

Islandora Scholar

- [Introduction](#)
- [Requirements](#)
- [Installation](#)
- [Scholar "Core"](#)
 - [Features of Scholar Core](#)
 - [Citation & Thesis Content Models](#)
 - [Object Display](#)
 - [Sherpa/RoMEO](#)
 - [Standard Metadata Display \(vs. COinS\)](#)
 - [Google Scholar and Library Catalogue search Links](#)
 - [PDF Viewer](#)
 - [PDF Configuration](#)
- [About CSL \(Citation Style Language\)](#)
- [Scholar Submodules](#)
 - [Importers and Populators](#)
 - [DOI Importer/Populator - special configuration](#)
 - [Islandora Bibliography and Exporter](#)
 - [Islandora Google Scholar](#)
 - [Scholar Embargo](#)
 - [Community Best Practices for Embargoing](#)
 - [Issues with Object Embargoes](#)
- [Complementary solution packs and utility modules](#)

Introduction

Islandora Scholar is a suite of modules designed to help Islandora function as an Institutional Repository (although some features are helpful in other use cases as well). It is unique compared to other Islandora modules in terms of the number of features it provides as well as the vast amount of submodules that it contains. It is helpful to think of Scholar as being a kind of scholarly content solution pack due to the new content models it provides (citation & thesis), but it also differs from other solution packs in that it provides new functionality that may be used with other content models (cmodels) as well.

Features:

- Citation Content Model and Thesis Content Model (and associated default forms)
- Display dynamically-formatted citations based on MODS data according to on user-uploaded CSL files
- (optional) tab on Citation objects containing the self-archiving policies of the associated periodical, based on [Sherpa/RoMEO](#)
- (optional) links on Citation and Thesis objects to search for them in Google Scholar or your local discovery layer
- Importers and populators to create objects based on DOIs, PMIDs, Endnote XML, or RIS files (see associated submodules)
- Ability to embargo objects or datastreams permanently or with a time limit (see Scholar Embargo submodule)
- Google Scholar-ready microdata in meta tags (see Islandora Google Scholar submodule)
- Exporting of collections of bookmarks as RIS, RTF or PDF files (see Islandora Bookmark and Exporter submodules)

Requirements

This module requires the following external modules/libraries:

- [Islandora](#)
- [Tuque](#)
- [Islandora Solr](#)

Islandora Scholar also requires the following submodules which are bundled in the /modules/ directory:

- [Citeproc](#) (a Drupal module, which also requires that the [citeproc-php](#) library is installed in the sites/all/libraries directory, such that the main CiteProc.php file is located at sites/all/libraries/citeproc-php/CiteProc.php.)
- [CSL](#)
- [Bibutils](#) (a Drupal module, which also requires that Bibutils is installed on the underlying operating system. Follow [these instructions](#) for installing Bibutils.)

Installation

See [this](#) for further information about standard installation of Drupal 7 modules.

Please note that additional features are available through optional Scholar submodules. They must be enabled separately, and they have their own separate requirements as well.

Namespaces



*Please note that the `citationCModel` and `thesisCModel` are in the "ir:" namespace, as opposed to the standard "islandora" namespace that other `cmodels` use. Due to this nonstandard namespace, Islandora instances that make use of **namespace restrictions** will need to enable the "ir" namespace, both in Islandora and in Islandora Solr, in order for the citation & thesis content models to work.*

Scholar "Core"

While "Scholar Core" is not an official community term, it can be a helpful way to keep track of the numerous features Islandora Scholar offers. In this context, "core" features are those that come from the Islandora Scholar module itself as opposed to coming from one of the numerous submodules which can be enabled or disabled independently.

Features of Scholar Core

Citation & Thesis Content Models

The core Islandora Scholar module provides two new content models: the **Citation Content Model** (`ir:citationCModel`) and the **Thesis Content Model** (`ir:thesisCModel`). The Citation model is intended for general scholarly works, while the Thesis model is for handling electronic theses and dissertations (ETDs). While the datastream structure and default display of these two Content Models is nearly identical, they can be tweaked separately to accommodate the different use cases one may have for theses vs. other scholarly works, such as altered displays, collection configuration, or separate metadata forms.

