UNIVERSITY of PRINCE EDWARD ISLAND

people • excellence • impact

Islandora 8 RDM Lifecycle

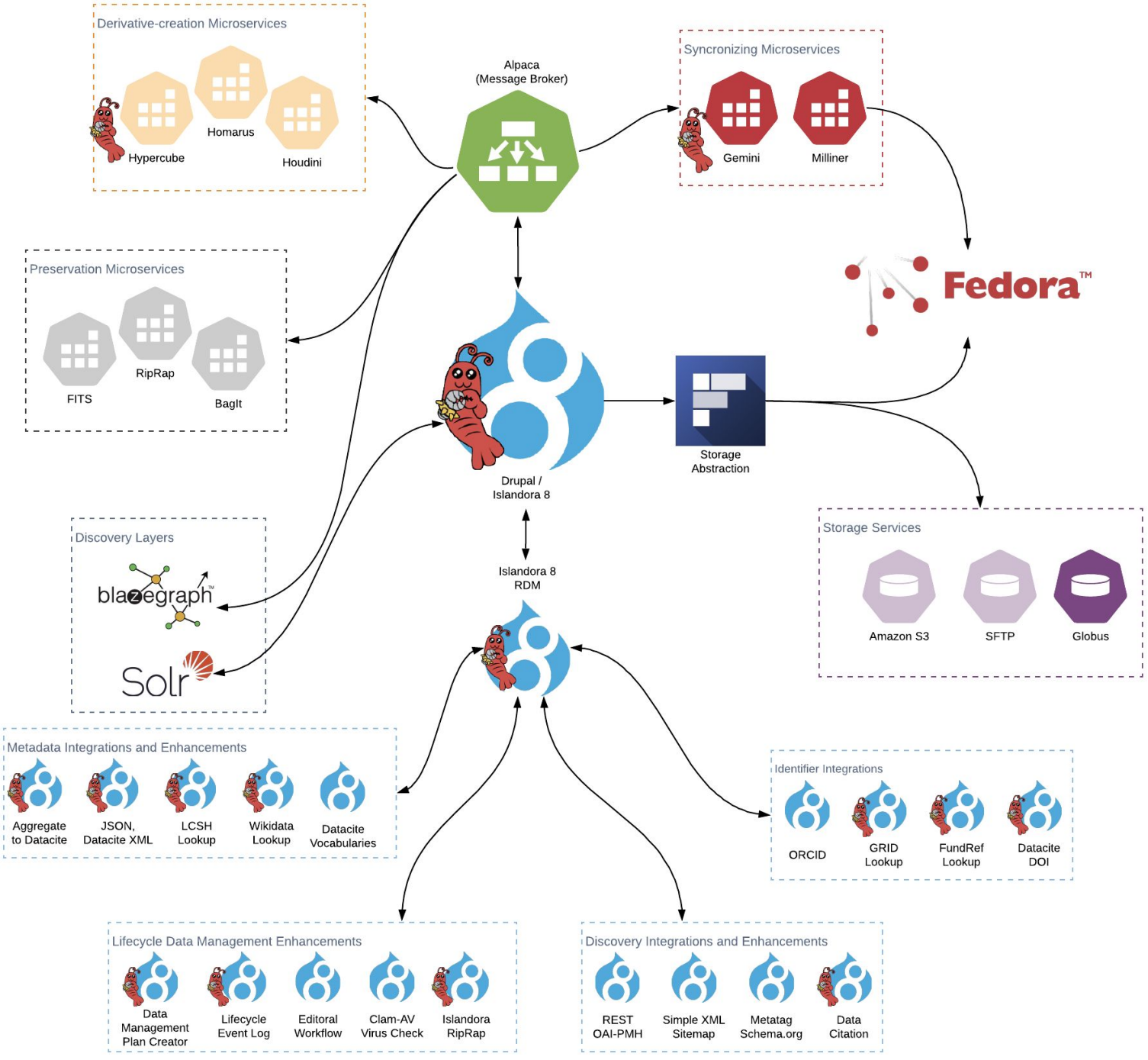
Feb. 4, 2020

Donald Moses, Alexander O'Neill, and
Rosie LeFaive, Alan Stanley, Noella MacIntyre
RDM-085
University of Prince Edward Island

Islandora 8 RDM Platform

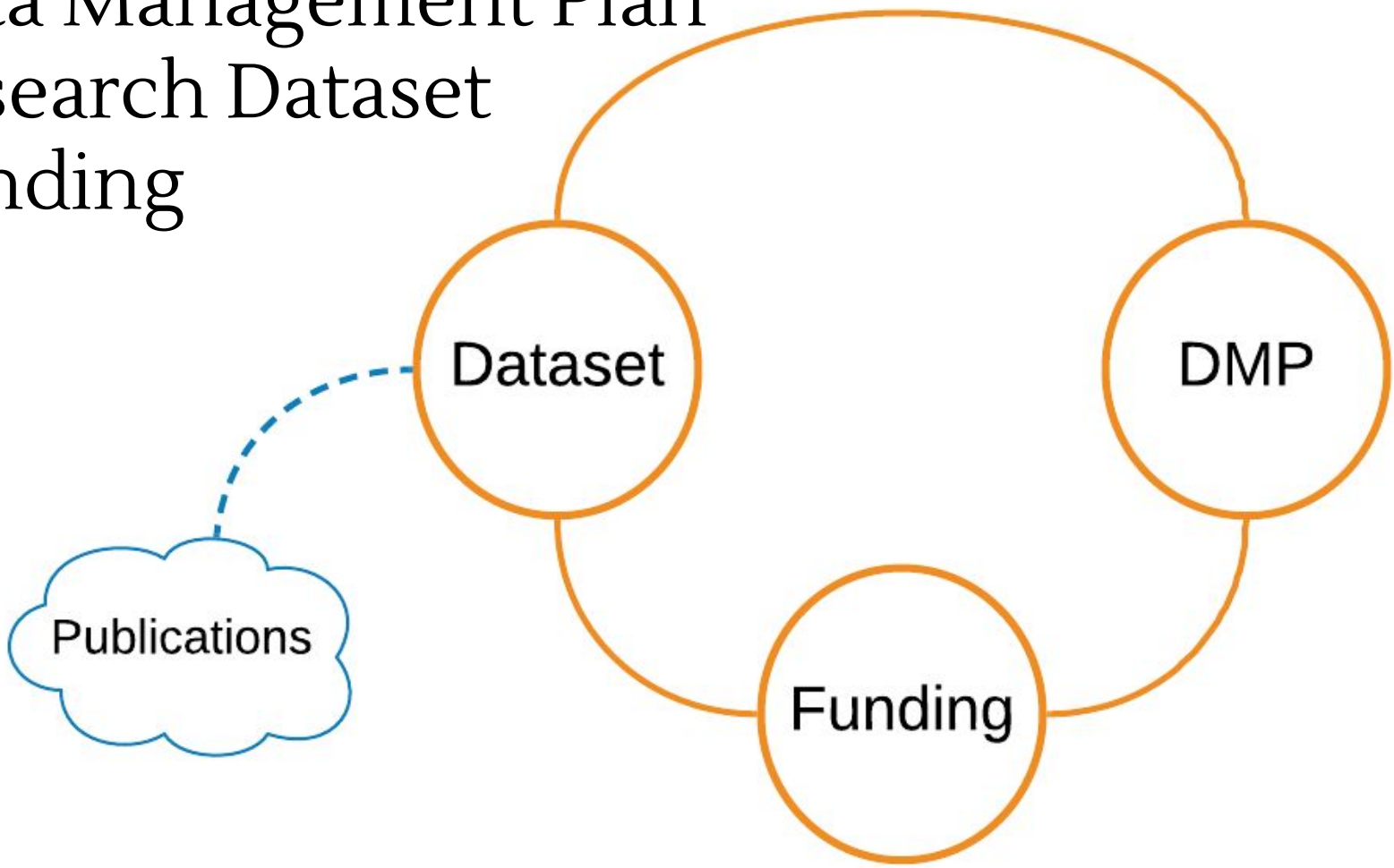
Develop features in Islandora 8 and integrations with other services that provides a comprehensive research data management platform that is informed by the FAIR principles.

