

Re-imaging the Institutional Repository

Presenters: Donald Moses, Dawn Hooper, and Paul Pound - University of Prince Edward Island
Potential OR Tracks - Fedora General Track or FUG Islandora Track.

Abstract

Institutional Repositories (IR) are often the first digital collection project an institution undertakes. At the University of Prince Edward Island strategic goals and open access policies have encouraged the development of an Islandora-based IR solution. Participants will gain an understanding of the IR's development; creation of an open access policy; workflow and batch ingest features; tools for exporting citations as bibliographies and records; integration of Sherpa/Romeo, COiNS, export to RefWorks, and social media tools; repository metrics; authority records and the development of Scholar profiles; and search/discovery and theming layers.

Institutional Repositories

The aims of an institutional repository (IR) are both practical and philosophical. From a practical aspect an IR provides a framework to manage, steward, and increase the visibility of the intellectual and research output of an institution and its members. An IR also plays a role in disrupting the traditional scholarly publishing model. IRs can serve as a publishing platform and leveraging social media or other metrics can act as an additional method of measuring impact.

The Evolution of UPEI's IR

The provision of an IR service has been a priority for the Robertson Library since 2008 with the release of University of Prince Edward Island's *Strategic Research Plan 2008-2018* included as a strategic priority the development of "a publicly accessible institutional repository which will contain citations for all research conducted at UPEI, as well as, in accord with copyright protocols, access to copies of our research documents, publications, and reports"¹. The first version of IslandScholar.ca was created shortly after and included the seeding of the repository with faculty citations harvested from RefWorks, authentication with UPEI's LDAP service, the ability for faculty to add their own citations and attach the related publication, and integration with Sherpa/Romeo, OpenURL and COiNS.

During February 2012 the University released a new policy entitled *Open Access & Dissemination of Research Output*² which encourages:

¹ *Strategic Research Plan 2008-2018* - <http://research.upei.ca/files/research/v9%20Senate%2022Apr08.pdf>

² *Open Access & Dissemination of Research Output* -

- scholars to deposit their scholarly research in IslandScholar,
- scholars to publish in journals that allow scholars to retain copyright in their work,
- scholars to consider open access and affordable / sustainable scholarly communication venues,
- scholars to deposit other types of research output in IslandScholar (preprints, research data, etc.),
- students to deposit their honours thesis, Master's or PhD thesis on acceptance.

The policy specifies additional requirements such as ensuring that the content was harvestable and that deposited research data be linked to the final scholarly output. To fulfill these needs and others, a set of requirements was drafted for an IslandScholar v2 and the new service was launched in October 2012.

Workflow

To ensure a sustainable workflow, the IR includes bulk ingest of citations in various formats including RefWorks, Endnote XML and RIS. In addition to batch ingest, the repository also has a suite of format specific forms that can be used to add citations individually using Islandora's XML forms framework. Regardless of the initial format, citation data is normalized and stored as MODS in the repository and is exposed to users in a standard citation format using the Citation Style Language (CSL)³, citeproc-js⁴, and a purpose-built converter⁵. When documents are attached to citations (eg. Word, OpenOffice, PDF) using IslandScholar's upload feature the original is sent through Islandora's microservices and transformed to PDF/a. The PDF/a is used for dissemination and longterm preservation.

User tools

The IR includes a number of tools for users including:

- integration with Sherpa/Romeo so that scholars can quickly check publisher copyright policies and self-archiving options
- COiNS compliant metadata so that users can connect to the full text of citations in databases their institution subscribe to (Check@, GETIT, etc.)
- "My List" feature allows users to add citations to a list for export as a bibliography or in another citation format
- RSS feeds of a scholar's citation are available from their profile
- export to RefWorks

<https://cab.upei.ca/sites/default/files/attachments/OpenAccessandDisseminationofResearchOutput.pdf>

³ Citation Style Language - <http://citationstyles.org/>

⁴ citeproc-js. <https://bitbucket.org/fbennett/citeproc-js/wiki/Home>

⁵ IslandScholar Github - converter.php.

https://github.com/roblib/islandora_scholar_upei/blob/master/modules/citeproc/generators/converter.php

Repository Metrics and Enabling Social Media

Given the growing importance of alternative metrics that scholars can use in their portfolios and for the institution to measure the impact of their research output, IslandScholar utilizes a pair approach to collecting data. Traditional web server statistics are collected both on the server and via the Drupal Google Analytics module. As well, IslandScholar keeps a separate count of citation views and fulltext downloads. While administrators can view server statistics, view/download metrics are publicly available from the scholar and departmental profile pages. To take advantage of the emerging altmetrics trends IslandScholar utilizes the AddThis⁶ service and is considering the integration of ImpactStory⁷.

Authorities and Scholar Profiles

IslandScholar implements authority records using the MADS⁸ schema for scholar and departmental/organizational records. Since the citations don't have a consistent way of identifying authors, we embed local identifiers into the citation for both scholars and departments. We relate these identifiers to their authority records, providing a reliable way of generating views based on authors and departments. Using the authority record we are able to generate a profile of the scholar which includes a thumbnail image, metadata from the authority record (biography, research interests, role, email, etc.), the citations associated with the scholar, a list of scholars in the associated department, and the grants associated with the scholar. The grants are generated from data extracted from Research Services, ingested into the repository, and associated with the scholar and the department.

Search and Discovery

The IslandScholar IR makes extensive use of Apache Solr for both browse and search interfaces. Both simple keyword and field-based searching is supported and results are displayed based on the default CSL style selected. As well, facets are available so that users can limit the results based on date range, type, department and collection.

One of the requirements of UPEI's new open access policy was to make the citation data open and harvestable by others. The project leveraged and customized an existing Islandora OAI module⁹. The module is integrated with IslandScholar and provides OAI-PMH access to collections within the repository. Scholar data is exposed as DC, while theses data is exposed as ETD-MS to meet the data harvesting requirements for theses by Library and Archives Canada.

Islandora IR Case Studies

⁶ AddThis - <http://www.addthis.com/>

⁷ ImpactStory - <http://impactstory.org>

⁸ Metadata Authority Description Schema (MADS). <http://www.loc.gov/standards/mads/>

⁹ islandora_oai module - https://github.com/Islandora/islandora_oai

The scholar module developed at UPEI has been used as a base for the development of other institutional repositories. New features have been added and the code has been refactored and migrated to Drupal 7. Additional import options based on PMID and DOI identifiers have been developed along with integration with Google Analytics for metrics gathering and custom reporting. The following IRs will be reviewed:

- IslandScholar
 - <http://islandscholar.ca>
- Fundación Juan March
 - <http://digital.march.es/ceacs-ir/en/>
- Grinnell College
 - <http://digital.grinnell.edu/drupal/>
- DGI Cloud
 - a Drupal 7 cloud-based IR service