

# Lab 2: Adding metadata—and using it

## 2.1. A collection of Word and PDF files—Part B

*In the Librarian Interface, open up the reports collection that you created in exercise 1.5. Remember that the extracted Title metadata for some documents was incorrect.*

### *Manually adding metadata to documents in a collection*

8. In the **Enrich** panel, manually add Dublin Core **dc.Title** metadata to those documents which have incorrect **ex.Title** metadata. Select *word03.doc* and double-click to open it. Copy the title of this document ("Greenstone: A comprehensive open-source digital library software system") and return to the Librarian Interface. Scroll up or down in the metadata table until you can see **dc.Title**. Click in the value box and paste in the metadata.
9. Now add **dc.Creator** information for the same document. You can add more than one value for the same field: when you press **Enter** in a metadata value field, a new empty field of the same type will be generated. Add each author separately as **dc.Creator** metadata.
10. Close the document (in Microsoft Word) when you have finished copying metadata from it. External programs opened when viewing documents must be closed before building the collection, otherwise errors can occur.
11. Next add **dc.Title** and **dc.Creator** metadata for a few of the other documents.
12. You will notice as you add more values, they appear in the **Existing values for ...** box below the metadata table. If you are adding the same metadata value to more than one document, you can select it from this list. For example, *pdf01.pdf* and *word03.doc* share the same Title; and many documents have common authors.

*If you build and preview your collection at this point, you will see that the **Titles** list now shows your new Titles. However, the **dc.Creator** metadata is not displayed. You need to alter the collection design to use this metadata.*

### *Document Plugins*

13. In the Librarian Interface, look at the **Document Plugins** section of the **Design** panel, by clicking on this in the list to the left. Here you can add, configure or remove plugins to be used in the collection. There is no need to remove any plugins, but it will speed up processing a little. In this case we have only Word, PDF, RTF, and PostScript documents, and can remove the **ZIPPlug**, **TEXTPlug**, **HTMLPlug**, **EMAILPlug**, **ImagePlug**, **ISISPlug** and **NULPlug** plugins. To delete a plugin, select it and click **<Remove Plugin>**. **GAPlug** and **MetadataXMLPlug** are required for any type of source collection and should not be removed.

### *Search indexes*

14. The next step in the **Design** panel is **Search Indexes**. These specify what parts of the collection are searchable (e.g. searching by title and author). Delete the **ex.Source** index, which is not particularly useful, by selecting it and clicking **<Remove Index>**.
15. Modify the **ex.Title** index to include **dc.Title** by selecting the index in the **Assigned Indexes** box and clicking **<Edit Index>**. Select **dc.Title** from the list of metadata, and click **<Replace Index>**. Searching this index will search both **dc.Title** and **ex.Title** metadata. If you want to restrict searching to just the manually added **dc.Title** metadata, edit the index again and deselect **ex.Title** from the list of metadata.

16. You can add indexes based on any metadata. Add a new index based on **dc.Creator** by clicking **<New Index>**. Select **dc.Creator** in the list of metadata, and click **<Add Index>**.

The next section is **Partition Indexes**. In this exercise, we will not make any changes to this.

### Browsing classifiers

17. The **Browsing Classifiers** section adds "classifiers," which provide the collection with browsing functions. Go to this section and observe that Greenstone has provided two classifiers, *AZLists* based on **ex.Title** and **ex.Source** metadata. These correspond to the *Titles* and *Filenames* buttons on the collection's access bar.

Remove the **ex.Source** classifier by selecting it and clicking **<Remove Classifier>**.

18. Modify the **ex.Title** classifier to use **dc.Title** instead. Select the classifier and click **<Configure Classifier...>**. In the **metadata** box, select **dc.Title** instead of **ex.Title**. Click **<OK>**.

19. Now add an **AZCompactList** classifier for **dc.Creator**. Select **AZCompactList** from the **Select classifier to add:** drop-down list and click **<Add Classifier...>**. A popup window **Configuring Arguments** appears. Select **dc.Creator** from the **metadata** drop-down list and click **<OK>**.

**AZCompactList** is like **AZList**, except that values that appear multiple times in the hierarchy are automatically grouped together and a new node, shown as a bookshelf icon, is formed.