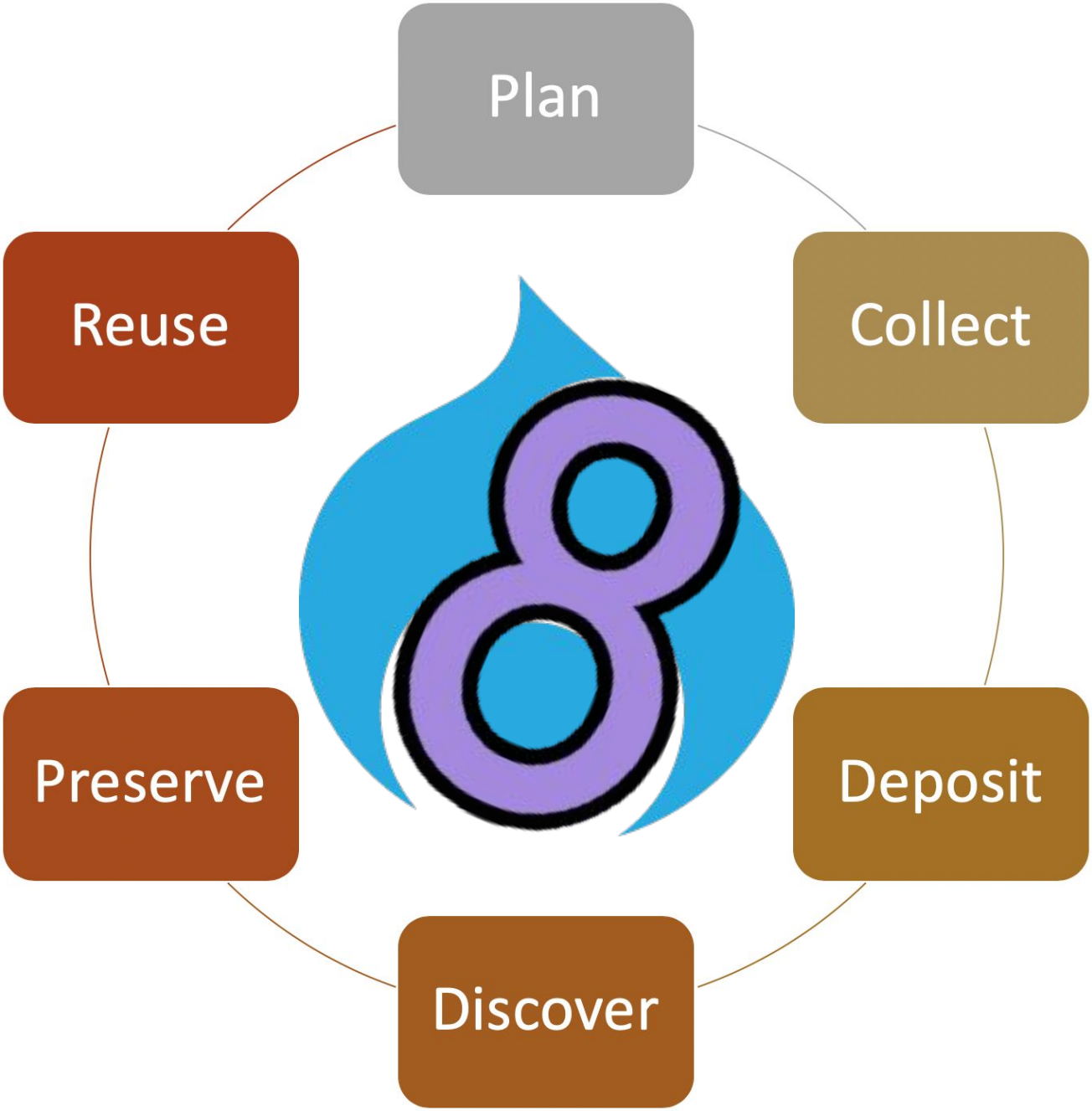
How we are FAIR.


