

Chapter 4 - Search & 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 8 - Installing Solr and GSearch](#)), and you have installed and activated the Solr module. The xslt and solrschema.xml documents that come packaged with [GSearch](#) 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, six additional blocks appear under Structure>Blocks:

+	Islandora simple search	Second sidebar	configure
+	Islandora advanced search	Second sidebar	configure
+	Islandora query	Second sidebar	configure
+	Islandora facets	Second sidebar	configure
+	Islandora sort	Second sidebar	configure
+	Islandora displays	Second sidebar	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 configurable search for users, where specific metadata fields can be searched and combined with boolean operators.

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

Islandora displays - displays current search query

Islandora sort - displays sorting options for search results.

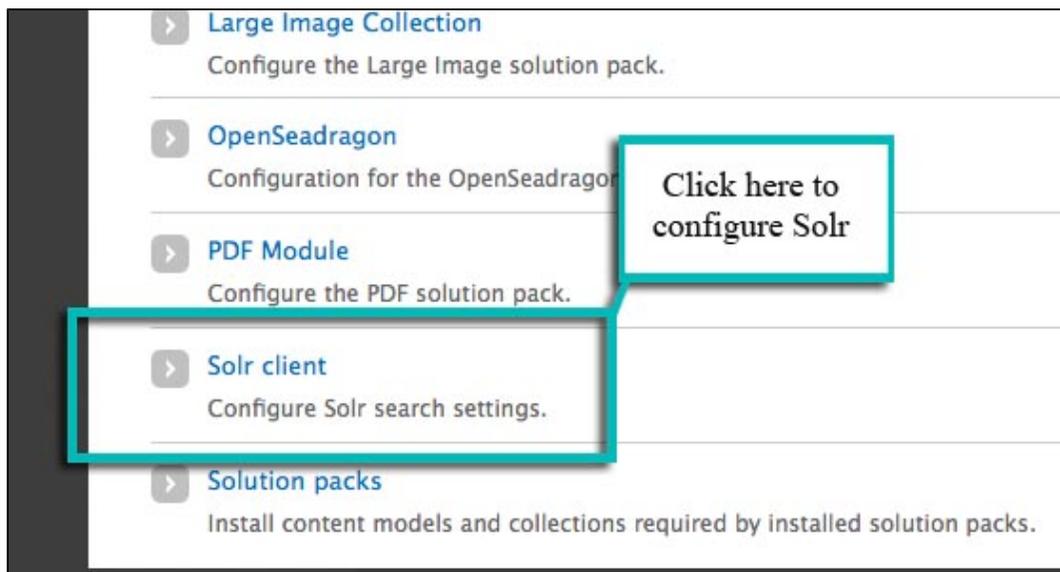
By placing these blocks in regions of your site and configuring the Solr module under Islandora>Solr client, you can facilitate both full-text searching and faceted searching of items ingested using Solution Packs. The basic configuration of Solr can be modified to change the weight of search fields, 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 - Installing Solr and GSearch](#). Additional information about customizing Islandora's Solr Search Module is provided in [Chapter 7 - Customizing Islandora](#).

