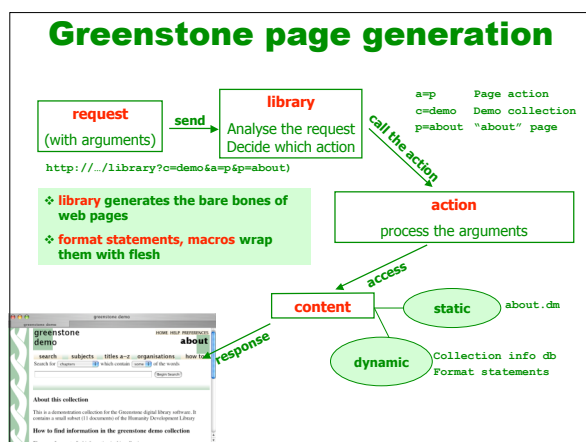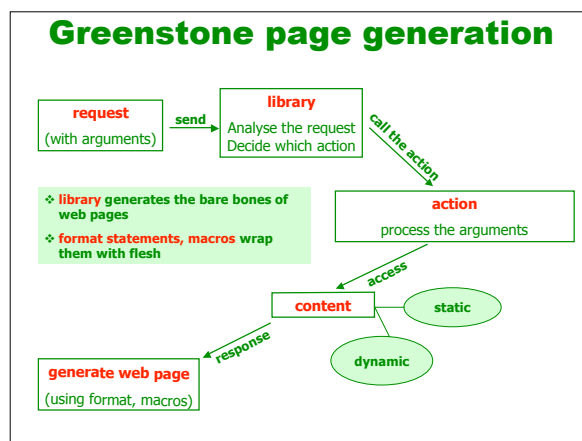## Interoperability

Course material prepared by

Greenstone Digital Library Project
University of Waikato, New Zealand

Agenda

❖ Interface customization
❖ Customizing parts of a page
❖ Greenstone in your organization
❖ Installing the web library (demo)
❖ Administration pages
❖ User management
❖ Restricting access to documents
❖ Building a collection
❖ The dreaded black screen
❖ More on building
❖ Multilingual support
❖ Language interfaces

Interface customisation

❖ Customising the Greenstone pages requires some knowledge of HTML
– Can't use an HTML editor because the pages aren't static: they are generated "on-the-fly"
❖ Customising the Greenstone pages involves:
– Macro files
❖ In the Greenstone "macros" folder
– Images
❖ In the Greenstone "images" folder
– Format statements
❖ GLI: "Format Features" part of the Design pane

Greenstone page generation

request (with arguments) — **send** → library: Analyse the request, Decide which action — **call the action** → action: process the arguments

❖ **library** generates the bare bones of web pages
❖ **format statements, macros** wrap them with flesh

action — **access** → content — static / dynamic

content — **response** → generate web page (using format, macros)

Greenstone page generation

request (with arguments) — **send** → library: Analyse the request, Decide which action — **call the action** → action: process the arguments

a=p   Page action
c=demo   Demo collection
p=about "about" page

http://…/library?c=demo&a=p&p=about

❖ **library** generates the bare bones of web pages
❖ **format statements, macros** wrap them with flesh

action — **access** → content — static (about.dm) / dynamic (Collection info db, Format statements)

content — **response** → (page)

Macros

❖ Macros
– Let you customize presentation without editing code
– Are resolved at run-time:
❖ Allows pages to be presented in different languages
❖ Allows variables to be printed into the page text (e.g. number of search hits)
❖ Macro files
– Stored in the Greenstone macros folder, named *.dm
❖ home.dm for the home page, query.dm for search page etc.
– Each file defines one or more "packages"
❖ A "package" is a group of macros
– Loaded on startup (= every request, for Web Library)
– Listed in the main.cfg file in Greenstone's "etc" folder

What do macros look like?

❖ Names begin and end with underscores (like _this_)

❖ Content is defined using curly brackets

macroname [Arguments] { Macro content }

– Can contain plain text, HTML, and other macro names
– Can contain conditional statements
  ❖ _If_ ...
– Can take arguments to restrict when they apply
  ❖ l=fr, c=demo etc.

Interface Language Macrofiles

english.dm

```
#------------------------------------
# text macros
#------------------------------------

_textimageabout_ {About page}
_textimagehelp_ {Help page}
_textimagepref_ {Preferences page}
_textimagegreenstone_ {Greenstone Digital Library Software}

_textimagesearch_ {Search for specific terms}
_textimageCreator_ {Browse alphabetical list of authors}
_textimageTitle_ {Browse alphabetical list of titles}
_textmatches_ {Matches }
_textbeginsearch_ {Begin Search}

_texticonsmalltext_ {View this section of the text}
_texticondetach_ {Open this page in a new window}
_texticonexpandtoc_ {Expand table of contents}
_texticonexpandtext_ {Display all text}
_texticonhighlight_ {Highlight search terms}

_textlanguage_ {Interface language: }
_textsimplemode_ {simple query mode}
_textadvancedmode_ {advanced query mode (allows boolean searching u
```