Derivative-creation Microservices
Hypercube
Homarus
Houdini
Alpaca
(Message Broker)
Syncronizing Microservices
Gemini
Milliner
Fedora™
Preservation Microservices
FITS
RipRap
BagIt
Drupal /
Islandora 8
Storage
Abstraction
Storage Services
Amazon S3
SFTP
Globus
Discovery Layers
blazegraph™
Solr
Islandora 8
RDM
Metadata Integrations and Enhancements
Aggregate
to Datacite
JSON,
Datacite XML
LCSH
Lookup
Wikidata
Lookup
Datacite
Vocabularies
Identifier Integrations
ORCID
GRID
Lookup
FundRef
Lookup
Datacite
DOI
Lifecycle Data Management Enhancements
Data
Management
Plan Creator
Lifecycle
Event Log
Editoral
Workflow
Clam-AV
Virus Check
Islandora
RipRap
Discovery Integrations and Enhancements
REST
OAI-PMH
Simple XML
Sitemap
Metatag
Schema.org
Data
Citation

Content Types

- Data Management Plan
- Research Dataset
- Funding




Plan
Collect
Deposit
Discover
Preserve
Reuse

PLAN

Data Management Plan

- DMPRoadmap as a model
- Drupal form
- Users can share unpublished DMPs with other users.
- Published DMPs become public
- DMPs can be exported

PLAN

Data Management Plan

Project Details * | Instructions | Write Plan | Add Collaborators

Project Title *

Collaborative Research: Quantifying the ecological role of Gulf of Maine deep s

Project description

Overview: Deep-sea and cold-water corals are known to be ecosystem engineers, creating foundation structures and habitat for a myriad of associated species. As autogenic engineers, these structure-forming species increase the complexity of the environments in which they grow and associated organisms rely on the physical growth of these corals to change the environment by providing shelter and altering resources. These corals often have a large branching morphology, which creates space for other organisms to seek refuge (for safety from predators, feeding, procreation and protection of young) and thus results in increased

Funding Organization

Funder

National Science Foundation

Crossref Funder DOI

http://dx.doi.org/10.13039/100000001

Grant Number (optional)

1111

PRINCIPAL INVESTIGATOR *

PRINCIPAL INVESTIGATOR

• Person 2 Collapse

Personal Name *

Waller, Rhian

Show row weights

PERSONAL IDENTIFIER

Personal Identifier Remove

PLAN

Data Management Plan

Project Details * | Instructions | Write Plan | Add Collaborators

DATA COLLECTION

What data will you collect or create?

B I | Format | Source

The project will produce several observational and experimental datasets, described in the list below. In addition to the datasets described below, educational resources produced by the project, including data and images, will be made available for public use on the COSEE.net website. Observational data will be collected on a Gulf of Maine research cruise planned to take place August 2020. Observational Datasets:

1. CTD: CTD data collected using a SeaBird SBE CTD package mounted on the ROV system; processing to be done using SeaBird's SeaSave software; data will include standard environmental measurements (such as pressure, temperature, salinity, fluorescence). File types: Raw (.con, .hdr, .hex, .bl) and processed and .cnv, .asc, .btl) ASCII files. Repository: BCO-DMO

2. Event log: Cruise scientific sampling event log; will include event numbers, start/end dates, times & locations of instrument deployments. File types: Excel file converted to .csv; scanned PDFs. Repository: BCO-DMO and Rolling Deck to Repository (R2R).

3. Cruise underway data: Routine underway data collected along the ship's track (including meteorological data, sea surface temperature, salinity, fluorescence, ADCP). Will be collected by the shipboard instrumentation. File types: .csv ASCII files. Repository: BCO-DMO and R2R.

4. Split Beam Sonar Data:

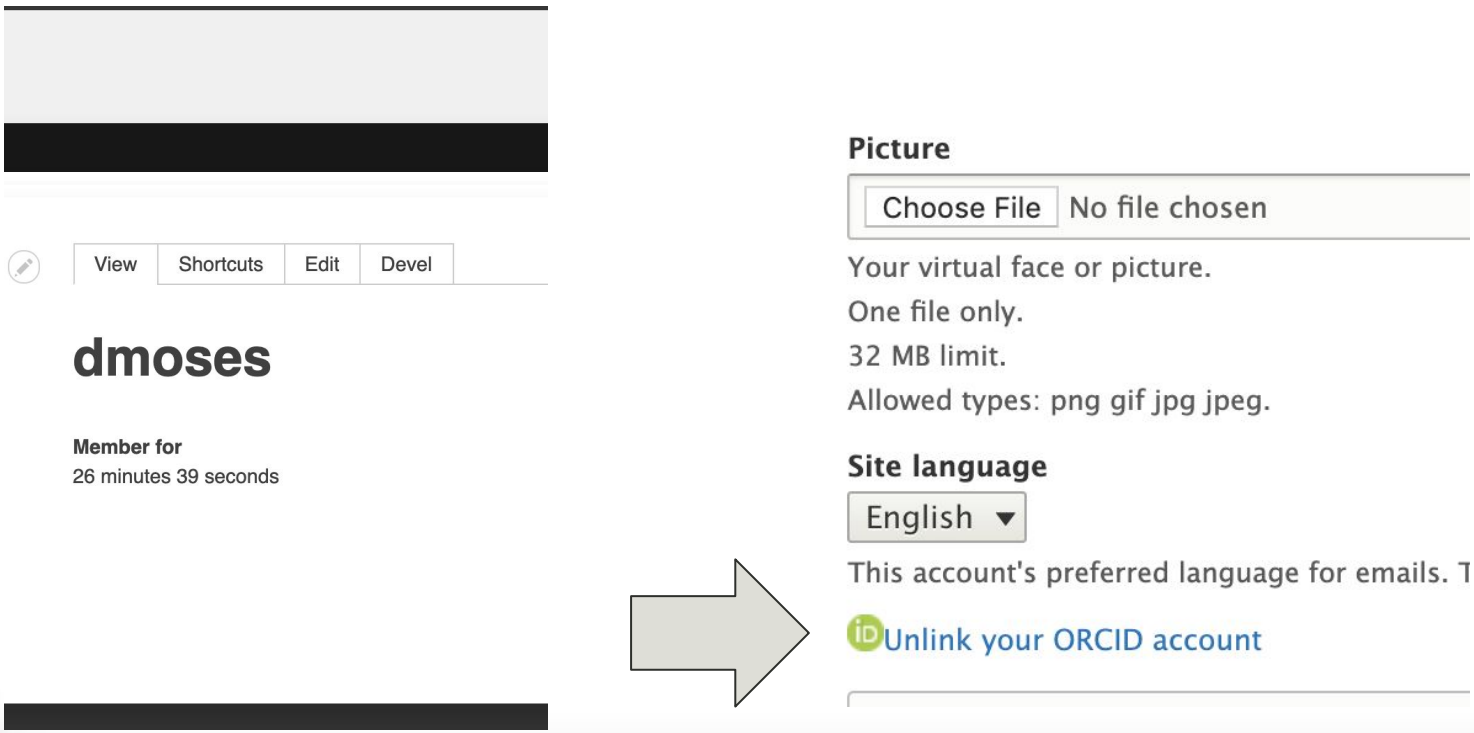
5. Deep sea coral samples: Corals will be sampled using the ROV during the cruise. Species identified, dive numbers, locations, depths, dates, and times will be recorded by hand on log sheets. Information from log will be transferred into an Excel spreadsheet. Images of each coral collected will be taken both in-situ by the ROV, and on the ship using a digital camera. File types: PDF files of scanned log sheets; Excel files of sampling logs; images (.jpg files). Repository: BCO-DMO.

