CLAWbarium

Islandora Camp HRM - July 20, 2018
Paul Pound and Donald Moses
University of Prince Edward Island
What We’ll Cover

1. Migration - Data Modelling
2. Migration - Data Extract/Enrich
3. Migration - Node creation
4. Search Configuration
5. Enhancements
6. Data Visualization
7. Challenges
8. Next Steps
Migration - Data Modelling

- **Drupal 6/F3 > Drupal 8/F4**
- Reviewed existing content model and map that to a new D8 content type
- **Islandora Herbarium Object**
  - Fields
    - Text fields map to DarwinCore or DC
    - Geo field
    - Tags (not required)
  - Media
    - Preservation Master
    - and derivatives
# Migration - Content Type

<table>
<thead>
<tr>
<th>Member of</th>
<th>field_member_of</th>
<th>Entity reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Municipality</td>
<td>field_dwc_municipality</td>
<td>Text (plain)</td>
</tr>
<tr>
<td>Municipality URI</td>
<td>field_dwciri_municipality</td>
<td>Text (plain, long)</td>
</tr>
<tr>
<td>Order</td>
<td>field_dwc_order</td>
<td>Text (plain, long)</td>
</tr>
<tr>
<td>Phylum</td>
<td>field_dwc_phylum</td>
<td>Text (plain, long)</td>
</tr>
<tr>
<td>Recorded By</td>
<td>field_dwc_recordedby</td>
<td>Text (plain, long)</td>
</tr>
<tr>
<td>Rights</td>
<td>field_dc_rights</td>
<td>Text (plain, long)</td>
</tr>
<tr>
<td>Scientific Name</td>
<td>field_dwc_scientificname</td>
<td>Text (plain)</td>
</tr>
<tr>
<td>Scientific Name URI</td>
<td>field_dwciri_scientificname</td>
<td>Text (plain, long)</td>
</tr>
</tbody>
</table>
Migration - RDF Mapping

Snippet from: [rdf.mapping.node.islandora_herbarium_object.yml](#)
Migration - Data Extract/Enrich

- Drush script to facilitate migration
  - Give script list of PIDs
  - Script extracts data from Original objects DarwinCore XML Datastream
  - Utilizes scientificName and municipality fields to perform external lookups
  - Used lookup results to enrich and correct ‘tree of life’ data and geographic data (eg. longitude and latitude)
Migration - New Nodes

- Drush script populates key node fields with extracted/enriched data
- Script imports archival image and populates the “Preservation Master” Media
- CLAW processes Preservation Master and generates a Service File and a Thumbnail
# Data Entry Form

## EVENT DATE

| 1967 - 07 - 17 | 05 : 10 : 10 a.m. |

## Rights

This material is provided for research, education, and private use only.

## Dublin Core Rights

## Municipality URI

http://geogratis.gc.ca/services/geoname/en/geonames/BACME

You must provide a URL to NRCAN datasource (formatted similar to http://geogratis.gc.ca/services/geoname/en/geonames/BACME in the field below. If this field is populated we will use data retrieved for it’s endpoint to populate the Municipality URI.

## Municipality

Tenmile House

Municipality text field (formatted similar to Mount Steward). If you provide a Municipality here and not a URL use the name here to create the URL and look up the Municipality/Location data.

## Title *

Symphytum asperum – Tenmile House

## Scientific Name URI

http://api.gbif.org/v1/species?name=Symphytum asperum
OpenSeaDragon Viewer

Fields with Field Formatters

Geofield - Geomap Field Formatter
<table>
<thead>
<tr>
<th>NAME</th>
<th>MEDIA TYPE</th>
<th>MIME TYPE</th>
<th>TAGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symphytum asperum–tn.jpg</td>
<td>Image</td>
<td>image/jpeg</td>
<td>Image, Thumbnail</td>
</tr>
<tr>
<td>Symphytum asperum</td>
<td>File</td>
<td>image/jp2</td>
<td>Image</td>
</tr>
<tr>
<td>Symphytum asperum–web.jpg</td>
<td>Image</td>
<td>image/jpeg</td>
<td>Image, Service File</td>
</tr>
<tr>
<td>Symphytum asperum–preservation–master</td>
<td>File</td>
<td>image/tiff</td>
<td>Preservation Master</td>
</tr>
</tbody>
</table>
Configuring Search - Solr

- Utilize existing Drupal 8 Solr Module
  - Admin > Config > Search and metadata > Search API
  - Added fields to index
    - > Default Solr content index
- Configure Displays
  - Admin > Structure > Content types > Islandora Herbarium Object > Manage display
  - Configure Facet blocks
# Solr Fields

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Date</td>
<td>field_dwc_eventdate</td>
<td>date</td>
</tr>
<tr>
<td>Family</td>
<td>field_dwc_family</td>
<td>string</td>
</tr>
<tr>
<td>Genus</td>
<td>field_dwc_genus</td>
<td>string</td>
</tr>
<tr>
<td>Kingdom</td>
<td>field_dwc_kingdom</td>
<td>string</td>
</tr>
<tr>
<td>Latitude</td>
<td>field_dwc_decimallatitude</td>
<td>string</td>
</tr>
<tr>
<td>Longitude</td>
<td>field_dwc_decimallongitude</td>
<td>string</td>
</tr>
</tbody>
</table>
## Solr Search Display Admin

EVA provided by Islandora CLAW

<table>
<thead>
<tr>
<th>FIELD</th>
<th>LABEL</th>
<th>FORMAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVA: Media EVAs – Thumbnail</td>
<td>Inline</td>
<td>Plain text</td>
</tr>
<tr>
<td>Type</td>
<td>Inline</td>
<td>Plain text</td>
</tr>
<tr>
<td>Scientific Name</td>
<td>Inline</td>
<td>Solr Facet Field Formatter</td>
</tr>
<tr>
<td>Vernacular Name</td>
<td>Above</td>
<td>Plain text</td>
</tr>
<tr>
<td>State Province</td>
<td>Inline</td>
<td>Plain text</td>
</tr>
<tr>
<td>Municipality</td>
<td>Inline</td>
<td>Solr Facet Field Formatter</td>
</tr>
<tr>
<td>Recorded By</td>
<td>Inline</td>
<td>Solr Facet Field Formatter</td>
</tr>
</tbody>
</table>
Solr Search Display

**Sambucus canadensis - Farmington**

*Type*: occurrence  

*Scientific Name*: Sambucus canadensis  

*State Province*: Prince Edward Island  

*Municipality*: Farmington  

*Recorded By*: Daryl Guignion  

*Event Date*: 1968-07-29
Enhancements

Modules

- **Solr Facet Field Formatter**
  - create links to solr facets
- **Herbarium Object Module**
  - Field formatters
    - Scientific name and municipality in graph view
  - Data lookup and enhancement
  - Data Visualization
Data Enhancements - lookups

• Populate Geo data using a [NRCAN API](#) – [Sample dataset](#)

• Populate Tree of Life data from [GBIF](#) – [Sample dataset](#)
Data Visualization - D3
Data Visualization - Geo

Geofield Google Map View
Challenges Encountered

- **CLAW still changing**
  - Content Modelling pull request while developing
- **Some bugs**
  - For example placing blocks didn’t work so had to use context for Solr facets block
- **Learning curve**
  - Drupal 8, Fedora 4, LDP etc.
Next Steps

• Theming
• Move beyond proof of concept
  – Decisions on domain name
  – Code cleanup (some assumptions should be moved to config)
  – Finalize content type
  – Final data ingest
• Interesting local developments UNB Herbarium Drupal 8 profile
Thank you
Paul Pound & Donald Moses
University of Prince Edward Island