Interface Language Macrofiles

chinese.dm

```
#------------------------------------
# text macros
#------------------------------------

_textimageabout_ [l=zh] {主页}
_textimagehelp_ [l=zh] {帮助页}
_textimagepref_ [l=zh] {选项页}
_textimagegreenstone_ [l=zh] {绿石数字化图书馆软件}

_textimagesearch_ [l=zh] {特定范围检索}
_textimageCreator_ [l=zh] {按作者名字的阿拉伯字母顺序浏览 }
_textimageTitle_ [l=zh] {按题名的阿拉伯字母顺序浏览}
_textmatches_ [l=zh] {匹配}
_textbeginsearch_ [l=zh] {开始查询}

_texticonsmalltext_ [l=zh] {  浏览该部份文本}
_texticondetach_ [l=zh] {  打开一个新的窗口显示本页}
_texticonexpandtoc_ [l=zh] {  扩展目录表}
_texticonexpandtext_ [l=zh] {  显示所有文本}
_texticonhighlight_ [l=zh] {  查询项以粗体显示}

_textlanguage_ [l=zh] {界面语言 : }
_textsimplemode_ [l=zh] {简单查询方式}
_textadvancedmode_ [l=zh] {高级查询方式(可以使用 !、&、 | 和括号来布尔查询)}
```
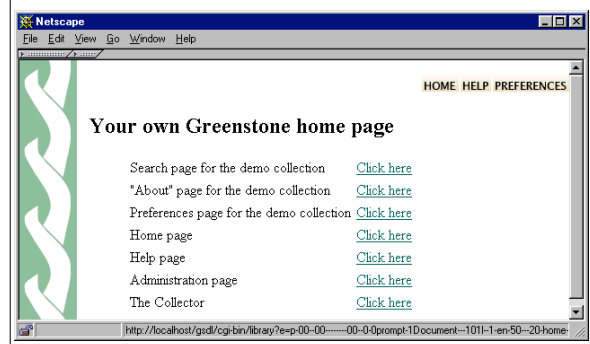
Customising the home page

C:\Program Files\Greenstone\etc\main.cfg   change *home.dm* to *yourhome.dm*

yourhome.dm

```
package home
_content_ {
<h2>Your own Greenstone home page</h2>
<ul>
<table>
<tr valign=top><td>Search page for the demo collection<br></td>
    <td><a href="_httpquery_&c=demo">Click here</a></td></tr>

<tr><td>"About" page for the demo collection</td>
    <td><a href="_httppageabout_&c=demo">Click here</a></td></tr>

<tr><td>Preferences page for the demo collection</td>
    <td><a href="_httppagepref_&c=demo">Click here</a></td></tr>

<tr><td>Home page</td>
    <td><a href="_httppagehome_">Click here</a></td></tr>

<tr><td>Help page</td>
    <td><a href="_httppagehelp_">Click here</a></td></tr>

<tr><td>Administration page</td>
    <td><a href="_httppagestatus_">Click here</a></td></tr>

<tr><td>The Collector</td>
    <td><a href="_httppagecollector_">Click here</a></td></tr>

</table>
</ul>
}
```

Macros used in home.dm

|	
httppagehome	Address of the home page
httppagehelp	… the help page
httppagestatus	… the administration page
httppagecollector	… the Collector page
httpquery&c=demo	… the search page for the demo collection
httppageabout&c=demo	… the about page for the demo collection
httppagepref&c=demo	… the preferences page for the demo collection
_content_{ … }	defines a macro called _content_ contains HTML, but '{', '}', '\', and '_' must be escaped with a backslash
_header_{ … }	HTML page header (contains background)
_footer_{ … }	HTML page footer

Collection-specific macros

❖ Macro files usually affect all collections

❖ To change the interface for one collection only:
  – Create a new macro file called extra.dm in the collection's "macros" folder
    ❖ Create this folder if it doesn't exist
  – Add collection specific macros to this file
    ❖ Package line in the beginning is mandatory
    ❖ Specify the collection name as an argument to the macro:
      content [c=collectionname] { }

Garish collection

❖ A very ugly version of the demo collection!
❖ Uses collection-specific macros and CSS files
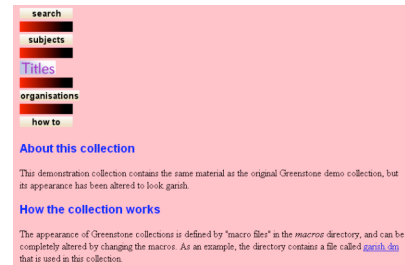
Garish: package "Global"

```
####################################
package Global
####################################
_starthighlight_  {<b>}
_endhighlight_    {</b>}
_navbarspacer_
    { <br/>
      <img src="_httpcimages_/horzline.gif"
         width="87" height="17">
      <br/> }
```

Highlight query terms in bold

Separates search and classifier buttons

Garish: package "Style"

```
#####################################
package Style
#####################################
_htmlhead_ [c=garish] {<html}
<head>
    <title>_pagetitle_</title>
    <link rel="stylesheet"
       href="_httpimg_/garish/style.css"
       type="text/css"/>
  </head>
<body>
}

_pagebanner_ [c=garish] {}
```

Page header macro
This overrides it to make all pages use the style sheet

Pagebanner normally includes the collection's logo, home/help/preferences buttons, and the image at top left that identifies the page.
This overrides it to suppress these images

Garish: package "Style"

```
#####################################
package Style
#####################################
_collectionspecificstyle_  {
<link rel="stylesheet"
     href="_httpcimages_/style.css"
     type="text/css"
     title="Greenstone Style"
     charset="UTF-8">
<style>
     body.bgimage \{ background-image: none; \}
     span.navlink:hover \{ background-image:
         url("_httpcimages_/bg_blue.gif"); \}
</style>}

_pagebanner_   {
<div class="pageinfo">
<p class="bannerlinks">_homelink_</p>
</div>
}
```

This uses a collection-specific style sheet

Pagebanner normally includes the collection's logo, home/help/preferences buttons, and the image at top left that identifies the page.
This overrides it to suppress all these and just leave in the home button

Interface customisation: images

❖ The look of the collection can be changed very easily by editing the Greenstone images
  – Right-click on an image in the web browser to find its name
  – Open in an image editor and edit to suit your collection
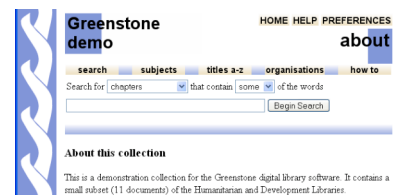
Bluestone?