20. Switch to the **Create** panel, and **build** and **preview** the collection.
21. Check that all the facilities work properly. There should be three full-text indexes, called *text*, *dc.Title*, and *dc.Creator*. The *Titles* list should display all the documents to which you have assigned **dc.Title** metadata (and only those documents). The *Creators* list should show one bookshelf for each author you have assigned as **dc.Creator**, and clicking on that bookshelf should take you to all the documents they authored.

### Renaming the search indexes

22. The default display text for the indexes in the drop-down list on the search page contains the content of the index. Now we will change this display text to make it nicer. Go to the **Format** panel by clicking its tab. This panel is split into several sections, each controlling some aspect of collection presentation.
23. Select **Search** in the left hand list. This section allows you to modify what text is displayed for the drop-down lists in the search form (indexes, subcollections, levels etc). Set the **Display text** for the **dc.Title,Title** index to be "titles", and that for the **dc.Creator** index to be "creators". Preview the collection by clicking the **Preview Collection**. The search form should display the new text.

### Classifying on multiple metadata

24. The new *Titles* list shows only those documents which have been assigned **dc.Title** metadata. For many documents, extracted Titles may be fine, and it is impractical to add the same metadata again as **dc.Title**. Fortunately there is a way we can use both metadata types in one classifier: specify a list of metadata names in the classifier.
25. In the **Browsing Classifiers** section of the **Design** panel, select the **AZList** for **dc.Title** in the **Assigned Classifiers** box and click **<Configure Classifier...>**. Note you can achieve the same result by double clicking on the classifier.
26. In the **metadata** field, type ",ex.Title" after the "dc.Title"—i.e. make it read

```
dc.Title,ex.Title
```

27. If you have already done the **Enhanced Word document handling** exercise, some of the documents will have extracted **ex.Creator** metadata, and some will have **dc.Creator**. To use both of these in the

Creators classifier, make a similar change to the **AZCompactList**: make the **metadata** field read `dc.Creator,ex.Creator`.

You may notice that **AZCompactList** has two options after the **metadata** option: **firstvalueonly** and **allvalues**. Manually added metadata can be used to replace or enhance automatically extracted metadata, and these options control exactly which pieces of metadata a document is classified by.

For example, say we have two documents. Document 1 has four Creators specified (`dc.Creator = dcA`, `dc.Creator = dcB`, `ex.Creator = exA`, `ex.Creator = exB`), while document 2 has three (`ex.Creator = exA`, `ex.Creator = exB`, `ex.Creator = exC`). The following table shows which metadata values each document is classified by, for the different classifier options:

<u>AZCompactList options</u>	<u>Document 1</u>	<u>Document 2</u>
-metadata dc.Creator,ex.Creator	dcA, dcB	exA, exB, exC
-metadata dc.Creator,ex.Creator -firstvalueonly	dcA	exA
-metadata dc.Creator,ex.Creator -allvalues	dcA, dcB, exA, exB	exA, exB, exC

28. **Build** the collection again and **preview** it. Now all of the documents should appear in the *Titles* list (and extracted Creators should appear in the *Creators* list).

Extracted metadata is unreliable. But it is very cheap! On the other hand, manually assigned metadata is reliable, but expensive. The previous section of this exercise has shown how to aim for the best of both worlds by using extracted metadata but correcting it when it is wrong. While this may not satisfy the professional librarian, it could provide a useful compromise for the music teacher who wants to get their collection together with a minimum of effort.

### ***Branding a collection with an image***

29. Switch back to the **Format** panel. The first section **General** appears. This allows you to modify the values you provided when defining the collection, if desired. You can also brand the collection using a suitable image.
30. Click on the **<Browse...>** button associated with **URL to 'about page' image:**, and browse to the image `sample_files Word_and_PDF wrdpdf.gif` on your computer. When you select this image, Greenstone automatically generates an appropriate URL for the image. **Preview** the collection: you should see the new image at the top left of the page.

## 2.2. A simple image collection

1. In the Librarian Interface, start a new collection (**File** → **New...**) called **backdrop**. Fill out the fields with appropriate information. For **Base this collection on:**, select the item **Simple image collection (image-e)** from the pull-down menu.