6. ROV Images: Still and video images will be taken during this cruise from the ROV, including 3D models generated from steropair data. File types: images (.jpg files), video (.mov files). Repository: BCO-DMO.

DEPOSIT

Researchers

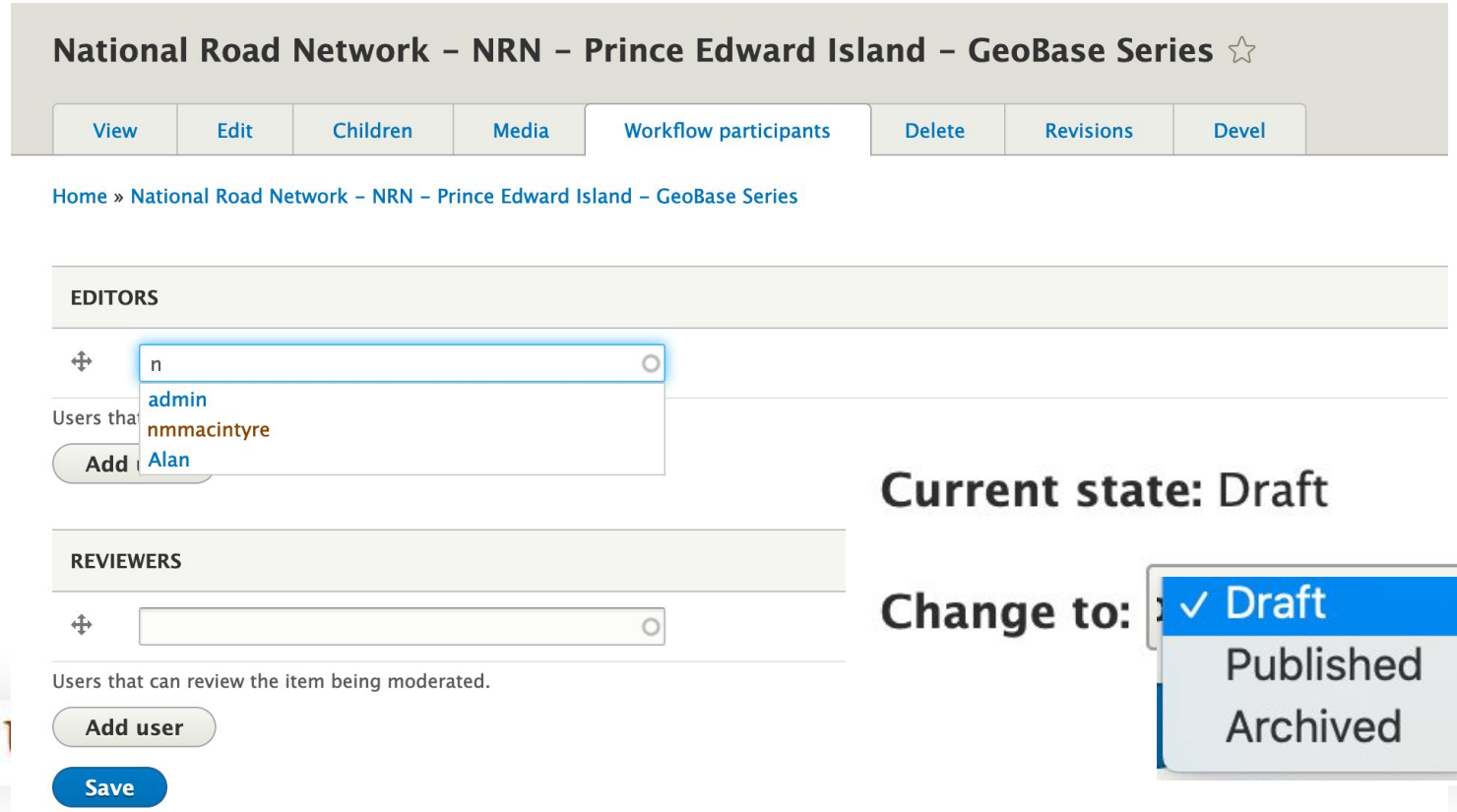
- Can log in using their ORCID



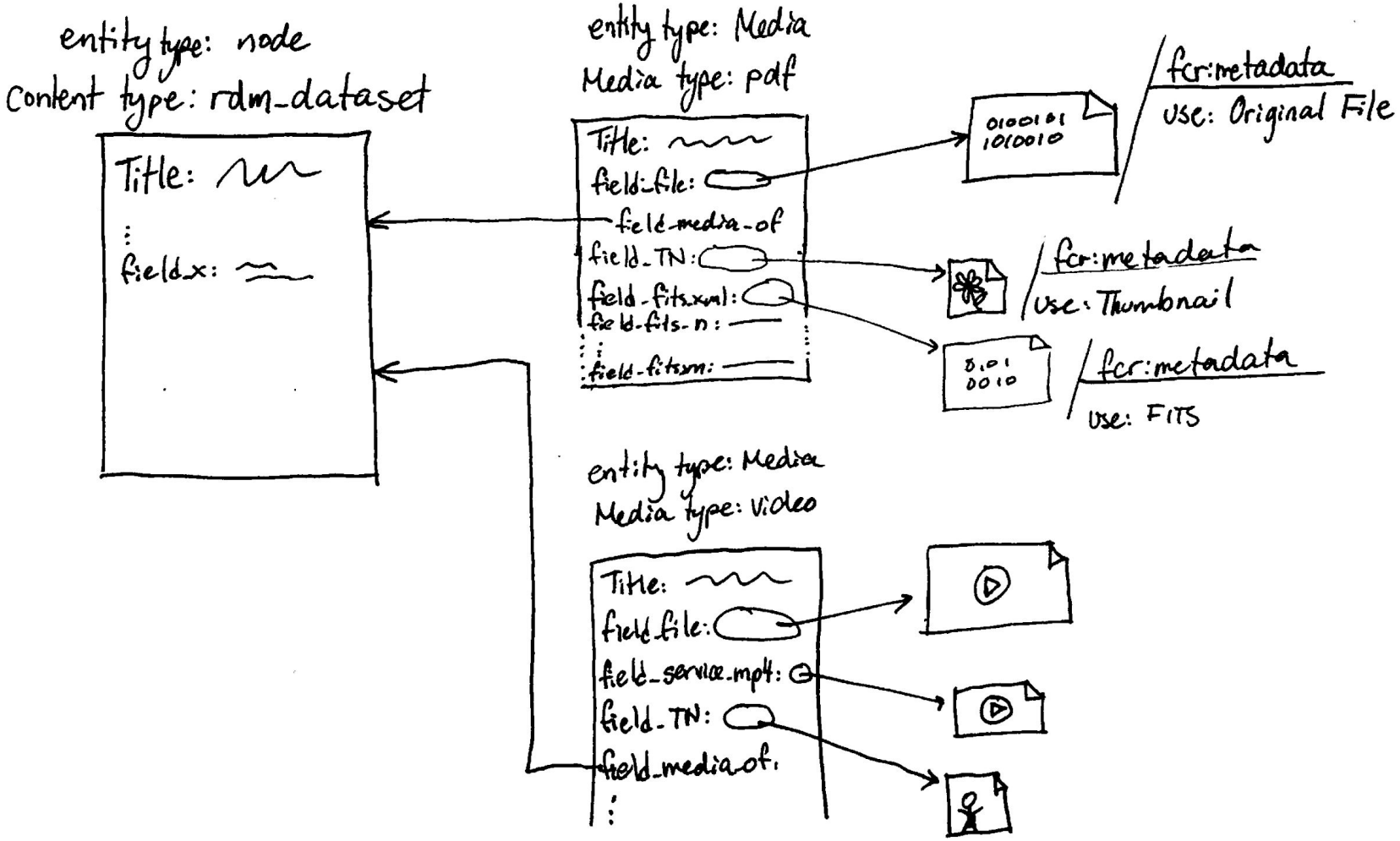
DEPOSIT

Researchers