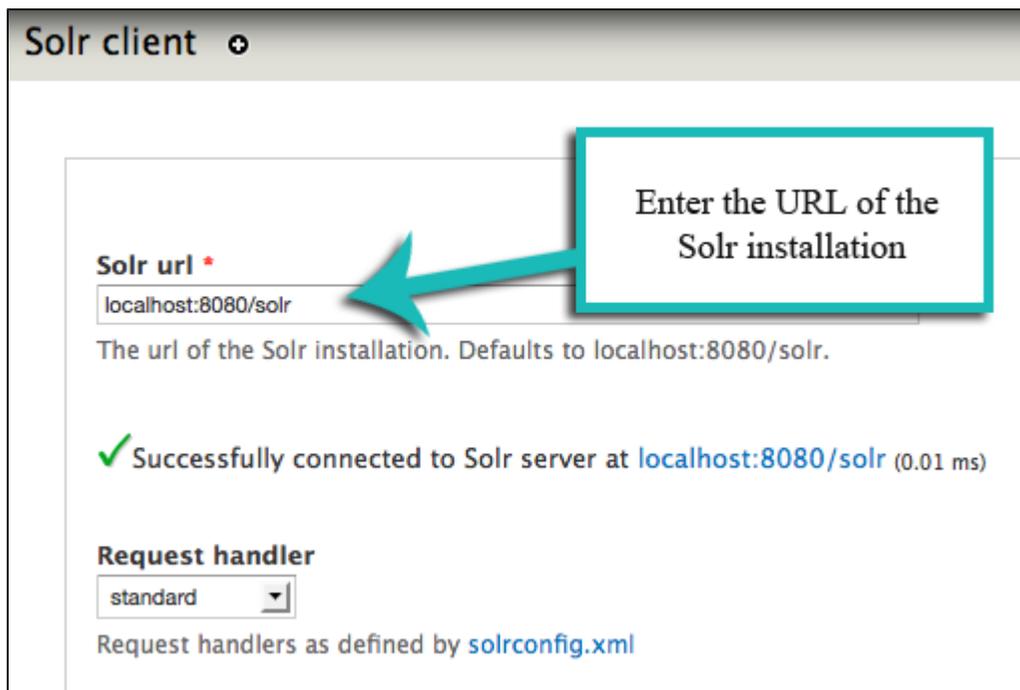
1. Navigate to the Configuration Panel

Start by going to the 'Islandora' page in the admin panel and click on the 'Solr Client' link.



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, or if using a local installation it should be localhost:8080/solr. If the URL is correct you will see a green check mark.



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.

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 List (default), Bookmark, Grid, and 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.

DISPLAY PROFILES

Primary display profiles [Show row weights](#)

DEFAULT	ENABLED	NAME	MACHINE READABLE NAME	CONFIGURATION
<input type="radio"/>	<input type="checkbox"/>	Bookmark	bookmark	
<input checked="" type="radio"/>	<input checked="" type="checkbox"/>	List (default)	default	
<input type="radio"/>	<input type="checkbox"/>	Grid	grid	
<input type="radio"/>	<input type="checkbox"/>	Table	table	

Preferred normal display profile for search results. Profiles are provided by third-party modules.

Secondary display profiles

CSV

Enabled secondary output/download types for search results.

Allow search results to be exported to CSV



4. Updating Default Display Settings

The default display settings are used to define the fields that will display for your search results.

Default display settings

- Sort settings
- Facet settings
- Advanced search block
- Query defaults
- Required Solr fields
- Other

Display fields

SOLR FIELD	SETTINGS
No fields defined	
Add another item	
<input type="text" value="dc.title"/>	<input type="text" value="(text_en_splitting)"/>
<input type="button" value="Add"/>	

Set labels for Solr fields to be included in the search results.

Limit results to fields listed above?
If checked, results displayed will be limited to the fields specified above. The order of the display fields is only enforced when this is enabled.
Note: some display profiles may not honour this value.

Results per page
20
The number of results to show per page.

Once you've added a field, there is a configuration option available for the field. You can enter a label for the field, adding linking, term highlighting and permissions.

Configure field: dc.title

Label
Title
A human-readable name.

Linking
 None
 Link this field to the object's page.
 Link the value to a Solr search result. (NOTE: Will likely break with very large values.)

Highlight
If a match is found on this field, the search term will be highlighted.
Note: Only text that has been both indexed and stored may be highlighted. While highlighting on non-tokenized fields is possible, the best results are achieved using tokenized fields. This checkbox may be grayed out if the Solr field cannot be highlighted.

▶ PERMISSIONS

Apply

Enter a label for the field.
You can link the field to the object.

Continue adding and configuring fields. You will want to limit your display to the fields you have defined. Ensure that you save the configuration.

The screenshot shows the Solr configuration interface for display settings. On the left is a sidebar with navigation options: Default display settings, Sort settings, Facet settings, Advanced search block, Query defaults, Required Solr fields, and Other. The main area is titled 'Display fields' and contains a table with columns 'SOLR FIELD', 'SETTINGS', and 'OPERATIONS'. A table entry shows 'dc.title' with 'Label: Title' and 'configure remove' links. Below the table is an 'Add another item' button and a text input field. A callout box points to the 'Add another item' button with the text: 'Add more fields using the same method.' Below the table, there is a section 'Set labels for Solr fields to be included in the search results.' with a checkbox 'Limit results to fields listed above?' and a text input field for 'Results per page' set to '20'. A callout box points to these two elements with the text: 'You should limit the results to the fields you've defined. You can also set the number of results to display per page.' At the bottom, there are two buttons: 'Save Solr configuration' and 'Reset to defaults'. A callout box points to the 'Save Solr configuration' button with the text: 'Don't forget to **Save**.'