Object Display

Citation and thesis objects have special components on their "landing page", including

- a formatted citation (using a configurable Citation Style Language or CSL - defaults to APA)
- a metadata "Details" tab that uses COinS metadata (may be disabled to use the Islandora Default metadata display, e.g. Islandora Solr Metadata)
- optionally, a link to search for this object in Google Scholar
- optionally, a link to search for this object in a library catalogue or discovery service
- optionally, a tab containing the copyright and self-archiving policies of the journal according to Sherpa/RoMEO.

These features can be configured on the Scholar admin page, at **Administration » Islandora » Solution Pack Configuration » Scholar** (`/admin/islandora/solution_pack_config/scholar`).

Scholar
BASIC CONFIGURATION
BIBUTILS
CITEPROC
DOI CONFIGURATION
PDF CONFIGURATION
XPATH CONFIGURATION

Home » Administration » Islandora » Solution pack configuration

ROMEO CONFIGURATION

Various parameters used by RoMEO

☐ Enable RoMEO attempts.
Whether or not we should even try getting information from RoMEO.

Sherpa/RoMEO API Key

Your institution's key to access RoMEO.

ISSN cache time

The amount of time in days for which to maintain an ISSN value obtained by scanning an object.

RoMEO cache time

The amount of time in days for which to maintain a copy of the markup for a given ISSN.

GOOGLE SCHOLAR SEARCH CONFIGURATION

Various parameters used by Google Scholar Search

☐ Render Google Scholar Search link
Enable Searching in Google Scholar.

LIBRARY CATALOG SEARCH CONFIGURATION

Various parameters used by your Library Catalog

☐ Render Library Catalog Search link
Enable Searching in your library discovery layer.

☐ Use Standard Metadata Display
Check to use standard Islandora metadata display instead of COinS.

☐ Let users choose display CSL
In object view let the users choose which CSL to use.

☐ Specify Document Versions
Use your own document versions when uploading a PDF

☐ Specify Use Permissions
Use your own use permissions when uploading a PDF

VIEWERS

Select a viewer

DEFAULT	LABEL	DESCRIPTION	CONFIGURATION
☒	None	Don't use a viewer for this solution pack.	
☐	pdf.js Reader	Mozilla pdf.js Reader.	

Preferred viewer for your solution pack. These may be provided by third-party modules.

Save configuration

Sherpa/RoMEO

Sherpa/RoMEO is a service which provides information about the copyright & self-archiving policies of various academic journals. When this feature is enabled, a "RoMEO" tab will appear on objects with the Citation Content Model which have a MODS identifier of type "issn". This can be used in staff workflows, to help encourage self-archiving. Results of RoMEO queries are cached in the Drupal database, but are not saved into object metadata.

Without an API key, Sherpa/RoMEO provides up to 500 requests per day. With an API key, requests are unlimited. [Sherpa/RoMEO API](#) registration is free-of-charge as of spring 2018.

[illegible]

Conditions

- [illegible]

Roman Column: 2nd c.



Standard Metadata Display (vs. COinS)

Islandora Scholar provides a special metadata display block for Citation and Thesis objects, using the [COinS](#) metadata standard. COinS is the standard that is used in OpenURL implementations, and is a standard set of terms for citation objects including "Title", "Authors", "Abstract", "Journal", "ISSN", etc. The COinS metadata ("Details") block pulls this information from the MODS datastream according to a hard-coded transform. To customize the metadata display, select "Use Standard Metadata Display", set Islandora to use "Islandora Solr" as a metadata display, and configure a [Solr metadata profile](#) for Citation and/or Thesis objects.

The COinS display is "expanded by default", unlike the Default DC or Islandora Solr metadata display blocks.

Google Scholar and Library Catalogue search Links

Scholar can include "search links" on the display of Citation and Thesis objects (formatted URLs using metadata from the object, to direct the user to a Google Scholar search or a library catalogue/discovery layer search). The metadata used to create these links can be configured by providing an XPath, which will be applied to the MODS datastream. By default, it will use the Primary Search XPath (defaults to `//mods:identifier[@type="doi"]`), which will find any DOI identifier present in the MODS record). If the Primary Search XPath does not find anything, it will fall back to use the Default Search XPath. If this does not provide a search string either, the object label will be used.