Agenda

- ❖ Interface customization
- ❖ Customizing parts of a page
- ❖ Greenstone in your organization
- ❖ Installing the web library (demo)
- ❖ Administration pages
- ❖ User management
- ❖ Restricting access to documents
- ❖ Building a collection
- ❖ The dreaded black screen
- ❖ More on building
- ❖ Multilingual support
- ❖ Language interfaces

Customizing parts of a page

① In the General part of GLI's Design pane, Browse for the "about page" icon and select your image

Customizing parts of a page

② Macro called _globallinks_
in **package global** in the **nav_css.dm** file

    javalinks {_homelink__helplink__preflink_}

Position of the buttons is in macro _pagebanner_ in the **style.dm** file. You can change the position of the buttons, add new ones, delete some, or display none at all

Customizing parts of a page

③ Macro called _imagethispage_, e.g. in **package help**
    package help
    imagethispage {_textHelp_}

the icon is language-specific, so it's in **english.dm**
    textHelp { help }

Labels like *subjects* are metadata-dependent, and defined in **english.dm** like this:
    labelSubject { subjects }

Customizing parts of a page

④ Button names are determined by the classifiers, e.g. in **collect.cfg**

    classify AZList -metadata Title   *button is named* titles a–z
    classify AZList -metadata Title -buttonname XXX   *button is named* XXX

Customizing parts of a page

⑤ The classifier list display is defined by a format string in the Format Features part of GLI's Design pane

Documentation and Help

❖ *Developer's Guide* (developer.pdf)

❖ Customizing the Greenstone user interface
(http://www.wrlc.org/dcpc/UserInterface/interface.htm)

❖ Documented example collections:
– Garish collection
– Authentication and formatting demo

Help!

Email us:
greenstone@cs.waikato.ac.nz

Join the mailing list:
https://list.scms.waikato.ac.nz/mailman/listinfo/
greenstone-users

Agenda

❖ Interface customization
❖ Customizing parts of a page
➡ ❖ Greenstone in your organization
❖ Installing the web library (demo)
❖ Administration pages
❖ User management
❖ Restricting access to documents
❖ Building a collection
❖ The dreaded black screen
❖ More on building
❖ Multilingual support
❖ Language interfaces

Greenstone in your organization

❖ Who needs to access your Greenstone collections?
– Just you: use the Local Library version

– Other people in your organization, over your network:
use the Web Library version
❖ Possible to use Local Library, but not recommended

– People outside your organization, using the internet: use
the Web Library version
❖ If your organization has a website already, you'll need to get
your ISP to install Greenstone on their servers
❖ If your organization doesn't have a website already, you'll need
to contract an ISP to install Greenstone and host your collections
for you

Local library vs Web library

❖ **Local library: stand-alone**
– **Serves collections on a standalone PC ...**
– **...and on others on the same network**
– **Includes built-in Web server**
❖ **Web library: uses external web server**
– **Apache, Microsoft PWS/IIS**

❖ **Windows: Both local library and web library**
– **All versions: (3.1, 3.11), 95, 98, NT, 2000, ME, XP**
– **Binaries supplied**
– **Normally use local library (else must set up server)**
– **Web library works with Microsoft PWS, IIS**
❖ **Unix, Mac OS X: Web library only**
– **Use Apache (or other web server)**
– **Linux binaries supplied**
– **Tested on SUN Solaris, Mac OS X**
– **Need GDBM (standard on Linux)**

Agenda
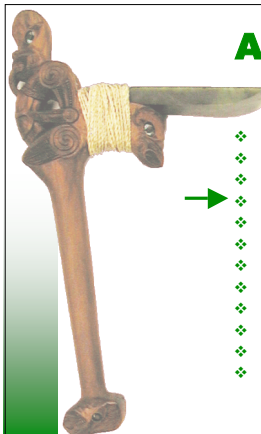
❖ Interface customization
❖ Customizing parts of a page
❖ Greenstone in your organization
➡ ❖ Installing the web library (demo)
❖ Administration pages
❖ User management
❖ Restricting access to documents
❖ Building a collection
❖ The dreaded black screen
❖ More on building
❖ Multilingual support
❖ Language interfaces

Web Library:
Installation demo

Configuring Apache

❖ Download Apache from httpd.apache.org

❖ Install, then add to the Apache httpd.conf file:

```
ScriptAlias /gsdl/cgi-bin "C:/Program Files/Greenstone/cgi-bin"
<Directory "C:/Program Files/Greenstone/cgi-bin">
    Options None
    AllowOverride None
</Directory>

Alias /gsdl "C:/Program Files/Greenstone"
<Directory "C:/Program Files/Greenstone">
    Options Indexes MultiViews FollowSymLinks
    AllowOverride None
    Order allow,deny
    Allow from all
</Directory>
```

❖ Access Greenstone at:

http://<host-name>/gsdl/cgi-bin/library.exe

Configuring PWS/IIS

1. Select *Advanced* to get the *Advanced Options* screen.

2. Select *Home* and click *Add*. Fill out the fields as follows:

Directory field:	*C:\Program Files\gsdl*
Alias field:	*gsdl*
Access permissions:	*Read*
Application permissions:	*None*
Click *OK*	

   This makes Greenstone files accessible to the webserver.

3. Back in *Advanced Options*, select *gsdl* and click *Add*. Fill out the fields as follows:

Directory field:	*C:\Program Files\gsdl\cgi-bin*
Alias field:	*cgi-bin*
Access permissions:	*None*
Application permissions:	*Execute*
Click *OK*	

   This allows the Greenstone program *library.exe* to be executed by the webserver.

4. Go to the URL *http://localhost/cgi-bin/library.exe*.

   Note: you need to specify the *.exe* file extension with PWS and IIS.

Web Library Troubleshooting
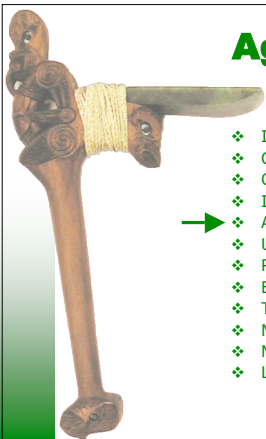
Problem	Try this
When I start Apache, it quits immediately.	Add a *ServerName localhost* directive to the Apache configuration file (see Section 4.1 of the Greenstone Installer's Guide).
When I point my browser at the digital library, it displays garbage – a binary file.	Check the *ScriptAlias* directive in the Apache configuration file, making sure it comes before the *Alias* directive (see Section 4.2 and 4.3 of the Greenstone Installer's Guide).
I get the Greenstone home page, but the Demo collection icon does not appear.	Run the program *library* (in the Greenstone cgi-bin directory) from the DOS prompt (or shell) to generate debugging information that will help you locate the problem.

