Lab 1

(Review)
Installing Greenstone, building collections, adding metadata

Tutorial exercise #1.1
Installing Greenstone (local library version)
- Self-installing CD-ROM
  - Greenstone
  - ImageMagick: yes
  - Ghostscript: yes
- Copy sample_files to desktop
- Try the reader's interface
- Librarian interface:
  - Navigate to Home Folder\Desktop\sample_files
  - Right-click

Tutorial exercise #1.2
Updating a Greenstone installation
- Removing Greenstone from a Windows system
- Reinstalling Greenstone on a Windows system
- Amalgamating different Greenstone collections

Tutorial exercise #1.3
Small collection of HTML files
- Invoke GLI: build a small collection of HTML files
  - Gather
  - Create
  - Look at extracted metadata
  - Set up shortcut in the Librarian interface

Tutorial exercise #1.4
Simple image collection
- Gather
- Create
- Enrich:
  - Add Dublin Core metadata set
  - Add dc.Title metadata
- Design:
  - Change format to display titles; re-build
  - Change thumbnail size; re-build
- Enrich:
  - Add dc.Description metadata
- Design:
  - Change format to display description
  - Add browser for descriptions (AZList)
  - Add searchable index based on descriptions

Tutorial exercise #1.5
Collection of Word and PDF files
- Gather
- Create
- Enrich:
  - Manually add metadata (dc.Title, dc.Creator)
  - Multiple-valued metadata
- Design:
  - Plugins
  - Indexes: text, title, author
  - Browsers: title (AZList), author (AZCompactList)
  - Classifying on multiple metadata
Tutorial exercise #1.6
Exporting a collection to CD-ROM/DVD

Lab 2
Designing collections

Tutorial exercise #2.1, 2.2
The Tudor collection of HTML files

#2.1 A large collection of HTML files—Tudor
Extracting more metadata from HTML
Blocking the stray images

#2.2 Enhanced collection of HTML files—Tudor
Adding hierarchically-structured metadata
Adding a Hierarchy classifier
Partitioning the full-text index
Controlling the building process

#4.1 Formatting the HTML collection—Tudor
#5.1 Pointing to documents on the web

Lab 3
Collections of bibliographic material

Building a collection from MARC records

Sample record

tutorial exercise #3.1
Bibliographic collection

#3.1 Bibliographic collection
- AZCompactList for Subjects
- Adding fielded searching
- Exploding the database
- Reformating the collection using the exploded metadata
Dublin Core crosswalk

Greenstone’s MARC to Dublin Core mapping
based on LOC’s MARC ⇒ DC crosswalk

### Subject

# 650 = “Subject: Topical Term”
650 -> Subject

# 653 = “Index Term: Uncontrolled”
653 -> Subject

### Publisher

# 260 = “Publication, etc”
260 -> Publisher

### Relation

# 787 = “Nonspecific Relationship Note”
787 -> Relation

### Rights

# 540 = “Reproduction Note”
540 -> Rights

### Creator

# 720 = “Uncontrolled Name”
#   with $e=author
720^e -> Creator
# 100 = “Personal Name”
100 -> Creator
# 110 = “Corporate Name”
110 -> Creator

### Source

# Greenstone has its own Source
# metadata...
# 786 = “Data Source Entry”
786 -> MarcSource

### Title

# 245 = “Title Statement”
245 -> Title

### Description

# 520 = “Summary, Note”
520 -> Description

### Type

# 655 = “Index Term - Genre/Form”
655 -> Type

### Identifier

# 024 = “Standard Identifier/number”
# greenstone uses its own Identifier metadata
024 -> MarcIdentifier
# 856 = “Electronic Location”
856 -> URL

### Language

# Greenstone has its own Language metadata...
# 546 -> MarcLanguage

---

**Tutorial exercise #3.2**

CDS/ISIS collection

#3.2 CDS/ISIS collection

- Delete browsing classifiers:
  - add AZList for Photographer
- Change format statement to display Photographer and Notes metadata
- Link to the raw CDS/ISIS record