PDF Viewer

Citation and Thesis objects can be associated with PDF datastreams (by default, the datastream ID is 'PDF', which differs from the standard Islandora use of 'OBJ' for the payload, or 'preservation master'). The PDF datastream can be displayed in an in-browser javascript viewer. The [Islandora PDF.js](#) module, which leverages [Mozilla's PDF.js](#) library, is an example (the only officially supported PDF viewing module as of 2018). Whether or not to use a viewer can be configured at the bottom of the Scholar admin page, in the "VIEWERS" section. If you see "No viewers detected", this means that there is a problem with your installation of Islandora PDF.js or that you have not enabled it. If no viewer is selected, citation and thesis object displays will show an image of the cover page of the PDF datastream for that object (PREVIEW datastream, configured below).

PDF Configuration

How uploaded PDFs are handled by Scholar can be configured at **Administration » Islandora » Solution Pack Configuration » Scholar » PDF Configuration** (/admin/islandora/solution_pack_config/scholar/pdf).

Like the [PDF solution pack](#), Scholar supports extracting searchable text from uploaded PDFs. This relies on the PDF having an embedded text stream (e.g. a born-digital PDF, or one with OCR already performed and embedded). Scholar will not generate OCR from raw scans. Consider converting text-filled images with no text streams to TIFFs, and using the [Book Solution Pack](#) with OCR enabled.

The sizes of the thumbnail (TN) and cover page (PREVIEW) datastreams, which are derived from the first page of the PDF using imagemagick, are also configurable here. Below, a screenshot of the PDF configuration screen.

Scholar
islandora-development.org
BASIC CONFIGURATION
BIBUTILS
CITEPROC
DOI CONFIGURATION
PDF CONFIGURATION
XPATH CONFIGURATION

Home » Administration » Islandora » Solution pack configuration » Scholar

TEXT

☒ Extract text streams from PDFs using pdftotext

The extracted text stream is added as FULL_TEXT datastreams and are indexed into Solr. Uploading a text file takes priority over text stream extraction.

Note: PDFs that contain visible text do not necessarily contain text streams (e.g. images scanned and saved as PDFs). Consider converting text-filled images with no text streams to TIFFs and using the [Book Solution Pack](#) with OCR enabled.

Path to pdftotext executable

✓ pdftotext executable found at `/usr/bin/pdftotext`

THUMBNAIL

Settings for creating PDF thumbnail derivatives

Width

The width of the thumbnail in pixels.

Height

The height of the thumbnail in pixels.

PREVIEW IMAGE

Settings for creating PDF preview image derivatives

Max width

The maximum width of the preview in pixels.

Max height

The maximum height of the preview in pixels.

Density used to generate preview

The imagemagick density value used to generate the preview. Generally, higher numbers are better but require more computation time.

Save configuration

About CSL (Citation Style Language)

Citation Style Language, or CSL, is an XML standard for describing the formatting of citation styles, such as APA, MLA, or Chicago. CSL styles are used by Scholar to transform MODS metadata into a formatted citation, which is displayed at the top of a Citation or Thesis object pages. The submodule Islandora Bibliography (described below), uses CSL to export a user-selected list of citations as a formatted bibliography.

Scholar comes with only the APA citation style built in. To configure new citation styles, go to the [Zotero CSL Repository](#) and download whatever CSL files you want to use. You can then upload CSL files at Configuration » Islandora » Islandora Utility Modules » CSL (`admin/islandora/tools/csl`) and choose a default style.

STYLE NAME		DELETE
APA	Is Default	Delete
Test Style	Set Default	Delete
Test Style2	Set Default	Delete
Name	Upload A CSL Style No file chosen	Add

If you want your users to be able to dynamically choose any available citation style from the citation/thesis object display interface, make sure to check the "Let users choose display CSL" checkbox at the bottom of the Islandora Scholar admin page at `admin/islandora/solution_pack_config/scholar`.

