Digital Resource Management: From Data to Knowledge

Greenstone Digital Library Software

Campbell Collections
Durban, South Africa
23 July 2009
Amos Kujenga
akujenga@gmail.com

An overview of managing the digital cycle.
This presentation gives an overview of Greenstone Digital Library software as a Content Management System. It also looks at its pros and cons, and if time permits a demonstration will be given at the end.
Digital Library collections building and distribution software

Freely available, Open Source Software (OSS)

- Issued under the terms of the GNU General Public License

First developed in 1996 by the New Zealand Digital Library Project at the University of Waikato’s Computer Science Department
Aims to empower users, particularly in universities, libraries, and other public service institutions to build their own digital libraries (to avoid being read-only societies)

The Greenstone developers hope that the software will encourage the effective deployment of digital libraries to share information and place it in the public domain
Current production version – 2.82

Distribution and Promotion Partners

- UNESCO
  - Information For All Programme (IFAP)
  - Training Workshops
- Human Info NGO, Antwerp, Belgium
  - Distribution of humanitarian collections on CD-ROM
- eIFL.net
  - Through establishment and coordination of the Southern African Greenstone Support Network (SAGSN)
Greenstone Requirements

Server is platform independent
- Runs on Linux, Windows, Solaris, MAC OS X

Java Run-time Environment (JRE)
- Version 1.4 or above required
- Required for Greenstone Librarian Interface (GLI)
- Must be installed before installing Greenstone

ImageMagick
- Needed in order to build image collections
Greenstone Requirements

C++ compiler
• Visual Studio or GCC
• Needed only if you wish to compile the source code

Perl
• Installed automatically

A Web Browser
• For Readers to access digital library/collections
Greenstone Features

Search
- Full text
- Wildcard
- Boolean
- Cross-collection

Collections can be served on CDs/DVDs
- Comes with installer
- Allows user to run from CD or install on Hard Disk

Automatic creation of access structures
- Browse
- Search
Greenstone Features

- **Access controls**
  - Password protection at document/collection level

- **Easy to install**

- **Graphical User Interface for building/managing collections** - Greenstone Librarian Interface (GLI)

- **Supports multiple file types**, e.g., DOC, PDF, MP3, JPG

- **Interoperable**
Greenstone Features

Easily customisable

z39.50 protocol support
- Supports protocol both as client and as server

Supports multiple languages
- In User/Reader Interface
- In Librarian Interface

Supports several metadata sets
Greenstone Features

Customisable metadata
- Includes metadata editor

Metadata-driven digital libraries

Provides several ways to build collections

Flexible browsing facilities
- Hierarchy
- List
- Generic
Greenstone Features

- Large User Base
- Large Technical Support Base
Multi-format digital collections

Bibliographic Collections

Institutional Repositories

Text
Image
Audio
Video
**Greenstone Use: The Facts**

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributed via SourceForge since</td>
<td>11/2000</td>
</tr>
<tr>
<td>Average downloads per month since then</td>
<td>4500</td>
</tr>
<tr>
<td>Currently running at</td>
<td>4500</td>
</tr>
<tr>
<td>Proportion of downloads that are documentation</td>
<td>33%</td>
</tr>
<tr>
<td>Proportion of downloads that are software</td>
<td>67%</td>
</tr>
<tr>
<td>Proportion Windows / Linux / Mac / Source</td>
<td>82% / 10% / 3% / 5%</td>
</tr>
<tr>
<td>Number of people on Greenstone email lists</td>
<td>750</td>
</tr>
<tr>
<td>Number of countries represented</td>
<td>82</td>
</tr>
<tr>
<td>Number of messages per month (excluding spam)</td>
<td>165</td>
</tr>
</tbody>
</table>

Source: http://www.greenstone.org/factsheet
Support Networks

Southern African Greenstone Support Network
http://www.eifl.net/cps/sections/services/eifl-foss/greenstone

Greenstone Support for South Asia
http://greenstonesupport.iimk.ac.in

Greenstone Support for South America
Under construction
User/Technical Support Resources

Documentation

On Greenstone Site:
- Manuals
- Demo collections
- Wiki
- Tutorial Exercises
- FAQs
User/Technical Support Resources

Help

Online Help in:

• User Interface
• Librarian Interface
Mailing Lists

- Global Greenstone Users’ List:
  - For general discussions about Greenstone
  - greenstone-users@list.scms.waikato.ac.nz
  - http://list.scms.waikato.ac.nz/mailman/listinfo/greenstone-users
  - Includes list archive

- Greenstone Developers’ List
  - For discussions about development of Greenstone
  - greenstone-devel@list.scms.waikato.ac.nz
  - http://list.scms.waikato.ac.nz/mailman/listinfo/greenstone-devel
  - Includes list archive

- SA Greenstone Discussion List
  - For general discussions about Greenstone
  - sagreenstone@lists.eifl.net
  - Includes list archive
Greenstone 3

A complete redesign and reimplemention of Greenstone

Research version of Greenstone (not recommended for production use)

When complete, will retain all advantages of Greenstone 2
Greenstone 3

Incorporates all features of current system

Backward compatible (can run and build existing collections without any need for modification)

