ReIndexing Content (for Browse or Search)

Overview

DSpace offers two options to index content for Browsing & Searching:

1. Traditional Browse & Search (via Lucene & Database indexes) - this is enabled by default
2. Faceted/Filtered Browse & Search (via Solr & DSpace Discovery) - available for XMLUI only and disabled by default

This particular page only describes the "Traditional Browse & Search" indexing processes. For more information on Faceted/Filtered Browse & Search, please see DSpace Discovery, in particular Discovery SOLR Index Maintenance.

Creating the Browse & Search Indexes

To create (or recreate) all the various browse/search indexes that you define in the Configuration Section there are a variety of options available to you. You can see these options below in the command table.

Command used:	`[dspace]/bin/dspace index-init`
Java class:	org.dspace.browse.IndexBrowse
Arguments short and long forms):	Description
`-r` or `-rebuild`	Should we rebuild all the indexes, which removes old tables and creates new ones. For use with `-f`. Mutually exclusive with `-d`
`-s` or `-start`	`-s <int>` start from this index number and work upwards (mostly only useful for debugging). For use with `-t` and `-f`
`-x` or `-execute`	Execute all the remove and create SQL against the database. For use with `-t` and `-f`
`-i` or `-index`	Actually do the indexing. Mutually exclusive with `-t` and `-f`.
`-o` or `-out`	`-o <filename>` write the remove and create SQL to the given file. For use with `-t` and `-f`
`-p` or `-print`	Write the remove and create SQL to the stdout. For use with `-t` and `-f`.
`-t` or `-tables`	Create the tables only, do no attempt to index. Mutually exclusive with `-f` and `-i`
`-f` or `-full`	Make the tables, and do the indexing. This forces `-x`. Mutually exclusive with `-f` and `-i`.
`-v` or `-verbose`	Print extra information to the stdout. If used in conjunction with `-p`, you cannot use the stdout to generate your database structure.
`-d` or `-delete`	Delete all the indexes, but do not create new ones. For use with `-f`. This is mutually exclusive with `-r`.
`-h` or `-help`	Show this help documentation. Overrides all other arguments.

Running the Indexing Programs

Complete Index Regeneration

Requires that you stop Tomcat first

Because this command actually **deletes** existing Browse Index tables, you **must** stop Tomcat (or your Servlet Container of choice) before executing `index-init`. After the indexing command completes, you can restart Tomcat.

By running `[dspace]/bin/dspace index-init` you will completely regenerate your indexes, tearing down all existing tables and reconstructing with the new configuration.

```
[dspace]/bin/dspace index-init
```

Updating the Indexes

By running `[dspace]/bin/dspace index-update` you will reindex your full browse & search indexes without modifying the DSpace table structure. (This should be your default approach if indexing, for example, via a cron job periodically). Because it does not "tear down" the existing tables, this command can be run while DSpace (and Tomcat or similar) is still running.

```
[dspace]/bin/dspace index-update
```

Destroy and Rebuild Browse Tables

You can destroy and rebuild the database, but do not do the indexing. Output the SQL to do this to the screen and a file, as well as executing it against the database, while being verbose. **WARNING: This is not really recommended unless you know what you are doing.**

At the CLI screen:

```
[dspace]/bin/dspace index \-r \-t \-p \-v \-x \-o myfile.sql
```

Indexing Customization

Browse Index Customization

DSpace provides robust browse indexing. It is possible to expand upon the default indexes delivered at the time of the installation. The System Administrator should review Browse Index Configuration to become familiar with the property keys and the definitions used therein before attempting heavy customizations.

Through customization is is possible to:

- Add new browse indexes besides the four that are delivered upon installation. Examples:
    - Series
    - Specific subject fields (Library of Congress Subject Headings). *(It is possible to create a browse index based on a controlled vocabulary or thesaurus.)*
    - Other metadata schema fields
- Combine metadata fields into one browse
- Combine different metadata schemas in one browse

Examples of new browse indexes that are possible. *(The system administrator is reminded to read the section on Browse Index Configuration)*

- **Add a Series Browse**. You want to add a new browse using a previously unused metadata element.
    - `webui.browse.index.6 = series:metadata:dc.relation.ispartofseries:text:single`
    - Note: the index # need to be adjusted to your browse stanza in the _dspace.cfg_ file. Also, you will need to update your *Messages. properties* file.
- **Combine more than one metadata field into a browse.** You may have other title fields used in your repository. You may only want one or two of them added, not all title fields. And/or you may want your series to file in there.
    - `webui.browse.index.3 = title:metadata:dc.title,dc:title.uniform,dc:relation.ispartofseries:title:full`
- **Separate subject browse.** You may want to have a separate subject browse limited to only one type of subject.
    - `webui.browse.index.7 = lcsubject.metdata:dc.subject.lcsh.text:single`

As one can see, the choices are limited only by your metadata schema, the metadata, and your imagination.

Because Browse Indexes are stored in database tables, remember to run `index-init` after adding any new definitions in the `dspace.cfg` to have the indexes created and the data indexed.

Search Index Customization

For information about configuring new Search Indexes, please refer to Configuring Lucene Search Indexes.