

SFU Library's Migration to Islandora

Mark Jordan

Simon Fraser University Library

Innovative Uses of Islandora Project Briefing
CNI Spring Meeting, 2015
Seattle, April 14, 2015

What's so innovative?

- Migration process
- New Islandora modules and tools developed to meet SFU's needs

Outline

- Rationale for migrating
- Migration use cases
- Metadata mapping and conversion
- Tool development

Rationale for migrating

CONTENTdm has served SFU Library well since 2007, but:

- We need a more flexible repository platform
- We tend to favor open source solutions where they are at least as good as the proprietary equivalents
- We want to standardize our repository platforms

Migration use cases

- Scalability (1.3 million objects)
- Non-disruptive replacement for the CONTENTdm API
- Preserve old URLs, auto-redirect to new Islandora equivalents
- Content types supported by CONTENTdm but not supported in Islandora

Metadata migration

- 130 collections in CONTENTdm (14 content types)
- Map CONTENTdm fields to MODS
- Some collections can't map to MODS
- Clean up non-standard metadata
 - incorrect data (in the wrong metadata element)
 - fields with mixed information types
 - date formatting
 - specialized / made up fields
- Develop library-wide metadata profiles

Before (CONTENTdm) ...

| CONTENTdm field | Content type | DCTERMS mapping |
|---|--------------|-----------------|
| BC Multicultural Photograph Collection at the Vancouver Public Library (/vpl) | Photo | |
| Title | | Title |
| Creator | | Creator |
| Subject | | Subject |
| Description | | Description |
| Date | | Date |
| Language | | Language |
| Identifier | | Identifier |
| Rights | | Rights |
| Format | | Format |
| Type | | Type |
| Full resolution | | |
| OCLC number | | |
| Date created | | |
| Date modified | | |
| CONTENTdm number | | |
| CONTENTdm file name | | |



... and after (mapped into MODS)

MODS mapping

```
<titleInfo><title>%value%</title></titleInfo>  
<name><namePart>%value%</namePart><role><roleTerm type="code" authority="marcrelator">cre</roleTerm><roleTerm type="text">creator</roleTerm><role></name>  
<subject><topic>%value%</topic></subject>  
<abstract>%value%</abstract>  
<originInfo><dateIssued encoding="w3cdtf" keyDate="yes">%value%</dateIssued></originInfo>  
<language><languageTerm type="text">English</languageTerm><languageTerm type="code" authority="iso639-2">eng</languageTerm></language>  
<identifier type="local" invalid="yes">%value%</identifier>  
<accessCondition type="use and reproduction"><%value%</accessCondition>  
  
<typeOfResource><still image>%value%</still image></typeOfResource>
```


Tool development

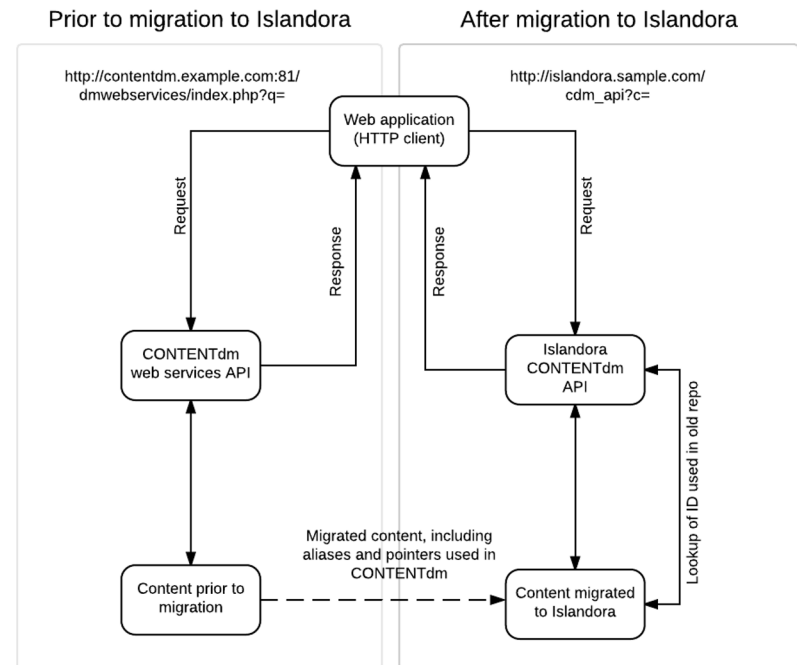
- Move to Islandora Kit (MIK)
- CONTENTdm API Emulator
- Collection Migrator
- URL Rerouter
- Islandora Feeds
- Islandora Newspaper Batch
- Misc utility modules

MIK (Move to Islandora Kit)

- Command-line tool for generating Islandora ingest packages
- Initial sources will be CONTENTdm collections and CSV/local files
- Extensible and customizable
- Open source (GPL 3)

CONTENTdm API Emulator

- For sites that use CONTENTdm Integration Drupal modules for custom and specialized front-ends (e.g., Komagata Maru: <http://komagatamarujourney.ca/>)
- API emulator translates CONTENTdm API read requests to corresponding Islandora API read requests.
- Sites continue to function as is without the presence of CONTENTdm instance.
- This allows us to sequence the work: first the migration, then update specialized front-ends.



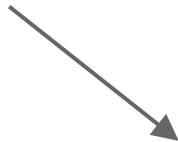
Islandora Feeds

- Uses the Drupal Feeds contrib module to import Islandora objects
- Imports CSV, TSV, etc. plus XML files
- Will be used for “database-like” collections

Title, Field 1, Field 2, Field 3

"A sample Islandora object", "The first field's value", "The second field's value", "The third field's value"

"Second title", "Second item's field 1 value", "Second item's field 2 value", "Second item's field 3 value"



| | |
|--------------|---------------------------|
| Title | A sample Islandora object |
| First field | The first field's value |
| Second field | The second field's value |
| Third field | The third field's value |

Islandora Newspaper Batch

- Developed by Discovery Garden Inc, sponsored by SFU Library
- Open source (GPL 3)
- Allows large ingests of newspaper content via web interface or command line
- Joins Islandora Batch and Islandora Book Batch modules for bulk ingestion of objects



SIMON FRASER UNIVERSITY
LIBRARY

Thank you!