Documentation and Help

❖ *Installer's Guide* (install.pdf)

   Versions of Greenstone, installation procedure, Greenstone collections, setting up the web server, configuring your site, personalizing your installation

❖ FAQ: sections on
   – Obtaining Greenstone
   – Installing Greenstone
   – Running Greenstone

Agenda

❖ Interface customization
❖ Customizing parts of a page
❖ Greenstone in your organization
❖ Installing the web library (demo)
❖ Administration pages
❖ User management
❖ Restricting access to documents
❖ Building a collection
❖ The dreaded black screen
❖ More on building
❖ Multilingual support
❖ Language interfaces

Administration Pages

❖ Access the administration pages by clicking on the button on the Greenstone home page

greenstone digital
library software

HELP PREFERENCES

select a collection

Greenstone
demo

|	
The Librarian Interface	Helps you create new collections, modify or add to existing ones, or delete collections
Administration Page	Allows you to add new users, summarizes the collections in the system, gives technical information on the Greenstone installation
About Greenstone	Tells you about the Greenstone software and the New Zealand Digital Library Project where it originated
Greenstone Documentation	Greenstone manuals

❖ Default admin user password: "admin"

Administration Pages

Edit
main.cfg file

View usage
and error logs

List of
collections
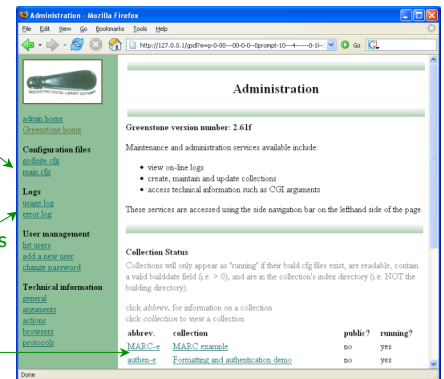
Agenda

❖ Interface customization
❖ Customizing parts of a page
❖ Greenstone in your organization
❖ Installing the web library (demo)
❖ Administration pages
➡ ❖ User management
❖ Restricting access to documents
❖ Building a collection
❖ The dreaded black screen
❖ More on building
❖ Multilingual support
❖ Language interfaces

User Management

❖ Users are added/edited/removed using the links on the Administration page

User management
list users
add a new user
change password

Show users, with
links to edit or
remove them
(admin only)

Add a new user to
the system (admin
only)

Change password (any
user with an account)

User Management

❖ User details:
– Username
– Password
– Groups (controls what the user can do)
– Comment

Add a new user

username	Usernames must be between 2 and 30 characters long. They can contain alphanumeric characters, '.', and '_'.
password	Passwords must be between 3 and 8 characters long. They can contain any normal printable ASCII characters.
account status	enabled
groups	Groups is a comma separated list, do not put spaces after the commas.
comment	

submit cancel

Agenda

❖ Interface customization
❖ Customizing parts of a page
❖ Greenstone in your organization
❖ Installing the web library (demo)
❖ Administration pages
❖ User management
➡ ❖ Restricting access to documents
❖ Building a collection
❖ The dreaded black screen
❖ More on building
❖ Multilingual support
❖ Language interfaces

Restricting Access to Documents

❖ Restrict users to view certain documents

❖ Define groups and users in admin pages

❖ Restrict levels
– collection (default) => the full collection
– document => individual documents

Restricting Access to Documents

❖ How to:

– Set up in collection configuration file
(collect_name/etc/collect.cfg)

– *authenticate* directive

– *auth_groups* directive

– Restrict documents
  ❖Public_documents
  ❖Private_documents

– Use document identifiers
  ❖Look at URL =>d argument

Restrict Collection

Authenticate Collection
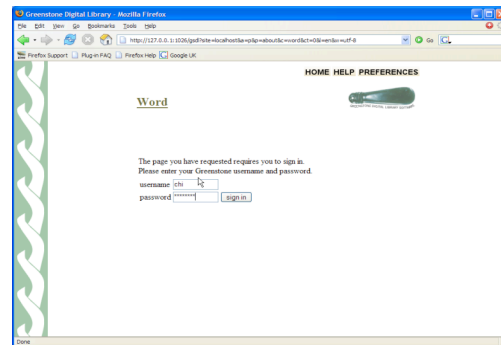
Documents Browsing

