

Chapter 4 - Search and Discovery in Islandora

Islandora uses Solr and the Solr module in combination with GSearch to provide search functions to users on your site. This chapter assumes either:

- a) You are using the **Virtual Machine Image**, or are exploring Islandora via Sandbox (where Solr is already installed and configured)
- b) Solr and Gsearch are installed (following the instructions in [Chapter 9 - Enabling Indexing & Searching with Solr](#)), and you have installed and activated the Solr module. The xslt and solrschema.xml documents that come packaged with the module should be used in configuration. These files are designed to work with our solution packs. Guidance for creating customizations to Solr are provided in the [Customizing GSearch and Solr](#) section.

What is Solr? And, why should I use it?

Solr makes it easy to create advanced search features in Islandora, like faceting (arranging search results in columns with numerical counts of key terms). The following comes from the Solr guide; a link to the guide is provided in the Selected Reading section of this guide:









Solr builds on another open source search technology---Lucene, a Java library that provides indexing and search technology, as well as spellchecking, hit highlighting and advanced analysis/tokenization capabilities. Both Solr and Lucene are managed by the Apache Software Foundation (www.apache.org).

The Lucene search library currently ranks among the top 15 open source projects and is one of the top 5 Apache projects, with installations at over 4,000 companies. Lucene/Solr downloads have grown nearly 10x over the past three years, with a current run-rate of over 6,000 downloads a day.

Islandora uses Solr to make objects in your Islandora installation discoverable. The Solr search module uses an XSLT in Gsearch to index the **FOXML** documents in your repository, and allows you to configure search fields for searching and faceting. Whenever you add a new object in Fedora, the Solr module updates your index, and makes those results available to your users.

Solr “Out of the Box”

The Solr module comes with support for DC and MODS searching and faceting to support our solution packs. When Solr is installed, four additional blocks appear under administer>site building>blocks:

| Block | Region | Operations |
|--|---|---------------------------|
| Sidebar 1 | | |
|  Islandora Current Query | Sidebar 1  | configure |
|  Islandora Facet Block | Sidebar 1  | configure |
|  Islandora Solr Advanced Search | Sidebar 1  | configure |
|  Islandora Solr Simple Search | Sidebar 1  | configure |

Islandora Simple Search - provides a simple full-text search of all items in the repository. A term will appear, no matter where it appears in the metadata.

Islandora Advanced Search - provides a faceted search for users, where terms can be targeted to chosen metadata fields.

Islandora Current Query - when a user is viewing a set of search results by facet, this block will show the current filters being applied in the search.

Islandora Facet Block - this block will show users the facets they can use to explore search results.

By placing these blocks in regions of your site and configuring the Solr module under administer>Islandora Solr client, you can facilitate both full-text searching and faceted searching of items ingested using Solution Packs. The basic configuration of Solr can be modified to change the weight of search fields, and extend the out-of-the-box functionality. Information for doing this is provided in Chapter 7: Customizing Islandora.

Configuring Solr out-of-the Box

In order for Solr to work for your collections, you will have to activate the blocks that you want, and configure them to display your desired results. The following instructions will show you how to configure Solr to work with solution packs. Additional information about installing Solr and how Islandora uses Solr is provided in Chapter 9: Enabling Indexing/Searching with Solr. Additional information about customizing Islandora’s Solr Search Module is provided in Chapter 7: Customizing Islandora.

1. Navigate to the Configuration Panel

Start by going to the ‘Administer’ page and click on the ‘Islandora Solr Client’ link under ‘Site Configuration.’

Site configuration

Adjust basic site configuration options.

- Actions
- Administration theme
- Clean URLs
- Configure Islandora
- Date and time
- Error reporting
- File system
- Image toolkit
- ImageAPI
- Input formats
- **Islandora Solr Client**
- Logging and alerts
- Performance
- Site information
- Site maintenance
- Tabs

Click here to
configure Solr

2. Verify the Solr URL and Request Handler

The Solr URL should be *ip.address.of.site:port/solr*. So, for example, if you're using the Islandora Virtual Machine Image, your Solr URL should be *192.168.56.195:8080/solr*. If the URL is correct you will see a green check mark.

Islandora Solr Client

Solr url:

localhost:8080/solr

The url of the Solr installation. Defaults to localhost:8080/solr.

✓ Successfully connected to Solr server at localhost:8080/solr

Request Handler:

standard

Request handlers as defined by [solrconfig.xml](#)

Limit results to namespace:

Enter a namespace ie 'demo' to restrict results to PIDs within that namespace

note: this setting may only be used with 'standard' request handler.