Greenways as green magnets

[View](#)[Document](#)[Manage](#)[MARCXML](#)

Style

☒ APA

Archivum Immunologiae et Therapiae Experimentalis
The Open University (Harvard)

Coutts, C., & Miles, R. (2011). Greenways as green magnets: The relationship between the race of greenway users and race in proximal neighborhoods. *Journal Of Leisure Researc*. Retrieved from http://purl.flvc.org/fsu/fd/FSU_migr_durp_faculty_publications-0013

Florida State University Libraries

2011

Greenways as green magnets: The relationship between the race of greenway users and race in proximal neighborhoods

Christopher Coutts and Rebecca Miles

The appropriate type of citation will be displayed according to the content of the <genre> tag in the MODS record for each Citation or Thesis object. The default thesis form includes a field called "Genre"; in the default Citation form, this is the "Publication Type" field near the top of the form. A case-sensitive entry is required in order to generate the correct formatting:

MODS <genre> term	Format of the citation generated by Islandora Scholar
blank/default/ any-term-not- in-a-controlled -vocab	Citation will be formatted like a journal article.
<genre>journal article</genre>	Citation will be formatted like a journal article.
<genre>book chapter</genre>	Citation will be formatted like a book chapter.
<genre>book section</genre>	Citation will be formatted like a book chapter.
<genre>book</genre>	Citation will be formatted like a book.

Scholar Submodules

Importers and Populators

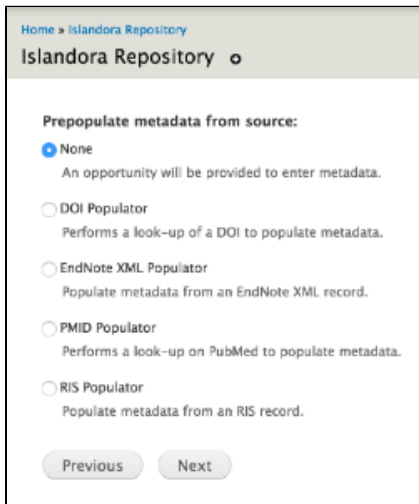
Scholar provides options for importing objects from various sources. These are provided by the following Scholar submodules:

- **PMID** ([PMID Populator](#) / [PMID Importer](#)) - Given a PubMed ID, imports metadata through the PubMed API
- **DOI** ([DOI Populator](#)/[DOI Importer](#)) - Given a Crossref DOI, imports metadata through the Crossref API
- **RIS** ([RIS Populator](#)/[RIS Importer](#)) - Parses citations from a Research Information Systems (RIS) formatted file
- **EndNote XML** ([EndNoteXML Populator](#)/[EndNoteXML Importer](#)) - Parses citations from an EndNote XML formatted file

A **Populator** is used when creating a single object, and will pre-fill a metadata form, allowing you to edit the data before ingesting. An **Importer** will automatically create one or more objects in a batch.

To use a Populator, follow the normal workflow to add an object to a collection. After selecting a content model, a step called **"Prepopulate metadata from source"** will display an option for each enabled Populator module. If one is selected, the user will be able to upload the file or provide the identifier, and then will be shown their associated metadata input form prepopulated with metadata from the source.

Be aware that some metadata from the source might not display in the form, but might be saved in the object's MODS datastream regardless. The transform from the sources to MODS XML are defined by each populator/importer module, but the metadata form that displays can be configured by an Islandora administrator. If the form is missing fields for certain MODS elements, such as within `<mods:relatedItem type="host">`, that information may not show up in the form that the user sees, but will still be saved in MODS. Populator modules can be used to create objects of any content type, and this may happen if creating an object with a form that was not designed for citations.



Home » Islandora Repository

Islandora Repository

Prepopulate metadata from source:

- ☒ None
An opportunity will be provided to enter metadata.
- ☐ DOI Populator
Performs a look-up of a DOI to populate metadata.
- ☐ EndNote XML Populator
Populate metadata from an EndNote XML record.
- ☐ PMID Populator
Performs a look-up on PubMed to populate metadata.
- ☐ RIS Populator
Populate metadata from an RIS record.

Previous Next

To use an Importer, log in, and browse to a Collection. Click the **"Manage"** tab, then the **"Collection"** sub-tab, then click **"Batch Import Objects"**. You will get a drop down menu with an option to use the "DOI Importer" and any other importers that are enabled. Upload a file or list of DOIs, and click "Import" to automatically import the batch of objects. This uses the Islandora Batch interface behind the scenes, and you will be redirected to a "queue" if the ingest was successful. See [How to Batch Ingest Files](#) for more information about ingest queues.