*When you base a collection on an existing one, it inherits all the settings of the old one, including which metadata sets (if any) the collection uses.*

2. Copy the images provided in `sample_files/images` into your newly-formed collection.
3. Change to the **Create** panel and **build** the collection.
4. **Preview** the result.
5. Click on **Browse** in the navigation bar to view a list of the photos ordered by filename and presented as a thumbnail accompanied by some basic data about the image. The structure of this collection is the same as **Simple image collection (image-e)**, but the content is different.
6. Back in the Librarian Interface, change to the **Enrich** panel and view the extracted metadata for *Bear.jpg*.

### *Adding Title and Description metadata*

7. We work with just the first three files (*Bear.jpg*, *Cat.jpg* and *Cheetah.jpg*) to get a flavour of what is possible. First, set each file's **dc.Title** field to be the same as its filename but without the filename extension:

Click on *Bear.jpg* so its metadata fields are available, then click on its **dc.Title** field on the right-hand side. Type in **Bear**.

Repeat the process for *Cat.jpg* and *Cheetah.jpg*.

8. Add a description for each image as **dc.Description** metadata.

What description should you enter? To remind yourself of a file's content, the Librarian Interface lets you open files by double-clicking them. It launches the appropriate application based on the filename extension, Word for .doc files, Acrobat for .pdf files and so on.

Double-click *Bear.jpg*: on Windows, the image will normally be displayed by Microsoft's Photo Editor (although this depends on how your computer has been set up).

Back in the **Enrich** pane, make sure that *Bear.jpg* is selected in the collection tree on the left hand side. Enter the text **Bear in the Rocky Mountains** as the value for the **dc.Description** field.

Repeat this process for *Cat.jpg* and *Cheetah.jpg*, adding a suitable description for each.

9. Go to the **Create** panel and click **<Build Collection>**. Once it has finished building, **preview** the collection. You will not notice anything new. That's because we haven't changed the design of the collection to take advantage of the new metadata.

### *Change Format Features to display new metadata*

10. Now we customize the collection's appearance. Go to the **Format** panel and select **Format Features** from the left-hand list. Leave the feature selection controls at their default values, so that **All Features** is selected for **Choose Feature**, and **VList** is selected as the **Affected Component**. In the **HTML Format String**, edit the text as follows:

- Change `_ImageName_:` to `Title:`
- Change `[Image]` to `[dc.Title]`
- After `[dc.Title]<br>` add `Description: [dc.Description]<br>`

*Metadata names are case-sensitive in Greenstone: it is important that you capitalize "Title" and "Description" (and don't capitalize "dc").*

11. The new format statement is displayed in the list of assigned format statements. The first substitution alters the fragment of text that appears to the right of the thumbnail image, the second alters the item of metadata that follows it. The addition displays the description after the Title.
12. Preview the collection by clicking the **<Preview Collection>** button. When you click on **Browse** in the navigation bar the presentation has changed to "Title: Bear" and so on. Each image's description should appear beside the thumbnail, following the title.

*After the first three items, the Title and Description become blank because we have only assigned Dublin Core metadata to these first three. To get a full listing, enter all the metadata.*

*Changes in the **Format** panel take place immediately and you can see the result straightaway by clicking **reload** (or **refresh**) in the web browser. If you modify anything in the **Gather**, **Enrich** or **Design** panels, you will need to rebuild the collection.*

### **Changing the size of image thumbnails**

13. Lets change the size of the thumbnail image and make it smaller. Thumbnail images are created by the **ImagePlug** plug-in, so we need to access its configuration settings. To do this, switch to the **Design** panel and select **Document Plugins** from the list on the left. Double-click **ImagePlug** to pop up a window that shows its settings. (Alternatively, select **ImagePlug** with a single click and then click **<Configure Plugin...>** further down the screen). Currently all options are off, so standard defaults are used. Select **thumbnailsize**, set it to **50**, and click **<OK>**.
14. **Build** and **preview** the collection.
15. Once you have seen the result of the change, return to the **Design** panel, select the configuration options for **ImagePlug**, and switch the **thumbnailsize** option off so that the thumbnail reverts to its normal size when the collection is re-built.