Restrict Document

Public_documents

Document:
HASHf48ef4f9081c1a3bbc7c68
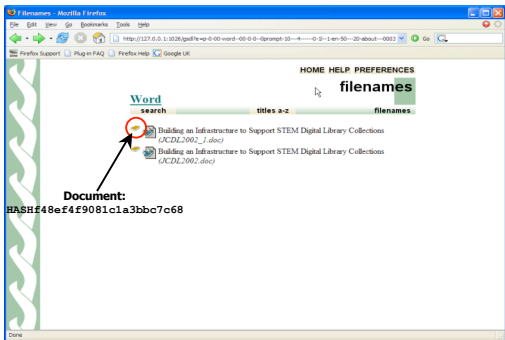
Authenticate document
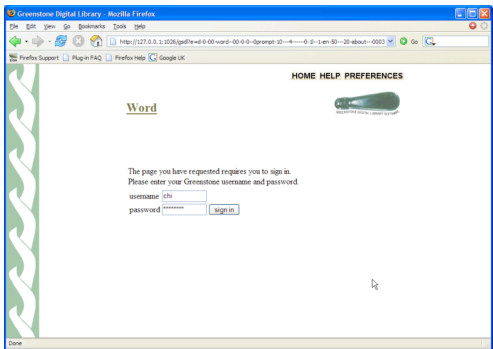
Document Display
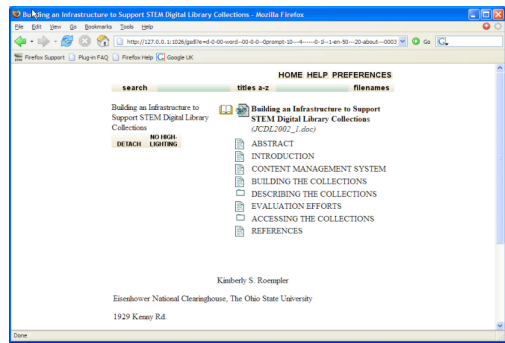
Agenda
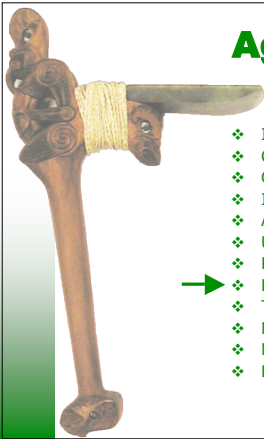
❖ Interface customization
❖ Customizing parts of a page
❖ Greenstone in your organization
❖ Installing the web library (demo)
❖ Administration pages
❖ User management
❖ Restricting access to documents
➡ ❖ Building a collection
❖ The dreaded black screen
❖ More on building
❖ Multilingual support
❖ Language interfaces

The building process

$GSDLHOME
collect
demo
import   archives   building   index   etc   perllib

Collection configuration file

Put material here ↑   ↑   ↑
        import   build   rename
                                 directory

Collection served from here (or to CD-ROM)

demo
import   archives   building   index   etc   perllib

Collection configuration file

import process
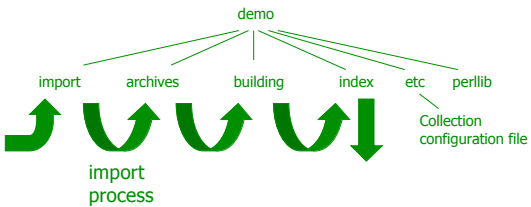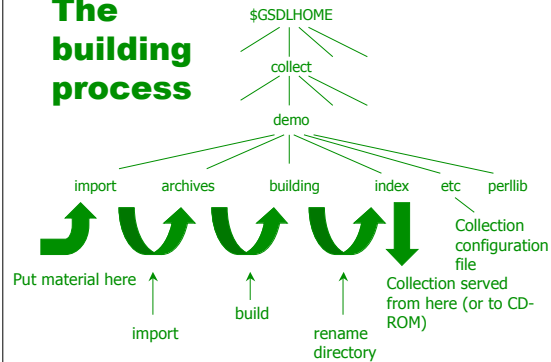
❖ Navigates *import* directory structure
❖ Assigns OIDs to documents
❖ Recognizes subsection structure

chapters, sections, subsections, pages, ...
used for (a) reading books, (b) search indexes

❖ Inserts metadata   *Dublin Core plus extensions*
❖ Converts to Greenstone Archive format   *uses plugins*
❖ Regularizes file structure

Slide 1:

demo

import　archives　building　index　etc　perllib

Collection
configuration file

build
process

❖ Creates indexes of full-text and/or metadata
❖ Compresses document text
❖ Classifies documents for browsing
❖ Generates a database for metadata, document
structure, and browsing classifier structure

Slide 2:

demo

import　archives　building　index　etc　perllib

Collection
configuration file

Rename
directory

❖ Delete current indexes – these are used to serve
the collection while the new index is being built
❖ Make the new index (in building directory) live (in
index directory).

Slide 3:

demo

import　archives　building　index　etc　perllib

Collection
configuration
file

❖ Controls import and build process
❖ Plugins for import
❖ Indexes, classifiers for build
❖ Collection metadata for serving

Slide 4:

demo

import　archives　building　index　etc　perllib

Put material here

Collection served from
here (or to CD-ROM)

misc subdirs	11 subdirectories	MG compressed text	collect.cfg
11 .htm	each with doc.xml	MG full-text indexes	
11 .jpg	+ associated .jpg	Gdbm database	mags.txt
248 .png	and .png files	Associated files	sub.txt
index.txt			org.txt

Slide 5:

demo

import　archives　building　index　etc　perllib

bostid	HASH0105.dir	(empty)	assoc	collect.cfg	classify
ecourier	HASH017d.dir		build.cfg	mags.txt	
faobetf	HASH63e6.dir		dtx	sub.txt	
index.txt	HASHaad6.dir		stt	org.txt	
wb	HASH0144.dir		stx		
	HASH026b.dir		text		
	HASH7df3.dir				
	HASHe52a.dir				
	HASH0173.dir				
	HASH54cf.dir				
	HASHa0a5.dir				
	archives.inf	(list of archived files)			

Slide 6:

Contents of demo/index

build.cfg　　　　used by receptionist to determine indexes

```
builddate  951855434
indexmap   section:text->stx  section:Title->stt  document:text->dtx
numbytes   3029746
numdocs    11
```

associated files	mg text		mg indexes		
assoc:	text:		stx: stt: dtx:		
HASH0141.dir	demo.ldb		demo.i	inverted file	
HASH0169.dir	demo.t	text	demo.tiw	doc weights	
HASH01a3.dir	demo.td	dictionary	demo.wa	approx weights	
HASH01b4.dir	demo.ti	text index	demo.idb	term dict	
HASH01ba.dir	demo.tsd	stats	demo.ib1	stem indexes:	
HASH01d6.dir			demo.ib2	casefolded,	
HASH0f76.dir			demo.ib3	stemmed, both	
HASH863c.dir					
HASH8b94.dir					
HASHc5b3.dir	document database				
HASHd803.dir					