Enter the Solr URL

Limit search results to a
particular namespace

Make sure the request handler is set to 'standard'. You can customize the request handler by editing *solrconfig.xml* to make other request handlers available. This assumes that you are using the default request handler, which will support all the metadata in our solution packs.

You can also restrict your search results to a particular namespace. This is useful if there are multiple sites using the same repository and you want to block search results from the other sites. Remember that the namespace is the first half of the PID – everything before the colon.

3. Set the Default and Secondary Display Profiles

You can modify the way results are displayed by configuring the default display profile. You can choose from Fields (default), Maps, and Simple Table. Switching back and forth between different display profiles is simple, so feel free to experiment and see which profile best suits your site.

Secondary display profiles provide optional secondary outputs for search results. Out-of-the-box, support for CSV output is provided. Selecting this option will place a CSV button next to your search results.

Default Display Profile:

☒ Fields (default)
☐ Maps
☐ Simple Table

Preferred normal display profile for search results. These may be provided by third-party modules.

Secondary Display Profiles:

☐ CSV

Enabled secondary output/download types for search results.

The Secondary Profiles that are appearing in the Solr Config screen are generated by the Solr Sample Configuration Module.

4. Configure Search Block Repetition

This simple option lets you choose the number of search fields that will appear in your search block. Each field may be able to search a number of categories (e.g. title, subject, etc.) so having more than one field allows the user to combine search terms easily.

5. Choose Search Terms

Here you can choose the search terms that will appear in the drop-down menus on the search block. Terms must be entered by their field names (in most cases this will be **Dublin Core**), though you can optionally specify a more human-readable label. To determine the appropriate syntax for your search terms, simply edit the metadata on any object in your repository (for instructions, see How to Edit an Object's Metadata in [Chapter 3 - Getting Started with Islandora](#)). Each field label will display the proper syntax for adding it to your Solr search configuration. A full list of the terms made available by the schema provided in the module package is provided in [APPENDIX D - SOLR SCHEMA \(SEARCH\) Term Reference](#). Note that you will want to use fields that have been indexed as "text" here.

```
dc.title ~ Title
dc.subject ~ Subject
PID ~ PID
type ~ Type
```

Search Terms:

dc.title ~ Title

Enter your search terms, one per line

Indicate what fields should appear in the dropdown menu of terms for the **Solr Advanced Search Block**. Enter terms on separate lines using the following pattern: field [tilde] preferred label. ie **dc.title ~ Title**

Review the schema.xml to see what terms are available. A list of the fields available when using the schema.xml Solr module is available in Islandora's online documentation.

6. Choose Facet Fields

Solr uses **faceting** to filter search results. Here, you can choose which fields you wish to allow faceting on. The format is the same as the search terms described above.

Facet Fields:

subject ~ Subject
date ~ Date
relation ~ Collection
title ~ Title

Enter facet fields,
one per line

Indicate which fields will appear in the **Islandora Facet Block**.
Enter terms on separate lines using the following pattern: field [tilde] preferred label. ie **dc.title ~ Title**

Note that you will want to use terms here that have been indexed as "strings" (see Appendix).

7. Define Labels for Your Search Results

You may not want to display every search-able field in the list of search results. Here, you can determine what labels will be used for fields in the list of results, and also which fields will display with search results. The format is the same as the search terms and facet fields described above. You can limit search results to this set of terms of selecting the 'Limit results to these fields?' box.

Labels for Returned Fields:

dc.subject ~ Subject
dc.type ~ Type

Enter labels for
returned fields,
one per line

Set labels for fields returned from query. Enter terms on separate lines using the following pattern: field [tilde] preferred label. ie **dc.title ~ Title**

Note that you will have to put the PID field in if you want a link back to the object from the search results. This means that configuring PID as one of the fields, particularly in your results is strongly recommended.

8. Minimum/Maximum Facet Fields and Results per Page

The final options let you control the minimum number of facets that must be available in order to display them to the user, and the maximum number of fields to display in the facet block. A negative maximum number will indicate unlimited fields. You can also limit the number of results to display per page.

Minimum Results to Display Facet:

2

Minimum facet count required for display

Set minimum number of facets required for display

Maximum Facets returned:

10

Set the number of terms that should be returned to
For example, if there are 100 possible subjects in a
Set the maximum number of terms displayed for the user

Results per page:

20

The number of results to show per page

☐ Debug Mode?
Dumps solr query to the screen for testing

Set the number of results per page

Save Solr Configuration Reset to defaults

You have successfully configured your Solr client.