The importers bundled with Scholar will only create Citation content model objects. If your collection policy is set up so that the collection can hold Citation Content Model objects, then the import will create these objects, and will create them in the namespace configured for Citation objects in the collection. If your collection policy is not set up to hold Citation Content Model objects, then the import will create objects in the `ir:` namespace. If namespace restrictions are enabled, these objects might not show up as expected.



Home » Islandora Repository

Islandora Repository

Importer

- ✓ DOI Importer
- EndNote XML Importer
- Pubmed Importer
- RIS Importer
- ZIP File Importer

DOI Importer/Populator - special configuration

There are two available APIs for retrieving metadata based on DOIs. The "Legacy" API is [Crossref's OpenURL API](#), which [requires registration](#) via an email address. The new API (as of Islandora 7.x-1.11) uses [Content Negotiation via doi.org](#). This API does not require registration, and will work retrieve metadata from CrossRef, DataCite, and mEDRA. New installations of the DOI importer module will default to using the new API; installations which have already been configured to use CrossRef will continue to use CrossRef.

Configure the API settings at **Home » Administration » Islandora » Solution pack configuration » Scholar » DOI Configuration** (`/admin/islandora/solution_pack_config/scholar/islandora_doi`).

Creating DOIs

Islandora Scholar does not provide the functionality of creating, or minting, DOIs. There are community instances running custom modules to do this, for example, https://github.com/SFULibrary/islandora_doi_framework.

Islandora Bibliography and Exporter

The **Islandora Bibliography** submodule extends **Islandora Bookmark** to allow it to handle collections of scholarly citations, which can be dynamically restyled with the help of the **Citeproc** and **CSL** submodules. They can also be exported in RIS, RTF or PDF format with the **Exporter** submodule.

Islandora Google Scholar

Getting indexed by Google Scholar is a primary concern for any institutional repository due to its overwhelming popularity with researchers. A properly configured IR should be seeing most of its visitors coming from Google Scholar, but Google Scholar has [many requirements for repositories who want to be indexed](#). The **Islandora Google Scholar** submodule embeds relevant meta tags in the html of Citation and Thesis object pages, so that Google Scholar learns about your repository and its contents. See the meta tags with the "citation_" prefix in the screenshot below:

```
<head>
<meta charset="utf-8" />
<meta name="citation_abstract_html_url" content="http://diginole.lib.fsu.edu/islandora/object/fsu%3A207165/" />
<link rel="shortcut icon" href="http://diginole.lib.fsu.edu/sites/fsu.digital.flvc.org/files/favicon.ico" type="image/vnd.microsoft.icon" />
<meta name="citation_online_date" content="2015" />
<meta name="citation_pdf_url" content="http://diginole.lib.fsu.edu/islandora/object/fsu%3A207165/datastream/PDF/view" />
<meta name="citation_publication_date" content="2016" />
<meta name="citation_author" content="Vandegrift, Micah" />
<meta name="Generator" content="Drupal 7 (http://drupal.org)" />
<meta name="citation_title" content="Faculty Senate Library Committee - Task Force on Scholarly Communications" />
<meta name="viewport" content="width=device-width, initial-scale=1" />
<meta name="MobileOptimized" content="width" />
<meta name="HandheldFriendly" content="true" />
<meta name="apple-mobile-web-app-capable" content="yes" />
<title>Faculty Senate Library Committee - Task Force on Scholarly Communications | fsu.digital.flvc.org</title>
<style>
```

Scholar Embargo

The **Scholar Embargo** submodule is very similar to (and can even be used in conjunction with) the **Islandora IP Embargo** module. Scholar Embargo allows you to suppress the visibility of objects in collections entirely, or allow the objects to display but disable a specified datastream on an object. Embargoes may be indefinite (meaning that they will persist until manually lifted) or temporary (meaning that they will persist until the specified expiration date chosen by an administrator, with email notifications 10 days before and on the day of an embargo's expiration).