5. Choose Search Terms in Advanced Search Block

Here you can choose the search terms that will appear in the drop-down menus on the advanced 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 with the **configure** option. 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.

Default display settings

Sort settings

Facet settings

Advanced search block

Query defaults

Other

Search

Public Core field names

Human readable label

Show row weights

SOLR FIELD	SETTINGS	OPERATIONS
+	dc.title	Label: Title configure remove
+	dc.subject	configure remove

Add another item

Add

Configure labels and permissions

Indicate what fields should appear in the dropdown menu of terms for the Advanced Search Block.
Note: It is recommended to use tokenized fields.

Default boolean operator

User configurable

AND

OR

Select a default boolean operator for the search query. User configurable exposes a dropdown which gives the user the choice between AND, OR and NOT.

You also have the option of setting permissions on a per-field basis, allowing only certain subsets of users access to search across different facets. These permissions, and the human-readable label for each Solr field, can be configured by clicking **configure** to open a new options window:

Configure field: dc.title

Label
Title
A human readable name.

▼ **PERMISSIONS**

Permissions

- anonymous user
- authenticated user
- administrator

Select what role can access this field.
Note: checkboxes may be grayed out for roles who do not have permission to search the Solr index.

Apply

Solr field permissions are dependent on a role having Drupal permissions to search the Solr index. Roles without this permissions may appear in this list, but they will not be selectable.

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. You can also use this screen to configure:

1. **Minimum Limit** - The minimum number of search results returned for a particular facet before that facet will be displayed. For instance, if the limit was set to '3' and a

search for "fish" returned only two results for 'Bass', 'Bass' would not be included as a facet.

2. **Soft Limit** - The number of facets to show when a search is first returned. This setting will return the most populous facets first, and include a button to expose more available facets.
3. **Maximum Limit** - Similar to a Soft Limit, but without the option to expand to show more terms beyond the limit set here.

The screenshot shows the Solr Facet settings interface. On the left is a sidebar with navigation links: Default display settings, Sort settings, Facet settings (highlighted with a red box), Advanced search block, Query defaults, and Other. The main content area is titled 'Facet fields' and includes a 'Show row weights' link. Below this is a table with columns 'SOLR FIELD', 'SETTINGS', and 'OPERATIONS'. The table contains one row for 'dc.title' with 'configure' and 'remove' links. Below the table is an 'Add another item' input field and an 'Add' button. A note below the table states: 'Indicate what fields should appear in as facets. Note: it is recommended to use non-tokenized solr fields (full literal strings)'. Three red boxes with numbers 1, 2, and 3 are overlaid on the page, pointing to the 'Minimum limit', 'Soft limit', and 'Maximum limit' settings respectively.

SOLR FIELD	SETTINGS	OPERATIONS
✚ dc.title		configure remove

1 **Minimum limit**
2
Minimum results required to display a facet

2 **Soft limit**
10
The number which should be displayed initially. If there are more, then the a "Show more" button will allow the rest up to the value below to be displayed. Use 0 to disable.

3 **Maximum limit**
20
Set the maximum number of terms that should be returned to the user. For example, if there are 100 possible subjects in a faceted result, you may wish to return only the top 10.

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

7. Choose Sort Fields

Using the same field formatting as Advanced Search and Faceting fields, you can select fields to make available to the user to sort their search results in the Sort block.

Default display settings

Sort settings

Facet settings

Advanced search block

Query defaults

Other

Sort fields

Show row weights

SOLR FIELD	SETTINGS	OPERATIONS
✚ dc.title		configure remove
✚ dc.subject		configure remove

Add another item

Add

Indicate what fields should appear in the Islandora sort block. To sort on relevancy, use the 'score' field.
Note: not all fields are sortable. For more info check the [Solr documentation](#).

