

# Greenstone Workshop, 2009

## Lab 1 Building collections and adding metadata

### 1.1. Building a small collection of HTML files

- Running the Greenstone Librarian Interface
- Starting a new collection
- Adding documents to the collection
- Building the collection
- Viewing the extracted metadata
- Viewing the internal links and external links
- Setting up a shortcut in the Librarian interface

### 1.2. A simple image collection

- Adding Title and Description metadata
- Change Format Features to display new metadata
- Changing the size of image thumbnails
- Adding a browsing classifier based on Description metadata
- Creating a searchable index based on Description metadata

### 1.3. A collection of Word and PDF files

- Viewing the extracted metadata
- Manually adding metadata to documents in a collection
- Document Plugins
- Search indexes
- Browsing classifiers
- Renaming the search indexes
- Classifying on multiple metadata

### 1.4. Exporting a collection to CD-ROM/DVD

## Lab 2: Designing collections

### 2.1. A large collection of HTML files—Tudor

- Extracting more metadata from the HTML
- Looking at different views of the files in the **Gather** and **Enrich** panels

### 2.2. Enhanced collection of HTML files—Tudor

- Adding hierarchically-structured metadata and a **Hierarchy** classifier
- Adding a hierarchical phrase browser (PHIND)
- Partitioning the full-text index based on metadata values
- Controlling the building process

## Lab 3: Collections of bibliographic material

### 3.1. Bibliographic collection

- Using fielded searching
- Exploding the database
- Reformatting the collection to use the exploded metadata

### 3.2. CDS/ISIS collection

## Lab 4: Greenstone formatting exercises

### 4.1. Formatting the HTML collection—Tudor

### 4.2. Formatting the Word and PDF collection

- Tidying up the default format statement
- Linking to Greenstone version or original version of documents
- Making bookshelves show how many items they contain
- Displaying multi-valued metadata
- Advanced multi-valued metadata

## Lab 5: Advanced collection configuration

### 5.1. Pointing to documents on the web

### 5.2. Enhanced Word document handling

- Using Windows native scripting
- Modes in the Librarian Interface
- Defining styles
- Removing pre-defined table of contents
- Extracting document properties as metadata

### 5.3. Enhanced PDF handling

- Modes in the Librarian Interface
- Splitting PDFs into sections
- Using image format
- Using **process\_exp** to control document processing (advanced)
- Opening PDF files with query terms highlighted

### 5.4. Section tagging for HTML documents

## Lab 6: Multimedia

### 6.1. Looking at a multimedia collection

### 6.2. Building a multimedia collection

- Manually correcting metadata
- Browsing by media type
- Suppressing dummy text
- Using **AZCompactList** rather than **AZList**
- Making bookshelves show how many items they contain
- Adding a Phind phrase browser
- Branding the collection with an image
- Using **UnknownPlugin**
- Cleaning up a title browser using regular expressions
- Using non-standard macro files
- Using different icons for different media types
- Changing the collection's background image

Building a full-size version of the collection  
Adding an image collage browser

## Lab 7: Scanned images

### **7.1. Scanned image collection**

Grouping documents by series title and displaying dates within each group  
Displaying scanned images and suppressing dummy text  
Searching at page level

### **7.2. Advanced scanned image collection**

Adding another newspaper to the collection  
XML based item file  
Using **process\_exp** to control document processing  
Switching between images and text

## Lab 8: Customization

### **8.1. Customization: macro files and stylesheets**

Collection specific customisation  
Changing the colour of the page title and page text  
Make your own Greenstone home page  
How to determine which images to replace (advanced)

## Lab 9: Sharing collections with OAI-PMH

### **9.1. Open Archives Initiative (OAI) collection**

Tweaking the presentation with format statements

### **9.2. Downloading over OAI**

Downloading using the Librarian Interface  
Downloading using the command line

## Lab 10: Miscellaneous

### **10.1. Downloading files from the web**

### **10.2. Editing metadata sets**

Running GEMS  
Creating a new metadata set  
Adding a new element to a metadata set

### **10.3. Building and searching with different indexers**

Build with Lucene  
Search with Lucene  
Build with MGPP  
Search with MGPP  
Use search mode hotkeys with query term  
A quick reference of the search mode hotkeys in MGPP

---

Copyright © 2005 2006 2007 2008 2009 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.”](#)