up.

Will object URLs stay the same?

- They will probably be different.
- There may be ways to preserve the old object URLs in metadata if needed, but requires investigation.

Will the Internet Archive book viewer go away?

- Yes, it will be replaced by Mirador.
- The way that Islandora Legacy handles paged content is that it takes what is usually a source PDF, extracts a large TIFF for each page, saves each of those pages as an individual object, and then stitches the page objects together in the BookReader to present to the viewer.
- The larger question is whether sites will want to continue having individual pages for paged objects or if it's preferable just to use PDFs, which will have implications for display.
- This is something sites can start thinking about now.

Will mass imports of metadata be easier in Islandora 2?

- Clarification: **Right now, when revising MODS, we use the Islandora Datastreams I/O module to import MODS files and if we try to import more than 100 or so files at a time, the system will time out. Will there be greater capacity in Islandora 2 for that kind of operation?**
- Islandora 2 does not use MODS XML, so sites won't be importing data that way anymore.
- Because every site will have its own container or VM, it may be possible to give administrators more backend access to get around the constraints of working in the browser.

You mentioned migration will be one institution at a time. Do you know the approximate schedule?

- When we actually have our images built (the image is the default install that can be deployed to a new site) and once we have our basic metadata fixes and mapping done, we can start looking at scheduling.

- Scheduling will be partly based on the size of the site. If a site has 3 TB of storage, we may leave it for later compared to a site that has 50 GB or less.
- If you are not a very busy site with frequent ingesting going on, we may start with you.
- We may also give larger sites access to start ingesting directly into their new environment while they wait for their existing content to be migrated.

Final reminders

- Admins are invited to review the document [In-House Migration Tasks](#), which contains some tasks that you can get started on immediately.
- DGI will also want sample data that can be used as a baseline to test migrations so if you would like to volunteer any of your data, this is something we will put a call out for.
- Feel free to reach out to the Arca Admin Centre at any time with questions or concerns.