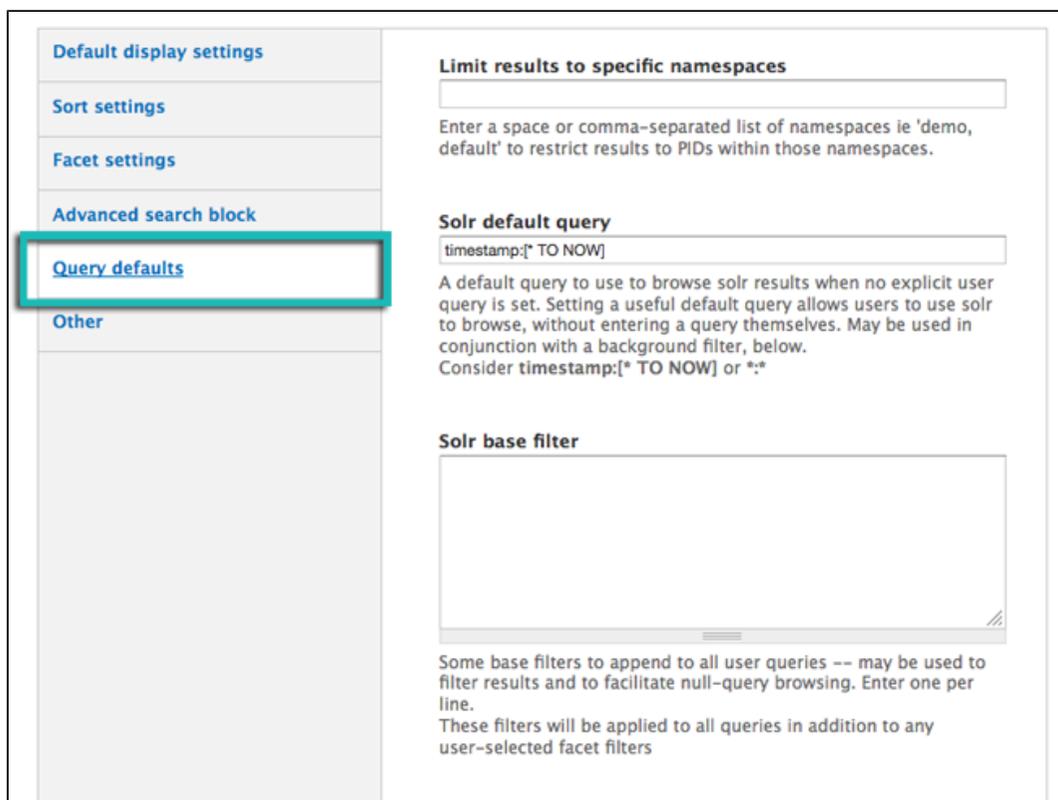
The field for a relevance based sort is "score"

8. Set Query Defaults

This section of the configuration panel provides advanced Solr query customization options:

- **Limit results to specific namespaces-** 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.

- **Solr Default Query** - This option allows you to specify a default query used to browse results when no explicit query has been entered. For example, if a user runs a search and then deletes their search term from the breadcrumbs, this default query will be applied in its place.
- **Solr Base Filter** - You can use this option as a blanket way to filter all Solr search queries. For example, you can apply date-based or collection-based restrictions. For more information on Solr, see the [Solr Reference Guide](#).



The screenshot displays the Solr configuration interface. On the left, a sidebar contains several menu items: 'Default display settings', 'Sort settings', 'Facet settings', 'Advanced search block', 'Query defaults' (highlighted with a red box), and 'Other'. The main content area is divided into three sections:

- Limit results to specific namespaces**: A text input field is empty. Below it, a note states: "Enter a space or comma-separated list of namespaces ie 'demo, default' to restrict results to PIDs within those namespaces."
- Solr default query**: A text input field contains the query `timestamp:[* TO NOW]`. Below it, a note explains: "A default query to use to browse solr results when no explicit user query is set. Setting a useful default query allows users to use solr to browse, without entering a query themselves. May be used in conjunction with a background filter, below. Consider `timestamp:[* TO NOW]` or `*:*`"
- Solr base filter**: A large, empty text area for entering filters. Below it, a note states: "Some base filters to append to all user queries -- may be used to filter results and to facilitate null-query browsing. Enter one per line. These filters will be applied to all queries in addition to any user-selected facet filters"

You have successfully configured your Solr client.