

Lab 3 (1 3/4 hours): Adding and using metadata

Part I – Building an image collection

The first image collection utilises the Librarian Interface's "base this collection on" feature to simplify the design process. The extra work section asks you to build an image collection from scratch.

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 **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 on the CD-ROM in the *test_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

1. 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 enter the **Format Features** zone by selecting this from the list of names on the left-hand side. Edit the format statement for the **VList** item:

Change 'Image name:' 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.]

2. 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.

3. 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.
4. 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.

5. Repeat this process for *Autumn.jpg* and *Azul.jpg*.
6. Now update the format statement to use the new **dc.Description** metadata. Switch to the **Design** panel 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**, place your cursor after the text that says

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

and add the following text:

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

Then click **<Replace Format>**.

7. **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.
8. 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**.
9. 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.

10. 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>**.
11. Switch to the **Create** panel, **rebuild** the collection, then **preview** it. Search for the term “autumn” as an example.

Controlling the building process:

12. 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.

Click on the **Create** panel and select **Import** from the list on the left-hand side. In the options that appear, select **maxdocs** and set its numeric counter to **3**. Now **rebuild**.

13. Preview the newly rebuilt collection's **browse** page. Previously this listed more than a dozen photos, but now there are just three—the first three files encountered by the building process, which happen to be the three images we assigned metadata to.

In this case the image collection, being intended for demonstration purposes, is small to begin with and doesn't take long to build in its entirety. Real collections often contain thousands of documents, and setting maxdocs to (say) 50 yields a good preview in a fraction of the time taken to build the whole collection.

14. Sometimes the building procedure does not process files in the way you expect, and it is helpful to study the text output generated during building. This shows which plugin processed which files, what indexes are being built, and so on. The volume of information displayed is controlled by the **verbosity** option on the **Create** panel.

First rebuild the collection and observe the current text that this generates.

15. Now click on the **Create** panel and select **Import** from the left-hand list. Select **verbosity** in the options that appear and set its numeric counter to **5**. **Rebuild** the collection and observe how much more information is generated.

16. Return the import setting back to the default values: unselect **maxdocs** and **verbosity**.

Part III – Extra work

Build a image collection starting from scratch—that is, start a new collection but instead of basing it on ‘image-e’, use the default settings for ‘New Collection’ and evolve the design manually until it behaves like the ‘backdrop’ collection you have been working on.