Agenda

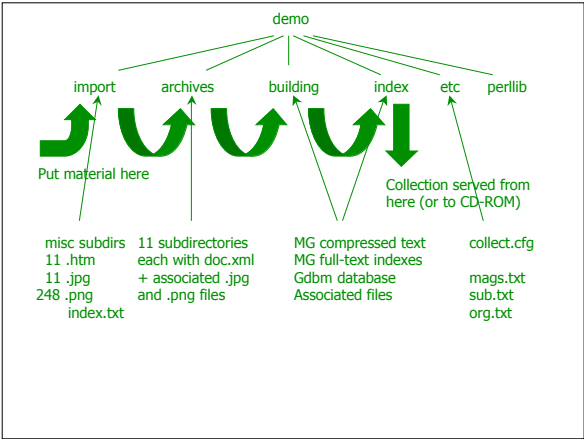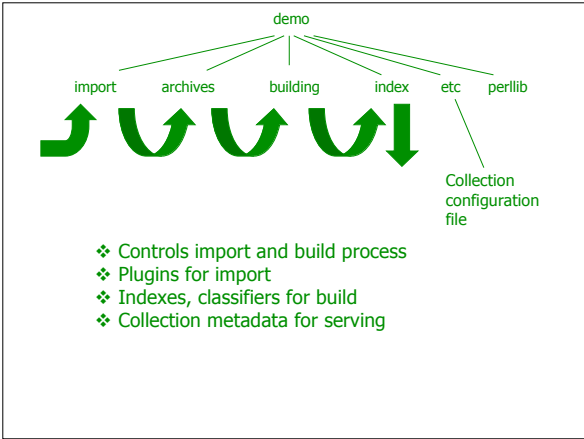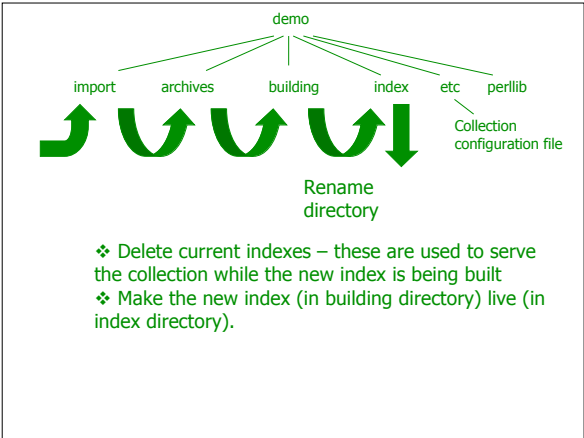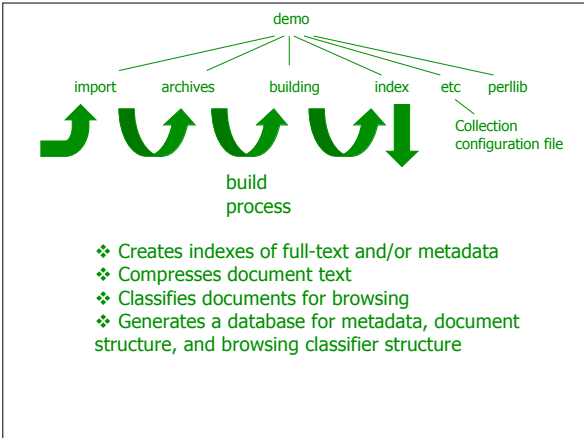- ❖ Interface customization
- ❖ Customizing parts of a page
- ❖ Greenstone in your organization
- ❖ Installing the web library (demo)
- ❖ Administration pages
- ❖ User management
- ❖ Restricting access to documents
- ❖ Building a collection
➤ ❖ The dreaded black screen
- ❖ More on building
- ❖ Multilingual support
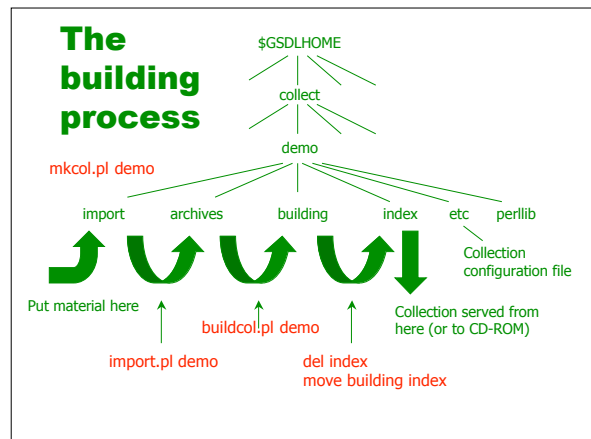- ❖ Language interfaces

The building process

$GSDLHOME

collect

demo

mkcol.pl demo

import   archives   building   index   etc   perllib

Collection configuration file

Put material here

buildcol.pl demo

Collection served from here (or to CD-ROM)

import.pl demo

del index
move building index

Start a command prompt

Don, Katherine

Internet — Mozilla Firefox
E-mail — Mozilla Thunderbird
Microsoft Office PowerPoint 2003
Microsoft Office Word 2003
Command Prompt
Secure File Transfer Client
Paint
Notepad

My Documents
My Recent Docume
My Computer
Control Panel
Printers and Faxes
Help and Support
Search
Run...
Windows Security

All Programs ▶

Log Off   Shut Down

Start   Lectures   Lectures

Run
Type the name of a program, folder, document, or Internet resource, and Windows will open it for you.
Open: cmd
OK   Cancel   Browse...