- Can share view/edit rights
- Can manage the editorial workflow



Research Dataset Model

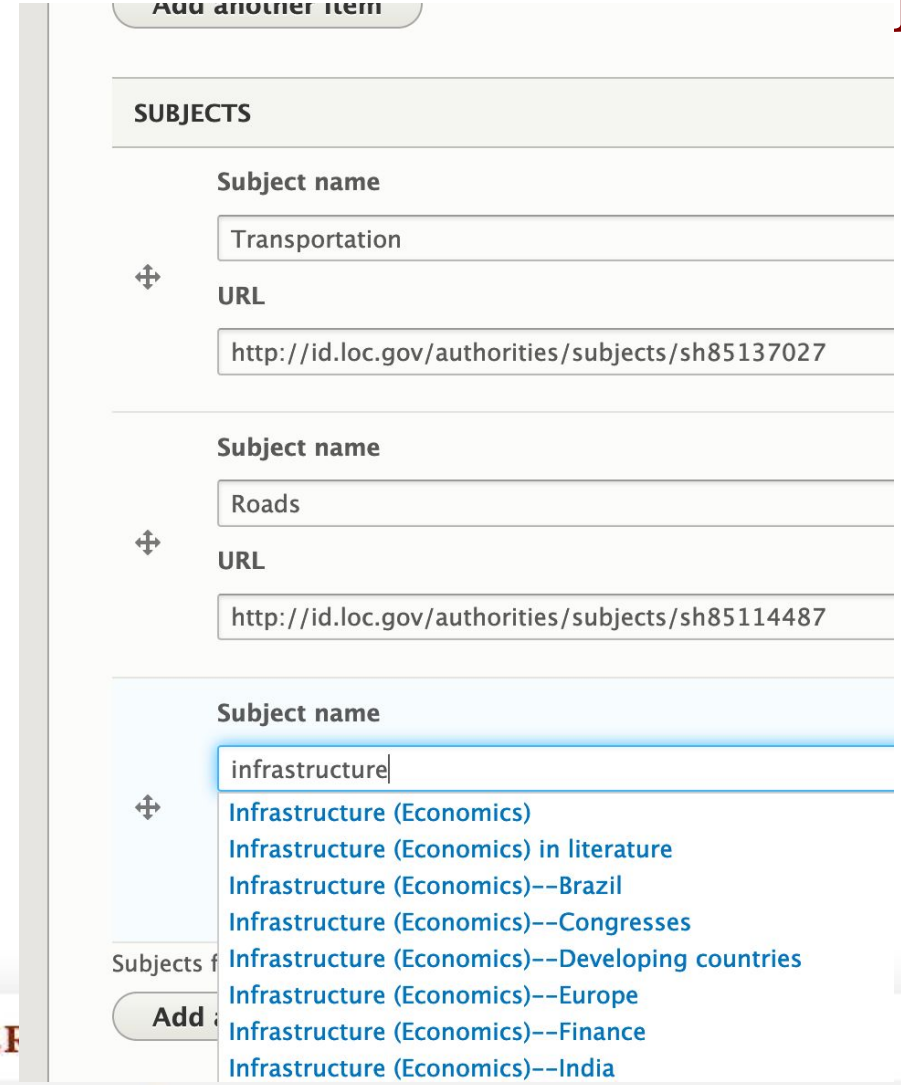


DEPOSIT

Research Dataset

- Metadata modeled on Datacite 4.2 schema
- Reflected in Drupal as fields, inline entity forms and paragraphs
- External Lookups
- Serialize as Datacite XML
- Metadata not yet fully mapped to Fedora
- Datacite DOI created on Publish