### **Adding a browsing classifier based on Description metadata**

16. Now we'll add a new browsing option based on the descriptions. In the **Design** panel, select **Browsing Classifiers** from the left-hand list. Set the menu item for **Select classifier to add:** to **AZList**; then click **<Add Classifier...>**.
17. A window pops up to control the classifier's options. Set the **metadata** option to **dc.Description** and click **<OK>**.
18. **Build** the collection, and **preview** it. Choose the new **Descriptions** link that appears in the navigation bar.

*Only three items are shown, because only items with the relevant metadata (**dc.Description** in this case) appear in the list. The original browse list includes all photos in the collection because it is based on **ex.Image**, extracted metadata that reflects an image's filename, which is set for all images in the collection.*

### **Creating a searchable index based on Description metadata**

19. Now we'll add an index so that the collection can be searched by descriptions. Switch to the **Design** panel and select **Search Indexes** from the left-hand list. Click the **<New Index>** button. Select **dc.Description** from the list of metadata to include in the index, leave **Indexing level:** at its default, "document", and click **<Add Index>**.

20. Switch to the **Create** panel, **build** the collection, then **preview** it. There is now a **Search** button in the navigation bar. As an example, search for the term "bear" in the *document:dc.Description* index (which is the only index at this point).
21. To change the text that is displayed for the index (document:dc.Description), go to the **Format** panel back in the Librarian Interface. Select **Search** from the left-hand list. This panel allows you to change the text that is displayed on the search form. Change the **Display text** for the **document:dc.Description** index to "descriptions" (or other suitable text). Go back to the browser and reload the search page. Your new text will appear in the search form.

## 2.3. Enhanced collection of HTML files—Tudor

We return to the Tudor collection and add metadata that expresses a subject hierarchy. Then we build a classifier that exploits it by allowing readers to browse the documents about Monarchs, Relatives, Citizens, and Others separately.

### Adding hierarchically-structured metadata and a Hierarchy classifier

1. Open up your **tudor** collection (the original version, not the **webtudor** version), switch to the **Enrich** panel and select the *citizens* folder (a subfolder of *englishhistory.net tudor* ). Set its **dc.Subject and Keywords** metadata to **Tudor period|Citizens**. The vertical bar ("|") is a hierarchy marker. Selecting a *folder* and adding metadata has the effect of setting this metadata value for all files contained in this folder, its subfolders, and so on. A popup alerts you to this fact. Click **<OK>** to close the popup.
2. Repeat for the *monarchs* and *relative* folders, setting their **dc.Subject and Keywords** metadata to **Tudor period|Monarchs** and **Tudor period|Relatives** respectively. Note that the hierarchy appears in the **Existing values for dc.Subject and Keywords** area.

If you don't want to see the popup each time you add folder level metadata, tick the **Do not show this warning again** checkbox; it won't be displayed again.

3. Finally, select all remaining files—the ones that are not in the *citizens*, *monarchs*, or *relative* folders—by selecting the first and shift-clicking the last. Set their **dc.Subject and Keywords** metadata to **Tudor period|Others**: this is done in a single operation (there is a short delay before it completes).

When multiple files are selected in the left hand collection tree, all metadata values for all files are shown on the right hand side. Items that are common to all files are displayed in black—e.g. **dc.Subject and Keywords**—while others that pertain to only one or some of the files are displayed in grey—e.g. any extracted metadata.

Metadata inherited from a parent folder is indicated by a folder icon to the left of the metadata name. Select one of the files in the *relative* folder to see this.

4. Switch to the **Design** panel and select **Browsing Classifiers** from the left-hand list. Set the menu item for **Select classifier to add:** to **Hierarchy**; then click **<Add Classifier...>**.
5. A window pops up to control the classifier's options. Change the **metadata** to **dc.Subject and Keywords** and then click **<OK>**.
6. For tidiness' sake, **remove** the **classifier** for **Source** metadata (included by default) from the list of currently assigned classifiers, because this adds little to the collection.
7. Now switch to the **Create** panel, **build** the collection, and **preview** it. Choose the new **Subjects** link that appears in the navigation bar, and click the bookshelves to navigate around the four-entry hierarchy that you have created.