Command Prompt

```
Command Prompt

Microsoft Windows 2000 [Version 5.00.2195]
(C) Copyright 1985-2000 Microsoft Corp.

C:\>_
```

The building process

```
Command Prompt

Microsoft Windows 2000 [Version 5.00.2195]
(C) Copyright 1985-2000 Microsoft Corp.
C:\>_
C:\> cd "C:\Program Files\Greenstone"

C:\Program Files\Greenstone> setup

C:\Program Files\Greenstone>perl -S
        mkcol.pl
        -creator me@here colname
```
Copy source into collect\colname\import
```
C:\>perl -S import.pl colname

C:\>perl -S buildcol.pl colname
```
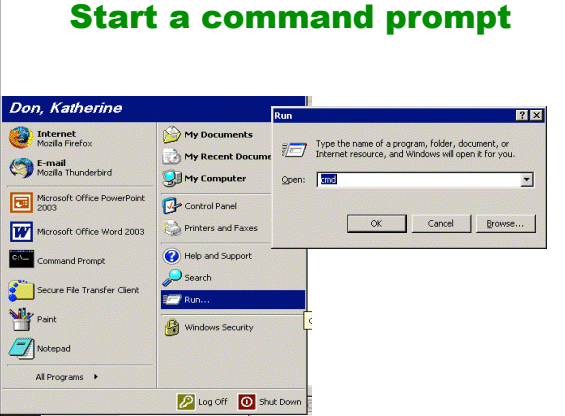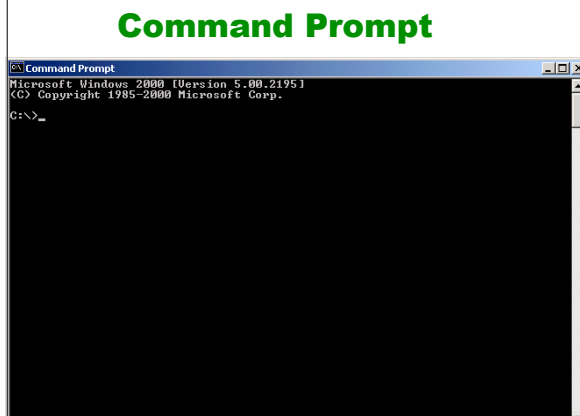Rename the "building" directory to "index"

Agenda

- ❖ Interface customization
- ❖ Customizing parts of a page
- ❖ Greenstone in your organization
- ❖ Installing the web library (demo)
- ❖ Administration pages
- ❖ User management
- ❖ Restricting access to documents
- ❖ Building a collection
- ❖ The dreaded black screen
➤ ❖ More on building
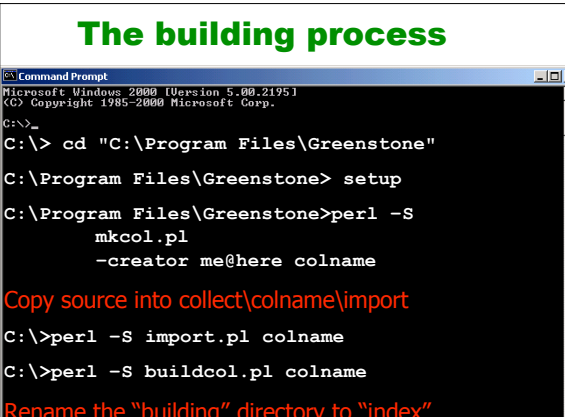- ❖ Multilingual support
- ❖ Language interfaces

Specifying metadata: XML metadata file

```xml
<?xml version="1.0" ?>
<!DOCTYPE GreenstoneDirectoryMetadata SYSTEM
"http://greenstone.org/dtd/GreenstoneDirectoryMetadata/1.0/GreenstoneD
irectoryMetadata.dtd">
<DirectoryMetadata>
  <FileSet>
    <FileName>nugget.*</FileName>
    <Description>
      <Metadata name="Title">Nugget Point, The Catlins</Metadata>
      <Metadata name="Place" mode="accumulate">Nugget Point</Metadata>
    </Description>
  </FileSet>
  <FileSet>
    <FileName>nugget-point-1.jpg</FileName>
    <Description>
      <Metadata name="Title">Nugget Point Lighthouse</Metadata>
      <Metadata name="Subject">Lighthouse</Metadata>
    </Description>
  </FileSet>
</DirectoryMetadata>
```

XML metadata format

Document type definition (DTD)

```
<!DOCTYPE GreenstoneDirectoryMetadata [
  <!ELEMENT DirectoryMetadata (FileSet*)>
  <!ELEMENT FileSet (FileName+,Description)>
  <!ELEMENT FileName (#PCDATA)>
  <!ELEMENT Description (Metadata*)>
  <!ELEMENT Metadata (#PCDATA)>
  <!ATTLIST Metadata name CDATA #REQUIRED>
  <!ATTLIST Metadata mode (accumulate|override) "override">
]>
```

Greenstone Archive Format: Example document

```xml
<?xml version="1.0" ?>
<!DOCTYPE GreenstoneArchive SYSTEM
"http://greenstone.org/dtd/GreenstoneArchive/1.0/GreenstoneArchive.dtd">
<Section>
  <Description>
    <Metadata name="gsdlsourcefilename">ec158e.txt</Metadata>
    <Metadata name="Title">Freshwater Resources in Arid Lands</Metadata>
    <Metadata name="Identifier">HASH0158f56086efffe592636058</Metadata>
    <Metadata name="gsdlassocfile">cover.jpg:image/jpeg:</Metadata>
    <Metadata name="gsdlassocfile">p07a.png:image/png:</Metadata>
  </Description>
  <Section>
    <Description>
      <Metadata name="Title">Preface</Metadata>
    </Description>
    <Content>
      This is the text of the preface
    </Content>
  </Section>
  <Section>
    <Description>
      <Metadata name="Title">First and only chapter</Metadata>
    </Description>
    <Section>
      <Description>
        <Metadata name="Title">Part 1</Metadata>
      </Description>
      <Content>
        This is the first part of the first and only chapter
      </Content>
    </Section>
  </Section>

</Section>
```

Greenstone archive format

Document type definition (DTD)

```
<!DOCTYPE GreenstoneArchive [
  <!ELEMENT Section (Description,Content,Section*)>
  <!ELEMENT Description (Metadata*)>
  <!ELEMENT Content (#PCDATA)>
  <!ELEMENT Metadata (#PCDATA)>
  <!ATTLIST Metadata name CDATA #REQUIRED>
]>
```