Research Dataset Metadata
LoC Lookup



DEPOSIT

UNIVER

people • excellence • impact

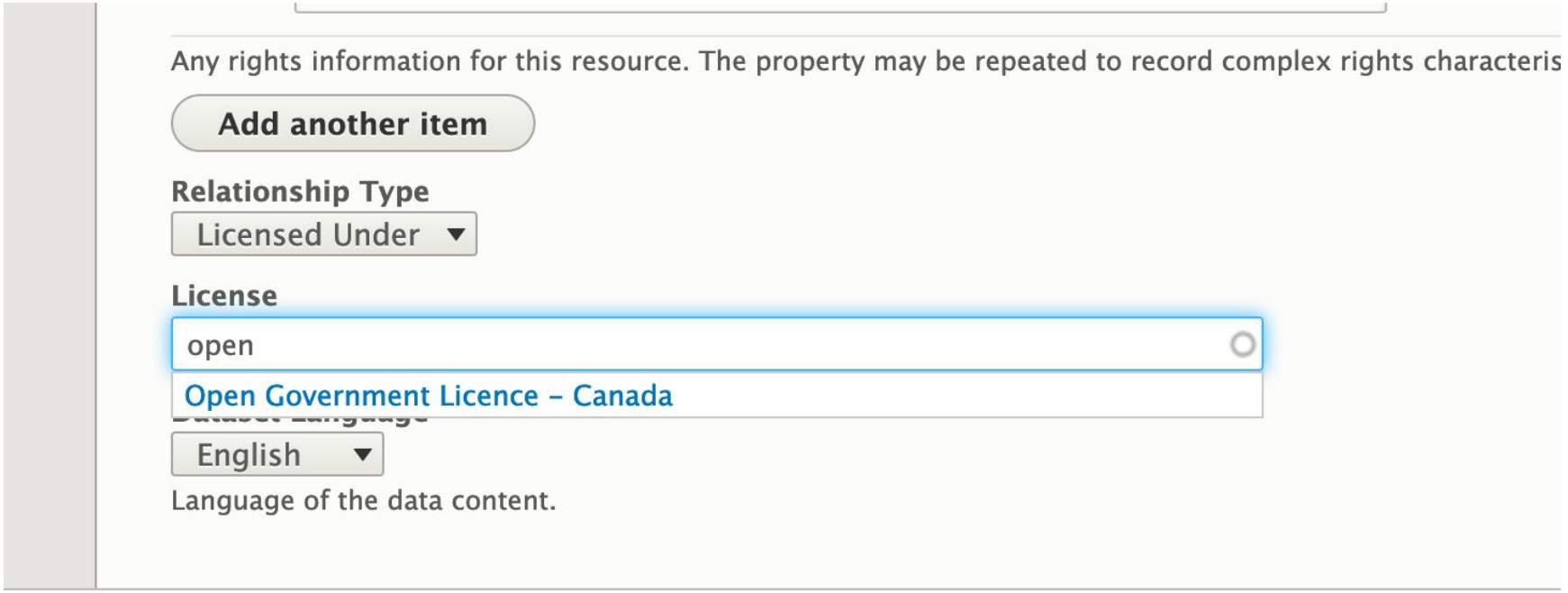
DEPOSIT

Research Dataset Metadata

This XML file does not appear to have any style information associated with it. The document tree is shown below.


```
▼<resource xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns="http://datacite.org/schema/kernel-4"
 xsi:schemaLocation="http://datacite.org/schema/kernel-4 http://schema.datacite.org/meta/kernel-4.1/metadata.xsd">
  <identifier identifierType="DOI">none</identifier>
  <resourceType resourceTypeGeneral="Dataset">Research Tool</resourceType>
 ▼<creators>
  ▼<creator>
     <creatorName nameType="Organizational">Statistics Canad - Statistical Geomatics Centre </creatorName>
   </creator>
  </creators>
 ▼<titles>
  ▼<title xml:lang="en">
     National Road Network - NRN - Prince Edward Island - GeoBase Series
   </title>
  </titles>
  <publicationYear>2015</publicationYear>
  <publisher>Statistics Canada</publisher>
 ▼<dates>
    <date dateType="Updated">2019-11-07</date>
    <date dateType="Collected">1979-07-01</date>
    <date dateType="Collected">2014-12-01</date>
  </dates>
 ▼<subjects>
    <subject>Transportation</subject>
    <subject>Roads</subject>
  </subjects>
 ▼<descriptions>
  ▼<description descriptionType="Abstract">
     The NRN product is distributed in the form of thirteen provincial or territorial datasets and consists of two linear entities (Road Segment and
     Ferry Connection Segment) and three punctual entities (Junction, Blocked Passage, Toll Point) with which is associated a series of descriptive
     attributes such as, among others: First House Number, Last House Number, Street Name Body, Place Name, Functional Road Class, Pavement Status,
     Number Of Lanes, Structure Type, Route Number, Route Name, Exit Number. The development of the NRN was realized by means of individual meetings
     and national workshops with interested data providers from the federal, provincial, territorial and municipal governments. In 2005, the NRN
     edition 2.0 was alternately adopted by members from the Inter-Agency Committee on Geomatics (IACG) and the Canadian Council on Geomatics (CCOG).
     The NRN content largely conforms to the ISO 14825 from ISO/TC 204.
   </description>
  </descriptions>
</resource>
```

Research Dataset Metadata Licensing Information



DEPOSIT

DEPOSIT

Depositing [Large] Files

- Explored different options
- Tested Globus Integration
- TUS/Uppy to support large file upload

Funding Information

- Funding information is its own content type and is integrated into the research datasets using the Entity Browser widget.
- Crossref FundRef

FundRef Lookup

Funding Organization

Funder

Social Sciences and Humanities Research

Irish Research Council for the Humanities and Social Sciences
Centre for Research in the Arts, Social Sciences and Humanities, University of Cambridge
Social Sciences and Humanities Research Council of Canada

Grant Number (optional)

