Islandora and Fedora 4; The Atonement.

Nick Ruest, York University, ruestn@yorku.ca
Daniel Lamb, discoverygarden inc., daniel@discoverygarden.ca
Melissa Anez, Islandora Foundation, manez@islandora.ca

Session Type (select one)

Presentation

Abstract

In the context of repository platforms, Islandora has a fair bit of age, and with that a fair bit of cruft. In the early winter of 2014/2015 the Islandora community began working on a project plan to outline what would be needed in a version of Islandora that would work with of Fedora 4, and what resources would be needed to build it. This presentation will provide an overview of the project, as well as an in-depth technical overview and demonstration of the new functionality.

Conference Themes

- Supporting Open Scholarship, Open Science, and Cultural Heritage
- Managing Research (and Open) Data
- Integrating with External Systems
- Managing Rights
- Developing and Training Staff
- Building the Perfect Repository

Keywords

Islandora, Fedora 4, Apache Camel, community, resources, engagement

Audience

The audience for this proposed talk is repository managers, developers, and librarians/archivists.

Background

This presentation fits in splendidly with the "Looking back, moving forward" theme of OR205. It looks over the Islandora project's history, how it has evolved, how the Islandora 7.x/Fedora 4.x project is informed by the project's history, and how it will adjust and move forward based on that history.

Presentation content

Community lead and built

The Islandora 7.x/Fedora 4.x integration project is building on a multi-year history of community driven development. In particular, the methods learned from the previous two releases of Islandora's software, which were managed and developed entirely by volunteers from the community, will provide a model for our upgration to Fedora 4.x.

One major deviation from our community-driven release model is the addition of paid developer time. Because of the sheer scope of the Fedora upgration, and the desire to re-architect some of the key components of Islandora for a slicker, easier-to-adopt product for the future, the project needed the services of at least one full-time developer who could take the lead on the technical aspects of the project and maintain a continuity of design, to which the community can contribute via use cases, discussion, and volunteer developers. Funding for this resource came from multiple sources in the Islandora community, with a number of supporting institutions stepping up to become Partners in the Islandora Foundation by contributing directly to the Fedora 4 integration.

Working with this one paid resource are a volunteer director, a project team with representatives from more than a dozen institutions, and an entire community of Islandora users and developers who are contributing their use cases and opinions to ensure that this newest version of Islandora serves the community that has come together to build it.

Technical overview

Islandora version 7.x-2.x is middleware built using Apache Camel to orchestrate distributed data processing and to provide web services required by institutions who would like to use Drupal as a front end to a Fedora 4.x repository.

This goal presents a unique set of challenges. Drupal is far more than a simple display layer; it is a full blown content management system designed to be built on top of a traditional relational database such as MySQL or Postgres, not a JCR repository. Additionally, there is a large amount of data processing and manipulation that must be performed for presentation and discovery, requiring the integration of additional third-party software such as Tesseract, ImageMagick, FFmpeg, and many others. If these processes are performed on the same servers that host Fedora and Drupal, is can be detrimental to the performance of the overall system, resulting in a unusable site during periods of content migration or manipulation.

To mitigate these issues, the overall design goals of the 7.x-2.x version of Islandora are:

- A properly modularized installation procedure so that Islandora can be consistently installed and configured in distributed environments.
- Asynchronous communication between Fedora and Drupal. This will be achieved through the use of persistent queues, which will also allow the stack to be easily distributed across multiple computers.
- Fedora treated as the source of the truly important data, only containing preservation objects and descriptive metadata. Metadata can exist either in Fedora's native RDF attached to the resource itself, or as standardized formats such as MODS, MADS, PBCORE, etc... that exist as resources in their own right.
- Data from Fedora transformed and indexed into the other major systems of the stack, most notably Drupal and Solr.
- Drupal content is represented as Nodes and Fields, allowing the content management system to utilize the relational database it is expecting instead of shimming in a completely different type of datastore. This will open up the entire Drupal module ecosystem to Islandora.
- Drupal's Services module will be used to expose RESTful services to middleware layer.

Conclusion

The Islandora 7.x/Fedora 4.x project may seem ambitious at first glance, but looking back on how our community has grown over the years, we are able to move forward with this project with confidence. The project will sit the community up nicely for the future with its rearchitected structure, and full integration into Drupal will make the project much more sustainable.

References

N/A