

Arca Admin Meeting Minutes - Sept 24, 2020

Chair: Brandon Weigel

Minutes: David C. Waddell (VPL)

1. Repeated Downtime - Explanation

DDoS (Distributed Denial of Service) attacks from botnets occurred in the earlier part of the year. Arca undertook a big project to put CloudFlare up across Arca. Cloudflare is a service used by 10% of the internet to act as a firewall against such attacks.

In July and August, several issues occurred at once: a corrupted disk at SFU + resource indexes used instead of Solr + poorly constructed views being deployed in large numbers = serious memory problems for Arca. All of these incidents came at once and created a perfect storm.

In response, infrastructure expanded onto a 4th server, where Solr now operates all on its own. Previously it was sharing services.

2. Solr core continues to be pared down to remove unique Solr fields.

A lot of uniquely named datastreams exist in some collections, possibly by accident. We need to reduce redundancies based on spelling, purpose, etc. Of course, some legitimate use cases exist as well. Example raised: Emily Carr had some use cases that Brandon was already aware of.

Action: Please let Brandon know about custom datastreams that your institution might need. Anything like licenses, other files being attached, or any other use cases should be brought to Brandon's attention before making a new kind of datastream.

If any Arca members are already using custom datastreams on their site, send Brandon a message. The Arca Admins can get together to help solve on a case by case basis.

3. The site is slower than we would like. Discovery Garden is looking into why Solr is so slow.

DiscoveryGraden is still working on the Solr indexes. They have gone through the sites looking for problematic configurations.

Action: We request members not change anything on any Islandora configuration pages (the page that deals with namespaces and other core configurations). Some checkboxes there can cause large problems.

4. A lot of bad bots are still getting through.

It is a small percentage of total attacks, but it has required Brandon to make more specific rules and at times to put the site under 'Attack Mode'. These restrictions impacted harvesters (like Library and Archives Canada and EBSCO) from accessing the site. Brandon has now placed these organizations on a whitelist, which allows them access during 'Under Attack Mode'. If this issue comes up again with other harvesters you work with, please let the Admin Centre know.

5. Theming and Accessibility

During a past Advisory Committee meeting, a decision was made on theming and accessibility plans.

Islandora 7 is approaching end of life. Most development activity has moved to Islandora 8, which includes a focus on accessibility.

Brandon continues to advocate for Arca's multi-tenancy use case among the Islandora development community.

This means that in the next two years, it will be both feasible and necessary to move over to Islandora 8. This makes expending resources on a re-theming and accessibility project unwise on 7 if it will become obsolete within 1 year.

Thus, all the budget for the theming and accessibility project is being held over for later years and Arca will do that work as part of the migration when the time comes.

6. More on Islandora 8

Islandora 7 support is scaling back. No new features are coming for 7, only bug fixes; all new development is happening in Islandora 8. So, the major issue holding back Arca's adoption of Islandora 8 is a clear path to multi-tenant repository support (Arca's parent and child approach). Within a year, Arca will prep for a migration.

Islandora 8 is very different from Islandora 7. No more MODS; everything will be in Drupal fields using an RDF schema. Many webinars and other resources exist on migration from 7 to 8. We will need to look into mapping our existing metadata into Islandora 8 fields. This will be a big project. The cleaner your metadata is now, the easier it will be when the migration comes. So cleaning it up while still in 7 makes sense.

Responding to a question: Islandora 8 avoids XML and everything is stored in Drupal, so things are much more straightforward for ingest forms.

Question: Are there implications for linked data?

Answer: We aren't in a position to consider this yet, but it will become possible to consider because of RDF.

7. This Year's Training

Messages have gone out. Orientation was completed and two basic trainings are coming up. Sign-ups are still open. Group seemed to indicate this was sufficient, along with the videos.

The Arca Sandbox remains available. Just ask Brandon for a login.

With no further discussion, meeting adjourned.