Funding Organization

Funder

Social Sciences and Humanities Research Council of Canada

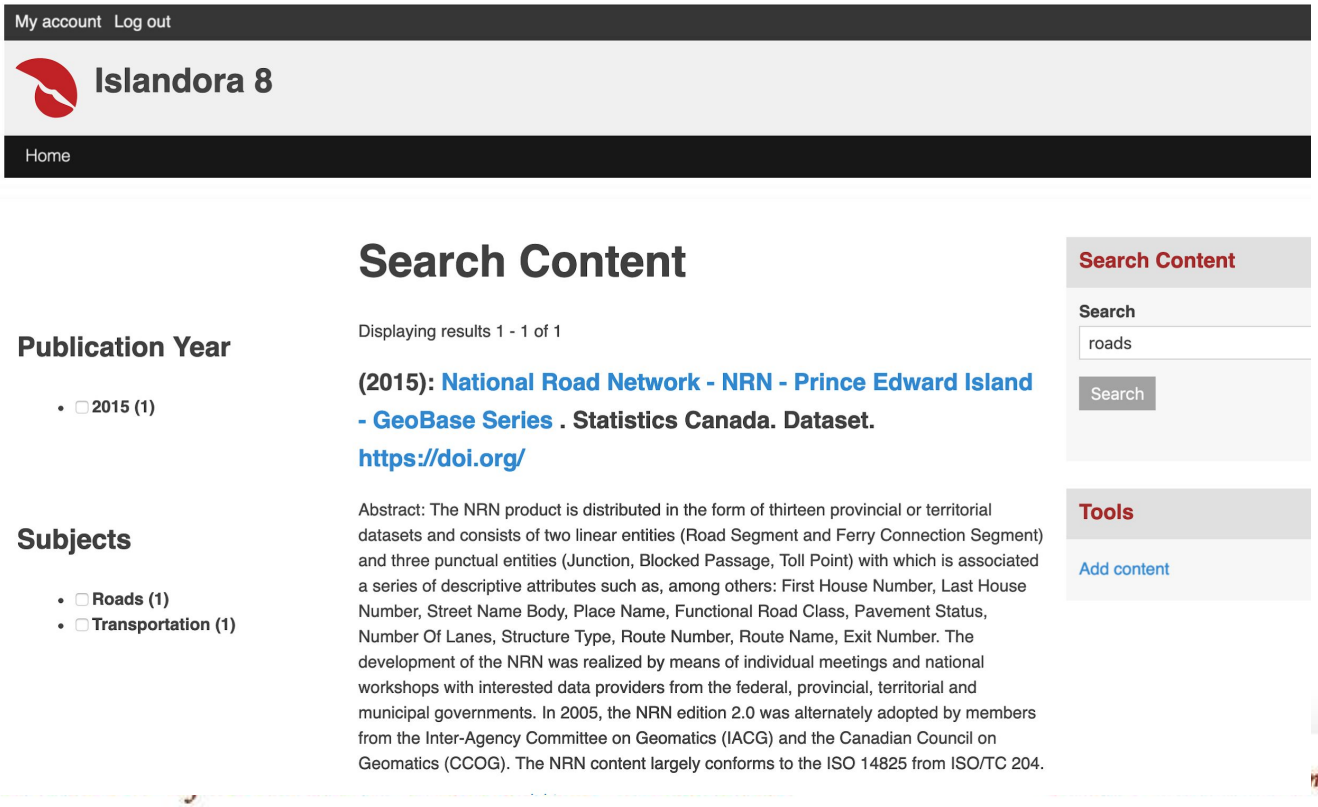
Crossref Funder DOI

http://dx.doi.org/10.13039/501100000155

DISCOVER

Data discoverability: Solr

- Dataset metadata indexed in Solr



DISCOVER

Data discoverability

- Dataset metadata is surfaced in Datacite Search platform
- Can be harvested via OAI-PMH
- Simple Sitemap XML
- Harvestable by Google Dataset Search

DISCOVER

Simple XML Sitemap

This text will be displayed on the *Add new content* page.

Submission form settings
Title

Publishing options
Published, Promoted to front page, Create new revision

Language settings
Site's default language (English)

Display settings
Display author and date information

Menu settings

Simple XML Sitemap
Included in sitemap variants: default

Sitemap variants

▼ *DEFAULT*

- ○ Do not index entities of type *RDM Dataset* in variant *Default*
- ◉ Index entities of type *RDM Dataset* in variant *Default*

Priority
0.5
The priority entities of this type will have in the eyes of search engine bots.

Change frequency
daily
The frequency with which entities of this type change. Search engine bots may take

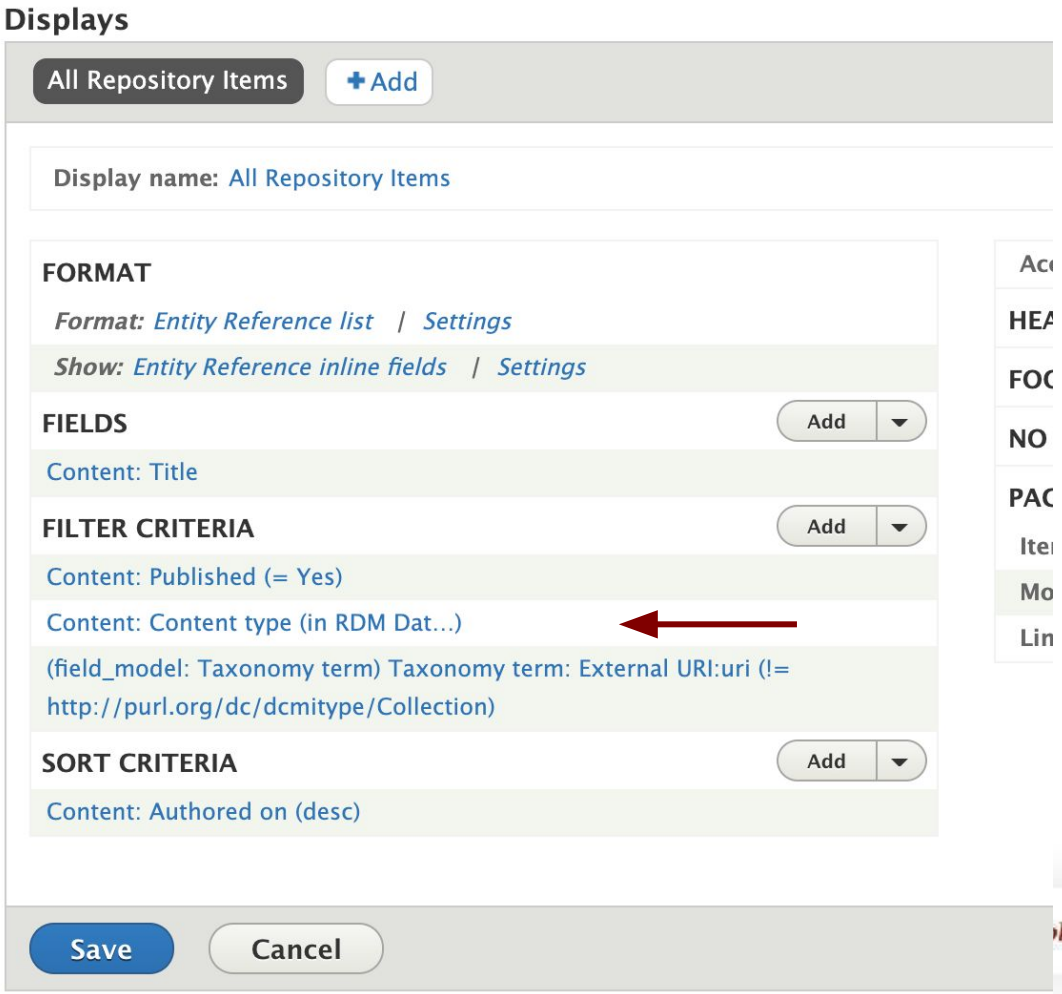
Include images
No
Determines if images referenced by entities of this type should be included in the si

☐ Regenerate all sitemaps after hitting *Save*
This setting will regenerate all sitemaps including the above changes.
Otherwise the sitemaps will be regenerated during a future cron run.