The two types of embargoes are object embargoes and datastream embargoes. Objects with an "object embargo" will be hidden from public users; they won't display in collections or search results, and will yield an access denied page when public users attempt to access them directly by PID. Only privileged users will be able to see them (more on that below). Datastream embargoes allow the object to display normally in collections and search results, but a particular datastream (specified when the embargo is placed) will be unavailable to public users to view or download. For citation and thesis objects this will be the 'PDF' datastream, but Scholar Embargoes can be applied to any content model, extending the list of possible datastreams. In order to put a Scholar Embargo on content models other than citations and theses, you must enable them from the "Islandora Scholar Embargo" tab of the Islandora Utility Modules admin menu at admin/islandora/tools/embargo (example below).

Select Content Models for objects to be embargoed
Objects defined by selected Content Models will be eligible for embargoing.

☐ PID	CONTENT MODEL
☒ lr:citationCModel	Citation Content Model
☐ islandora:collectionCModel	Islandora Collection Content Model
☐ islandora:sp-audioCModel	Islandora Audio Content Model
☐ islandora:sp-basic_image	Islandora Basic Image Content Model
☐ islandora:pageCModel	Islandora Page Content Model
☐ islandora:bookCModel	Islandora Internet Archive Book Content Model
☐ islandora:compoundCModel	Islandora Compound Object Content Model
☐ islandora:sp-large_image_cmodel	Islandora Large Image Content Model
☐ islandora:sp-web_archive	Islandora Web ARCHive Content Model
☐ islandora:sp-videoCModel	Islandora Video Content Model
☐ lr:thesisCModel	Thesis Content Model
☐ islandora:sp-pdf	Islandora PDF Content Model
☐ islandora:newspaperCModel	Islandora Newspaper Content Model
☐ islandora:newspaperissueCModel	Islandora Newspaper Issue Content Model
☐ islandora:newspaperPageCModel	Islandora Newspaper Page Content Model
☐ islandora:OACCModel	Islandora Annotation CModel
☐ islandora:oralhistoriesCModel	Islandora Oral Histories Content Model

Configure EmbargoReset to default

You may also use the "Manage Embargoed Items" tab in the Islandora Scholar Embargo admin menu to see all objects that are currently under embargo.

LABEL	PID	TYPE	EMBARGOED UNTIL	MANAGE
Citation – Temporary PDF Embargo	fsu:6965	PDF	Apr 08, 2036	Manage
Thesis – Indefinite Object Embargo	fsu:7270	Object	Indefinite	Manage
Thesis – Temporary PDF Embargo	fsu:7273	PDF	Apr 08, 2034	Manage
Thesis – Indefinite PDF Embargo	fsu:7274	PDF	Indefinite	Manage
Thesis – Temporary Object Embargo	fsu:7278	Object	Apr 08, 2036	Manage
Citation – Temporary Object Embargo	fsu:7969	Object	Apr 08, 2036	Manage
Citation – Indefinite Object Embargo	fsu:7982	Object	Indefinite	Manage
Citation – Indefinite PDF Embargo	fsu:8006	PDF	Indefinite	Manage

Community Best Practices for Embargoing

It is worth noting that the current best practice in the world of institutional repositories (citation needed) seems to be exposing the object display and metadata, but blocking the viewing or downloading of the full text record (typically a PDF file). This is so users can know what information is in the system through collection displays or search results, even if the system blocks them from fully accessing it. Many faculty and students like to check to make sure their records are in the repository and still embargoed, and when they can't find their record they may misunderstand and think it hasn't been loaded.

Issues with Object Embargoes

In order to place embargoes on objects, you must grant the "Manage embargo on any objects" permission to the appropriate user roles under the Islandora Scholar Embargo section of admin/people/permissions. This will allow a user to place object and datastream embargoes on objects, as well as remove or update datastream embargoes. Due to issues with the connection between Drupal permissions and the way Scholar Embargo implements its embargo policies through XACML, **currently only the object's owner will be able to remove or update object embargoes**. This is because the XACML policy works independently of Drupal permissions, and the embargo XACML policy will block everyone except the object's owner from managing the object at all once embargoed regardless of what Drupal permissions they have. A [fix for this is in the works](#) and should be part of the Islandora 7.x-1.8 release.

Complementary solution packs and utility modules

The Islandora code base includes a number of solution packs and utility modules that help to enhance the functionality of Islandora Scholar as a feature-rich institutional repository:

- [Entities Solution Pack](#) - provides Scholar (person) and Department (organization) objects
- [Islandora Badges Module](#)
- [Usage Stats Utility Module](#)