---

**Lab 4**

Greenstone formatting exercises

---

**Tutorial exercises #4.1, 4.2**

Greenstone formatting exercises

#4.1 Formatting the HTML collection—Tudor
  - Replacing the default format statement
  - Formatting the subject hierarchy

#4.2 Formatting the Word and PDF collection
  - Understanding the default format statement
  - Linking to the original version of documents
  - Making bookshelves show the number of items
  - Displaying multi-valued metadata
  - Advanced multi-valued metadata

---

**Lab 5**

Advanced collection configuration
**Tutorial exercises #5.1, 5.2**  
Advanced collection configuration

- **#5.1**  
  Pointing to documents on the web

- **#5.2**  
  Enhanced PDF handling
  - Modes in the Librarian Interface
  - Splitting PDFs into sections
  - Using image format
  - Using `process_exp` to control processing
  - Opening PDFs with query terms highlighted

---

**Tutorial exercises #5.3, 5.4**  
Advanced collection configuration

- **#5.3**  
  Enhanced Word document handling
  - Using Windows native scripting
  - Defining styles
  - Extracting document properties as metadata

- **#5.4**  
  Section tagging for HTML documents

---

**Lab 6**

**Sharing collections with OAI-PMH**

- **Tutorial exercise #6.1, 6.2**  
  OAI collection
  
  Using OAI-PMH, build a Greenstone collection based on metadata exported from the OAI server rocky.dlib.vt.edu.

  - **#6.1**  
    Open Archives Initiative (OAI) collection
    - Making the collection
    - Formatting it

  - **#6.2**  
    Downloading over OAI
    - Downloading using the Librarian Interface
    - Downloading using the command line

---

**Lab 7**

**Multimedia and scanned images**

- **Tutorial exercise #7.1, 7.2**  
  The Beatles multimedia collection

  - **#7.1**  
    Looking at a multimedia collection
  
  - **#7.2**  
    Building the 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
    - Using `UnknownPlug`
    - Cleaning up a title browser using regular expressions
    - Using non-standard macro files
    - Changing the collection’s background image
**Tutorial exercise #7.3, 7.4**

**The Niupepa newspaper collection**

### #7.3 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

### #7.4 Advanced scanned image collection
- Creating item files
- Using multiple copies of a plugin
- Switching between images and text

---

**Niupepa newspaper collection**

What the .item files mean:

<table>
<thead>
<tr>
<th>Number</th>
<th>Volume</th>
<th>Issue</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9-1-1)</td>
<td>Te Waka Volume 1 Number 1</td>
<td></td>
<td>p.1 gif, p.2 gif, p.3 gif, p.4 gif</td>
</tr>
<tr>
<td>(9-1-2)</td>
<td>Te Waka Volume 1 Number 2</td>
<td></td>
<td>p.5 gif, p.6 gif, p.7 gif, p.8 gif</td>
</tr>
<tr>
<td>(10-1-1)</td>
<td>Te Whetu Volume 1 Number 1</td>
<td></td>
<td>p.1 gif, p.2 gif, p.3 gif, p.4 gif, p.5 gif</td>
</tr>
<tr>
<td>(10-1-2)</td>
<td>Te Whetu Volume 1 Number 2</td>
<td></td>
<td>p.6 gif, p.7 gif, p.8 gif</td>
</tr>
<tr>
<td>(10-1-3)</td>
<td>Te Whetu Volume 1 Number 3</td>
<td></td>
<td>p.9 gif, p.10 gif</td>
</tr>
</tbody>
</table>

---

**Lab 8**

**Customization**

**Tutorial exercise #8.1**

**Customization: Macro files and stylesheets**

- Collection-specific macros
- Collection-specific styles
- Collection-specific stylesheets
- Make your own Greenstone home page