Save content type Delete

DISCOVER

Data discoverability: OAI-PMH



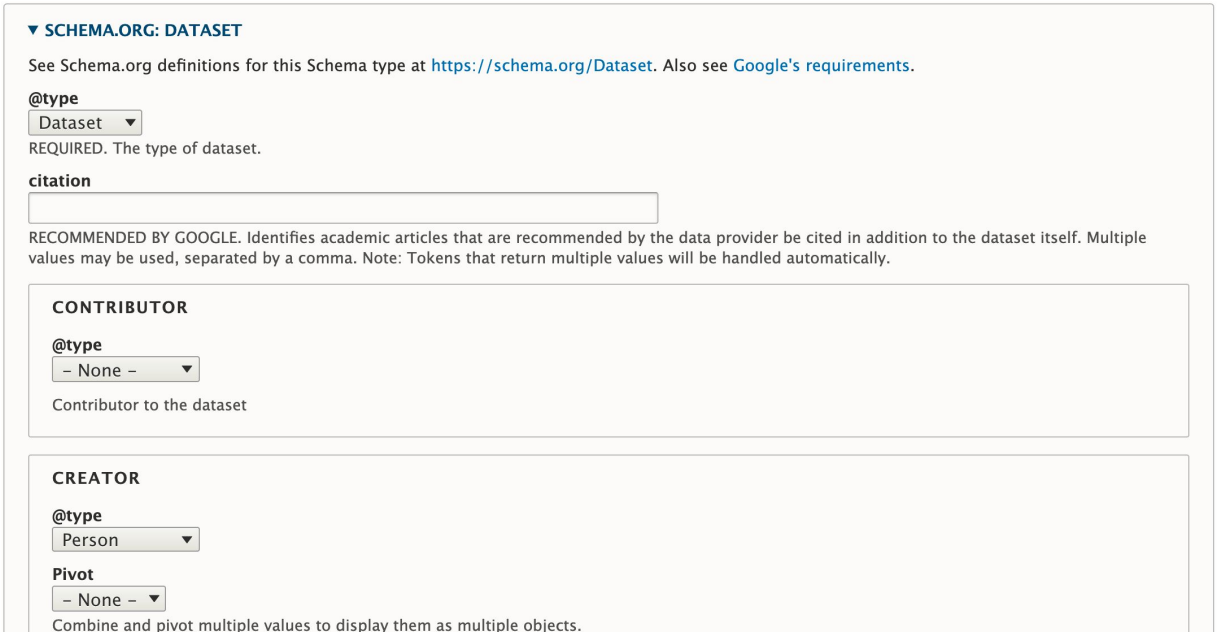
DISCOVER

Data discoverability: OAI-PMH


```
▼<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd" name="OAI-PMH">
  <responseDate>2020-02-03T12:43:29Z</responseDate>
  <request verb="ListRecords" metadataPrefix="oai_dc">http://rdm-upei.researchspaces.ca/oai/request</request>
 ▼<ListRecords>
   <resumptionToken completeListSize="20" cursor="0" expirationDate="2020-02-03T13:43:29Z">14</resumptionToken>
  ▼<record>
   ▼<header>
     <identifier>oai:rdm-upei.researchspaces.ca:node-13</identifier>
     <datestamp>2019-11-06T16:49:44Z</datestamp>
     <setSpec>oai_pmh:entity_reference_1</setSpec>
    </header>
   ▼<metadata>
    ▼<oai_dc xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
     http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
      <dc:title>Atlantic Canada Newspaper Survey - Dataset</dc:title>
      <dc:creator/>
      <dc:creator>University of New Brunswick at Saint John</dc:creator>
      <dc:subject>Newspapers</dc:subject>
      <dc:subject>Advertisers</dc:subject>
      <dc:subject>Material culture</dc:subject>
     ▼<dc:description>
        The Atlantic Canada Newspaper Survey (ACNS) project was designed to provide a machine-readable database of goods and
        services offered for sale in important east coast regional newspapers before 1900. This newspaper research database has as
        its focus nine broad categories of commodities based upon Statistics Canada's "Trade of Canada Commodity Classification".
        In the 1980s work on the database was supported by the Canadian Museum of Civilization. The project received SSHRC funding
        in 1986-1987 as part of its Research Tools Program. In June 2013, the Canadian Heritage Information Network was no longer
        able to host the database due to accessibility issues. CHIN exported the dataset and provided a copy to UPEI in the hopes
        that it would be useful.
      </dc:description>
      <dc:publisher>University of Prince Edward Island</dc:publisher>
      <dc:contributor>,,</dc:contributor>
      <dc:date>2017</dc:date>
      <dc:type>Dataset</dc:type>
      <dc:identifier>10.11571/upei-roblib-data/researchdata:448</dc:identifier>
      <dc:language>en</dc:language>
      <dc:coverage>,</dc:coverage>
    </oai_dc>
   </metadata>
  </record>
```

DISCOVER

Data discoverability: Google Dataset Search




```
<script type="application/ld+json">{
    "@context": "https://schema.org",
    "@graph": [
        {
            "keywords": "Newspapers,Advertising,Material culture",
            "contributor": [],
            "description": "The Atlantic Canada Newspaper Survey (ACNS) project was designed to provide a machine-
readable database of goods and services offered for sale in important east coast regional newspapers before 1900.
This newspaper research database has as its focus nine broad categories of commodities based upon Statistics
Canada's "Trade of Canada Commodity Classification". In the 1980s work on the database was supported by the Canadian
Museum of Civilization. The project received SSHRC funding in 1986-1987 as part of its Research Tools Program. In
June 2013, the Canadian Heritage Information Network was no longer able to host the database due to accessibility
issues. CHIN exported the dataset and provided a copy to UPEI in the hopes that it would be useful."
        }
    ]
}</script>
```

PRESERVE

Preservation Strategies

- Extract/Store/Index technical metadata with FITS microservice
- Integration with Islandora Riprap for checksum monitoring
- Bagit Microservice
- File Normalization - in progress

REUSE

Reuse

- Metadata in a variety of formats
- License is exposed
- Islandora 8 RDM Ansible Playbook
  - https://github.com/roblib/rdm-playbook
- Modules contributed back to the community

REUSE

RDM and Islandora 8

- Build capacity in the community around research data
- Reuse / repurpose existing community code
- Provide migration paths for existing Islandora users
- Demonstrate the flexibility of the framework

UNIVERSITY
of Prince Edward
ISLAND
Thank you
Contact Info:
dmoses@upei.ca
aoneill@upei.ca
rlefaive@upei.ca
Github repo
https://islandora-rdm.researchspaces.ca/
people ▪ excellence ▪ impact