Document database

demo/ index/ demo.ldb

```
[42]
<section>HASH863cfd85c90056aeb66bc3.7.1
----------------------------------------------------------------------
[HASH863cfd85c90056aeb66bc3.7.1]
<doctype>doc
<hastxt>1
<Title>National park restoration in Chad: luxury or necessity ?
<docnum>42
----------------------------------------------------------------------
[HASH863cfd85c90056aeb66bc3.8]
<doctype>doc
<hastxt>0
<Title>Developing World
<childtype>VList
<contains>".1;".2
<docnum>43
----------------------------------------------------------------------
[CL1]
<doctype>classify
<hastxt>0
<childtype>VList
<Title>Subject
<numleafdocs>17
<thistype>Invisible
<contains>".1;".2;".3;".4;".5;".6
----------------------------------------------------------------------
[CL1.2]
<doctype>classify
<hastxt>0
<childtype>VList
<Title>Communication, Information and Documentation
<numleafdocs>1
<contains>".1
<mdoffset>1
```

Agenda

- ❖ Interface customization
- ❖ Customizing parts of a page
- ❖ Greenstone in your organization
- ❖ Installing the web library (demo)
- ❖ Administration pages
- ❖ User management
- ❖ Restricting access to documents
- ❖ Building a collection
- ❖ The dreaded black screen
- ❖ More on building
- → ❖ Multilingual support
- ❖ Language interfaces

Multilingual support

❖ Greenstone works well with English documents, but what about documents in other languages?
  – Greenstone uses Unicode internally, so it can handle these too
  – Searching and browsing works (with a few limitations)

Multilingual issues: build-time

❖ Source filenames should be in ASCII (English)
❖ GLI metadata entry: Unicode only
  – Install Microsoft's IME (Input Method Editor) for your language

Multilingual issues: build-time

❖ GLI: Make sure you're using a Unicode font that includes your language!

Multilingual issues: build-time

❖ Source file content:
  – Unicode, *or*
  – Encoding that Greenstone can convert to Unicode

Multilingual issues: run-time

❖ Browser must have the correct fonts installed
  – Not usually a problem these days (Unicode fonts)
❖ Browsing:
  – Many of Greenstone's classifiers don't support Unicode (GenericList is one that does)

  – Sorting may be incorrect

Multilingual issues: run-time

❖ Searching:
  – Enter search terms in Unicode (using IME), *or*
  – Select native encoding from the Preferences page, and enable "Auto-select" or select the native encoding in the browser

  – Stemming: English only

Agenda

* ❖ Interface customization
* ❖ Customizing parts of a page
* ❖ Greenstone in your organization
* ❖ Installing the web library (demo)
* ❖ Administration pages
* ❖ User management
* ❖ Restricting access to documents
* ❖ Building a collection
* ❖ The dreaded black screen
* ❖ More on building
* ❖ Multilingual support
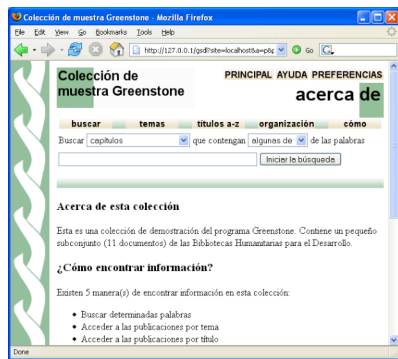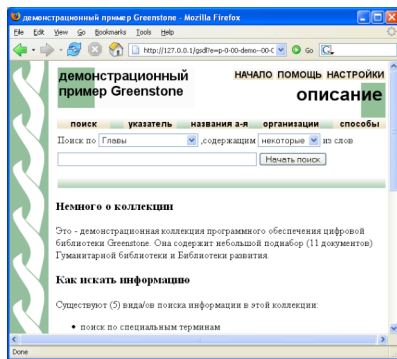* ➡ ❖ Language interfaces

Language Interfaces

❖English:

Language Interfaces

❖Spanish:

## Language Interfaces

❖Russian:

Language Interfaces

❖ Text fragments must be translated

❖ Button images must be generated

❖ Maintenance nightmare!
* – About 750 text fragments in the user interface
* – No one person knows all languages
* – Little/no overlap between developers and translators
* – Developers and translators geographically dispersed
* – Translations must keep pace with software developments
* – We need an automated system to help…

Language Interfaces

❖ Greenstone Translator Interface (GTI)
* – Web-based system
* – Keeps track of which text fragments have changed, or are new
* – Translators are presented with pages containing the text fragments to update or translate
* – When submitted, these are saved to a separate installation of Greenstone
* – Mistakes in existing translations can be corrected
* – Translators can download the translated files
* – Translators must log in to add/edit translations

GTI

❖ Select a file to work on:
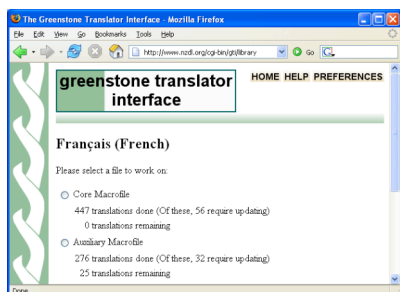
GTI

❖ Updating a translation:

Français (French) -- Core Macrofile

447 translations done (Of these, 56 require updating)

0 translations remaining

Correct existing translations

Download file

View this file in action

Updating text fragment *document.etsec*

English	Last updated 06-Mar-05
CONTRACT\ nTEXT	

Français (French)	Last updated 23-Jan-04
RÉDUIRE\nLE TEXTE	

SUBMIT

Can you help?

❖ Translating the Greenstone interface is easy!
  – No special technical skills are required
  – We generate the button images for you

❖ Translation is done at your own pace

❖ Promotes your language

❖ Improves Greenstone

❖ E-mail to get an account!

Join the Greenstone community

❖ Mailing list
❖ Wiki
❖ User-contributed documentation
❖ Documentation, documentation, documentation
  – new exercises, documented example collections (public-domain material?), fix up exercise text …
❖ Give workshops!
❖ Resources:
  – we undertake contracts
  – DLConsulting (John visiting Hawaii in August)
❖ Become a Friend of Greenstone
❖ Become a Regional Support Center