

## Lab 2 (3/4 hour): Adding and Importing Metadata

### Part I – Building an image collection

1. Start a new collection (File→New) called **backdrop**. Fill out the fields with appropriate information. For “**Base this collection on**”, pull down the item **Simple image collection (image-e)** from the menu. Then click <OK>.  
  
[You won’t be asked to choose a metadata set because the new collection inherits the one (or ones) used by the seed collection.]
2. Copy the images provided in the *workshop\_files* folder *sample\_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 ‘image-e’, but the content is different.
6. Change to the **Enrich** panel, and view the extracted metadata for *Ascent.jpg*.

*We will now manually add our own metadata and use it to give users a new way to browse the collection:*

7. The collection (image-e) on which **backdrop** is based uses only extracted metadata. We will add a further metadata set that matches our needs. For this exercise we use Dublin Core, a modest-sized metadata set that has been designed for a broad range of uses. Go to the main menu bar to the Librarian Interface and click **Metadata Sets->Import Set**.
8. In the window that pops up, select **dublin.mds** and click <import>.
9. If you are not already viewing the **Enrich** panel, do so now by clicking this tab. The metadata for each file now shows the Dublin core ‘dc.’ fields as well as the extracted ‘ex.’ fields.
10. Let’s work with just the first three files (*Ascent.jpg*, *Autumn.jpg* and *Azul.jpg*) to get a flavour of what is possible. First, we set each file’s dc.Title field to be the same as its filename but without the filename extension.
11. Click on *Ascent.jpg* so its metadata fields are available, then click on its **dc.Title** field on the right-hand side. Click on the **Value field** text box, enter **Ascent**, and

click <**append**>. (The **Previous Values** box will become more useful when more entries have been added.)

12. Repeat this process for *Autumn.jpg* and *Azul.jpg*.

## Part II – Customising the collection’s appearance

13. Building or previewing the collection at this point won’t reveal anything new. That’s because we haven’t changed the design of the collection to take advantage of the new metadata. Go to the **Design** panel (by clicking on its tab) and select **Format Features** from the left-hand list. Leave the **Editing Controls** at their default value, so that **Choose Feature** remains blank 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]’

Note: Metadata names are case-sensitive in Greenstone: it is important that you capitalize “Title” (and not “dc”). Next click <**Replace Format**>. 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.

Go to the **Create** panel and click <**Build Collection**>. Now **preview** the collection. Now when you click on **browse** in the navigation bar the presentation has changed to “Title: Ascent” and so on. Because we only assigned metadata to the first three items, after this the title becomes blank because the subsequent items have no dc.Title metadata. You need to spend more time entering the metadata to get a full listing.

[Note: For some design parameters the collection must be rebuilt before the effect of changes can be seen. However, changes to format statements take place immediately and you can see the result straightaway by pressing reload in the web browser. Above, you were asked to build before previewing just to simplify the explanation.]

14. Let’s 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 **plugin ImagePlug** to pop up a window that shows its settings. (Alternatively, select ImagePlug with a single click and then click <**Configure Selected Plugin**> further down the screen). Currently all options are off, so standard defaults are used. Select **thumbnailsize**, set it to **50**, and click <**OK**>.

**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 thumbnail size option off so the thumbnail will revert to its normal size when the collection is re-built.

16. Now add metadata that describes the photos in the collection. Again, for illustration, we focus on the first three images (*Ascent.jpg*, *Autumn.jpg* and *Azul.jpg*). Switch to the **Enrich** panel and select *Ascent.jpg*. We'll store our description in the **dc.Description** metadata element, so select it now in the right-hand panel.

What description should we 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 *Ascent.jpg*: 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 Librarian Interface enter the text **Moon rising over mountain landscape** as the **dc.Description** field's **value** and click **<append>** to have it added.

17. Repeat this process for *Autumn.jpg* and *Azul.jpg*.
18. Rebuild the collection to incorporate the new metadata.
19. Now update the format statement to use the new **dc.Description** metadata. Switch to the **Design** panel and enter the **Format Features** zone by selecting this from the list of names on the left-hand side and ensure **VList** is selected. In the **HTML Format String**, place your cursor after the text that says

```
[dc.Title]<br>
```

and add the following text:

```
Description: [dc.Description]<br>
```

Then click **<Replace Format>**.