### Adding a hierarchical phrase browser (PHIND)

Next we'll add an interactive hierarchical phrase browsing classifier to this collection.

8. Switch to the **Design** panel and choose the **Browsing Classifiers** item from the left-hand list.
9. Choose **Phind** from the **Select classifier to add:** menu. Click **<Add Classifier...>**. A window pops asking for configuration options: leave the values at their preset defaults (this will base the phrase index on the full text) and click **<OK>**.
10. **Build** the collection again, **preview** it, and try out the new **Phrases** option in the navigation bar. An

interesting PHIND search term for this collection is "king". Note that even though it is called a phrase browser, only single terms can be used as the starting point for browsing.

### ***Partitioning the full-text index based on metadata values***

*Next we partition the full-text index into four separate pieces. To do this we first define four subcollections obtained by "filtering" the documents according to a criterion based on their **dc.Subject and Keywords** metadata. Then an index is assigned to each subcollection. This will enable users to restrict a search to a subset of the documents.*

11. Switch to the **Design** panel, and click **Partition Indexes**. This feature is disabled because you are operating in **Librarian** mode (this is indicated in the title bar at the top of the window).
12. Switch to **Library Systems Specialist** mode by going to **Preferences...** (on the **File** menu) and clicking **<Mode>**. Read about the other modes too.
13. Return to the **Partition Indexes** section of the **Design** panel. Ensure that the **Define Filters** tab is selected (the default). Define a subcollection filter with name **monarchs** that matches against **dc.Subject and Keywords**, and type **Monarchs** as the regular expression to match with. Click **<Add Filter>**. This filter includes any file whose **dc.Subject and Keywords** metadata contains the word *Monarchs*.
14. Define another filter, **relatives**, which matches **dc.Subject and Keywords** against the word **Relatives**. Define a third and fourth, **citizens** and **others**, which matches it against the words **Citizens** and **Others** respectively.
15. Having defined the subcollection filters, we partition the index into corresponding parts. Click the **Assign Partitions** tab. Select the citizens subcollection and click **<Add Partition>**. Next select monarchs, and click **<Add Partition>**. Repeat for the other two subcollections, so that you end up with four partitions, one based on each subcollection filter.

The order they appear in the **Assigned Subcollection Partitions** list is the order they will appear in the drop down menu on the search page. You can change the order by using the **<Move Up>** and **<Move Down>** buttons.

16. **Build** and **preview** the collection.
17. The search page includes a pulldown menu that allows you to select one of these partitions for searching. For example, try searching the *relatives* partition for *mary* and then search the *monarchs* partition for the same thing.
18. To allow users to search the collection as a whole as well as each subcollection individually, return to the **Partition Indexes** section of the **Design** panel and select the **Assign Partitions** tab. Select all four subcollections by checking their boxes and click **<Add Partition>**.
19. To ensure that the combined index appears first in the list on the reader's web page, use the **<Move Up>** button to get it to the top of the list here in the **Design** panel. Then **build** and **preview** the collection.
20. Search for a common term (like *the*) in all five index partitions, and check that the numbers of words (not documents) add up.
21. The text in the drop down box on the search page is based on the filters each partition was built on. To change the text that is displayed, go to the **Search** section of the **Format** panel. The single filter partitions have sensible default text, but the combined partition does not. Set the **Display text** for the combined partition to "all". **Preview** the collection.
22. In the Librarian Interface, return to **Librarian** mode, using **Preferences...** (on the **File** menu).

### ***Controlling the building process***

*Finally we look at how the building process can be controlled. Developing a new collection usually involves numerous cycles of building, previewing, adjusting some enrich and design features, and so on. While prototyping, it is best to temporarily reduce the number of documents in the collection. This can be accomplished through the **maxdocs** parameter to the building process.*

23. Switch to the **Create** panel and view the options that are displayed in the top portion of the screen. Select **maxdocs** and set its numeric counter to **3**. Now **build**.
24. Preview the newly rebuilt collection's **Titles** page. Previously this listed more than a dozen pages per letter of the alphabet, but now there are just three—the first three files encountered by the building process.
25. Go back to the **Create** panel and turn off the **maxdocs** option. **Rebuild** the collection so that all the documents are included.

## 2.4. Bibliographic collection—Part A

*This exercise looks at using fielded searching in a collection. Fielded searching is best used for metadata rich collections. Here we use bibliographic data in MARC format. We also "explode" the database, enabling editing of the metadata with the Librarian Interface.*

1. Start a new collection called **Papers Bibliography** which will contain a collection of example MARC records of the working papers published at the [Computer Science Department](#), [Waikato University](#). Enter the requested information and base it on **-- New Collection --**.
2. In the **Gather** panel, open the `sample_files marc` folder, drag `CMSwp-all.marc` into the right-hand pane and drop it there. A popup window asks whether you want to add **MARCPlug** to the collection to process this file. Click **<Add Plugin>**, because this plugin will be needed to process the MARC records.
3. In the **Document Plugins** section of the **Design** panel, remove the plugins **TextPlug** to **NULPlug** by selecting each one in the **Assigned Plugins** list and clicking **<Remove Plugin>** (**ZIPPlug**, **GAPlug** and **MARCPlug** remain). It is not strictly necessary to remove these redundant plugins, but it is good practice to include only plugins that are needed, to avoid unwanted (and unexpected) side effects.
4. Now select **Browsing Classifiers** within the **Design** panel and **remove** the default classifier for **Source** metadata.
5. In the **Search Indexes** section, **remove** the **ex.Source** index. In this collection all records are from the same file, so **ex.Source** metadata, which is set to the filename, is not particularly interesting or useful.
6. Switch to the **Create** panel, **build** the collection, and **preview** it. Browse through the *Titles* and view a record or two. Try searching—for example, find items that include **graphics**.
7. Back in the Librarian Interface, go to the **Browsing Classifiers** section of the **Design** panel. Select **AZCompactList** from the **Select classifier to add:** drop down menu, and click **<Add Classifier...>**. In the popup window, select **ex.Subject** as the metadata item. Click **<OK>**.

*AZCompactList is like AZList, except that terms that appear multiple times in the hierarchy are automatically grouped together and a new node, shown as a bookshelf icon, is formed.*

8. **Build** the collection and **preview** the result.

### Using fielded searching

9. Collections built with MGPP (the default indexer) provide the option of fielded searching. In the browser, go to the *PREFERENCES* page. You will notice that there is a **Query style:** option which enables you to switch between "normal" and "fielded" search. Change to fielded search now and click on the **Search** button. The search form has changed to a fielded form.
10. You can specify which search form types are available for a particular collection, and which one is the default, using the **SearchTypes** format statement. In the **Format** panel select **Format Features** from the left-hand list. Select the **SearchTypes** format statement from the list of assigned formats, and set the contents to **form**. This will make only fielded searching available for this collection.

*Search type options include **form** and **plain**. You can specify one or both separated by a comma. If both are specified, the first one is used as the default: this is the one that the user will see when they first enter the collection.*

11. **Preview** the collection again. Notice that the collection's home page no longer includes a query box. (This is because the search form is too big to fit here nicely.) To search, you have to click **Search** in the navigation bar. Note that the *PREFERENCES* page has changed so that the "normal" query style is

no longer the default value.

12. Look at the search form in the collection. There are two fields that can be searched: *text* and *Title*. Add some more fields to search on by going back to the Librarian Interface.
13. In the **Design** panel, go to the **Search Indexes** section. Add a new index based on **ex.Subject** by clicking **<New Index>**, selecting **ex.Subject** in the list of metadata, and clicking **<Add Index>**.
14. **Rebuild** the collection and **preview** the results. Notice the extra field in the **... in field** drop-down menus in the search form. You can do quite complicated queries by searching for words in different fields at the same time.
15. To change the text that is displayed in the drop-down menus of the search form, go to the **Search** section of the **Format** panel. Here you can change the display text for the indexes.

## 2.5. CDS/ISIS collection

This exercise is similar to the **Bibliographic collection** exercise, except that a CDS/ISIS database is used instead of a MARC database, and we do not explode the database.