20. Preview the result (you do not need to build the collection first, because changes to format statements take effect immediately). Each image's description should appear beside the thumbnail, following the title.
21. Now we'll add a new browsing option based on the descriptions. Switch to the **Design** panel and select **Browsing Classifiers** from the left-hand list. Set the menu item for **select classifier to add** to **AZList**; then click **Add classifier**.
22. A window pops up to control the classifier's options. Set the menu item for metadata to **dc.Description**. Beneath this item, select the **buttonname** option and enter **descriptions** in its text box. Finally click **<OK>**. Now switch to the **Create** panel, 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.

23. We now build a searchable index based on dc.Description metadata. Switch to the **Design** panel and select **Search Indexes** from the left-hand list. Enter the text **descriptions** as the **Index Name**, select **dc.Description** and click **<Add Index>**.
24. Switch to the **Create** panel, **build** the collection, then **preview** it. Search for the term "autumn" as an example.

### Part III – Working with bibliographic data

*We now turn attention to configuring collections based on bibliographic content. We use MARC as the sample input format:*

25. Start a new collection called **Beatles Bibliography** which will contain a collection of MARC records from the U.S Library of Congress on the Beatles. Enter the requested information and base it on "New Collection". There is no need to include any metadata sets.
26. In the **Gather** panel, open the *sample\_marc* folder in *workshop\_files* and drag **locbeatles50.marc** into the right-hand pane and dropping it there.
27. In the **Document Plugins** section of the **Design** panel, add **MARCPlug** to the list. Leave its options at their preset defaults.
28. Remove the plugins **TextPlug** to **PSPlug** (**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.

[Note: you must delete plugins one at a time. Alt-R is a hot key for this, and you can work faster by clicking on a plugin, deleting it with Alt-R, and continuing in this fashion.]

29. Now select **Browsing Classifiers** from within the **Design** panel and **remove** the default classifier for **Source** metadata. In this collection all records are from the same file, so Source metadata, which is set to the filename, is not particularly interesting.
30. Switch to the **Create** panel, **build** the collection, and **preview** it. Browse through the **titles a-z** and view a record or two. Try searching—for example, find items that include **George Martin**.
31. Add an **AZCompactList** classifier for the **Subject** metadata. Select this item from the relevant menu of the **Browsing Classifiers** section of the **design** panel and

click **<Add Classifier>**. In the popup window, select **ex.Subject** as the metadata item, activate the **mingroup** option and set its field to **1**.

**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. Setting **mingroup** to 1 means that the bookshelf appears even when there is just one item, and is done here to provide a more uniform display.

32. **Build** the collection and **preview** the result.
33. Make each bookshelf node show how many entries it contains by appending this to the **Format Features** for **VList** format statement in the **Design** panel:

```
{If}{ [numleafdocs] , <td><i> ( [numleafdocs] ) </i></td> }
```

Click **<Replace Format>**, switch to the **Create** panel, and click **<Preview>** (no need to rebuild).

34. Next add fielded searching. In the **Design** panel select **Search Types** from the left-hand list and activate the **Enable Advanced Searches** options.
35. **Rebuild** the collection and **preview** the results. 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 to control the advanced searching options.

#### Part IV – Extra work

*To finish the collection off, brand it with an image that will be used to represent the collection on the Greenstone page, and appear at the top of each page of the collection:*

36. From the **General** section of the **Design** panel, press the **<browse>** button next to the label **URL to about page icon** and use the resulting popup file browser to access the folder:

My Computer → D → workshop\_files → sample\_marc

Select *beatles\_logo.jpg* and press open.

[This assumes that 'D' is the letter used for the drive where the workshop files are stored: this may vary depending on how the computer has been configured.]

37. Repeat this process for the **URL to home page icon**, selecting the same image.
38. Now **rebuild** the collection and **preview** it.