1. Start a new collection called **ISIS Collection**.
2. Drag the files from `sample_files isis` (excluding the `format_tweaks` folder) into the collection.
3. **Build** and **preview** the collection. The default indexes, classifiers and formats are not very useful for this data. There is no metadata searching, and the *Titles* classifier is completely empty. The filenames classifier is useless because all records come from the same file.
4. In the **Search Indexes** section of the **Design** panel, remove the useless Source and Title indexes, and add new indexes for Photographer<sup>all</sup>, Country<sup>all</sup> and Notes<sup>all</sup> metadata.

CDS/ISIS metadata has subfields, and these are represented using ^.

5. In the **Browsing Classifiers** section, remove the existing (useless) classifiers for **Title** and **Source**, and add a new **AZList** for **Photographer**.
6. **Rebuild** and **preview** the collection.
7. In the **Format Features** section of the **Format** panel, change the **VList** format statement to display **Photographer** and **Notes** metadata. Change it to look like:

```
<td valign=top>[link] [icon] [/link]</td>
<td valign=top><b>[ex.Photographer^all]</b><br/>[ex.Notes^all]</td>
```

8. Make fielded searching the default by changing the **SearchTypes** format statement to **form,plain** (instead of **plain,form**).

**ISISPlug** stores a nicely formatted version of the record as the document text, and this is what is displayed when we view a record. Lets tidy it up a little more.

9. Remove the *DETACH* and *NO HIGHLIGHTING* buttons by setting the **DocumentButtons** format statement to empty.
10. Remove the "Untitled" at the top of the document by setting the **DocumentHeading** format statement to empty.
11. Finally, lets link to the raw record, which is stored as **ISISRawRecord** metadata. Edit the **DocumentText** format statement to look like the following. (This format can be copied from `sample_files isis format_tweaks document_text.txt` .)

```
<p>[Text]</p>
{If}{_cgiargshowrecord_,
<p><b>CDS Record:</b><br/><tt>[ISISRawRecord]</tt></p>
<center><a href=\'_gwcgi_?e=_cgiarge_&a=d&d=_cgiargd_\'>Hide CDS Record</a></center>,
<center><a href=\'_gwcgi_?e=_cgiarge_&a=d&d=_cgiargd_&showrecord=1\'>Show CDS Record</a></center>
}
```

12. Preview the collection.

## 2.6. Editing metadata sets

GEMS (Greenstone Editor for Metadata Sets) can be used to modify existing metadata sets or create new ones. GEMS is launched from the Librarian Interface when you want to create a new metadata set, or edit an existing one. In this exercise, we run GEMS outside of the Librarian Interface.

### *Running GEMS*

1. Start the Greenstone Editor for Metadata Sets (GEMS):

**Start All Programs Greenstone Digital Library Software Greenstone Editor for Metadata Sets**

2. GEMS starts up with no metadata set loaded. You can start a new set, or open an existing one, from the **File** menu.

### *Creating a new metadata set*

3. In this exercise, we will create a new metadata set. In order to save time, we will base it on an existing one: Development Library Subset. From the **File** menu, select **New...** (**File New...**). A popup window appears: **New Metadata Set**. Fill in the fields. Use "My Metadata Set" for the **Metadata set title:**, "my" for the **Metadata set namespace:**, and select "Development Library Subset Example Metadata" from the **Base this metadata set on:** drop down list. Click **<OK>**.
4. The new metadata set will be displayed. The left hand side list the elements (and sub-elements, if any) for the set, and the right hand side displays the set or element attributes. Since the new set was based on the Development Library Subset metadata set, it already contains all the elements from that set.

### *Adding a new element to a metadata set*

5. Right click on **My Metadata Set** in the left hand tree (or in the blank space in the left hand side) and choose **Add Element** from the menu that appears. In the popup window, type "Category" for the new element name, and click **<OK>**. The new element will appear in the list.
6. In the right hand side, the default attributes will appear for the new element. "Label" and "definition" are used in the Librarian Interface when displaying metadata elements and their descriptions. These attributes can be set in multiple languages.
7. Save the new metadata set by **File Save**, then close the GEMS by **File Exit**.

---

Copyright © 2005 2006 2007 by the [New Zealand Digital Library Project](#) at [the University of Waikato](#), New Zealand

Permission is granted to copy, distribute and/or modify this document under the terms of the [GNU Free Documentation License](#), Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "[GNU Free Documentation License.](#)"