Further Information: http://www.greenstone.org/greenstone3-home
SA Greenstone Support Network (SAGSN)

Based on feasibility study carried out in 2005

Co-ordinated by eIFL.net as part of the eIFL-FOSS Programme (http://www.eifl.net/cps/sections/services/eifl-foss/greenstone)

Sponsored by Koha Foundation

Four National Centres established in Lesotho, Malawi, Namibia, and Zimbabwe
Strengthen the existing national centres and networks in southern Africa through centre specific investments and activities

Start longer term support of other African countries' use of Greenstone through the expertise located in existing national centres through an array of means ranging from remote assistance to on-site visits as well as facilitating locally organised training

Extend the network by facilitating candidate national centres in organising start-up training for their country along with neighbouring ones and setting up support

Actively seek the network's expertise and support to grow the numbers of operational Greenstone digital collections, benefiting both creation and access of African local content

Work towards a self-reliant and self-sufficient network of national centres in a broad landscape of libraries, archives and museums by 2010
The term “open source” refers to software that is made available for free and that includes the original (uncompiled) source code used to create it so that users can freely modify it.

http://www.opensource.org/docs/definition.php
Open Source Software (OSS)

Advantages

Low cost
- Initially free – “As free as a free puppy”
- No huge annual subscriptions

Can be freely distributed
- Can increase capacity / expand usage as needed

Can be customised to fit local needs
Open Source Software (OSS)

Other OSS For Building Digital Libraries:

- DSpace
  - http://www.dspace.org

- Eprints
  - http://www.eprints.org

- Fedora
  - (Flexible Extensible Digital Object and Repository Architecture)
    - http://www.fedora.info
Greenstone Interfaces

User/Reader Interface

Librarian Interface

[Images of Greenstone software interfaces]
Greenstone User/Reader Interface

- Accessible via any web browser
- Multilingual Support (Available in +30 Languages)
- Highly Customisable
Greenstone User/Reader Interface

Default Interface

Customised Interface
Greenstone Librarian Interface (GLI)
Greenstone Librarian Interface (GLI)

- **Download** files/web pages for your collection
- **Gather** source files from your local machine
- **Enrich** content with metadata
- **Design** searching & browsing facilities

- **Create** (or build) the collection
- **Format** the output of your collection
Greenstone Librarian Interface

- Graphical Interface used by librarians to build/maintain collections
- Easy to use
- Java based
- Available in English, French, Spanish, and Russian
- Includes a metadata editor
- Remote management of collections via Remote GLI
At least one metadata set is associated with each collection

Dublin Core used as default set

Greenstone automatically extracts metadata from source files through plugins during the collection building process

New sets can be defined using the built-in Greenstone Editor for Metadata Sets (GEMS)
Greenstone Editor for Metadata Sets (GEMS)

Open/Edit Existing Metadata Set

Create New Metadata Set
Open Archives Protocol for Metadata Harvesting (OAI-PMH) support

- Can harvest documents and include them in collections
- Incorporates server that can serve collections over OAI-PMH

METS

- Any collection can be exported to METS
- Greenstone can take in documents in METS form
- Greenstone METS Profile approved by the METS Editorial Board

DSpace

- Export collections to DSpace
- Import collections from DSpace
Access Controls

Password protection at:

- Collection-level
- Document-level
Customisation

Look and Feel
• Done mainly through Macro files, Images, & Format statements (in GLI)

Program functionality
• Can be altered by modifying C++ and Perl source code that is freely available

Requires some level of knowledge of HTML

Cannot be done using an HTML editor since the pages are not static
Enables Greenstone to be used to build Institutional Repositories

Allows Users to add documents to collections via the User/Reader Interface
The Upside of Greenstone

- Open Source Software
- Large User Base
- Large Technical Support Base
  - Fair amount of documentation
  - Several active mailing lists
- Easy to use/get going
  - Automatically creates installer
  - Gives user choice to run from removable media or install to hard disk
- Distributable Collections
  - Converts files to HTML while keeping the original at the same time
- Long-term document preservation features
  - Supports OAI-PMH
  - METS
- Interoperable
### The Upside of Greenstone

**Multilingual**
- Documents & Interface

**Multiplatform**
- User not tied to one platform

**Fairly mature**
- +12 years old now

**Minimal System Requirements**
- Ideal for use in developing world

**Highly configurable**

**General Purpose DL software**
- Can be used in many DL applications

**Flexible Metadata Support**
- Supports many sets
- Allows editing of existing sets/creation of new ones
# The Downside of Greenstone

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item-centric</td>
<td>• Difficult to combine objects</td>
</tr>
<tr>
<td>Complicated URLs</td>
<td>• Users can only be added one-by-one</td>
</tr>
<tr>
<td>Authentication</td>
<td>• Users can only be added one-by-one</td>
</tr>
<tr>
<td>User Administration</td>
<td>• Users can only be added one-by-one</td>
</tr>
<tr>
<td>Access Controls</td>
<td>• Modes that control access levels in GLI can be changed by anyone</td>
</tr>
</tbody>
</table>
The Downside of Greenstone

- Workflow process is implied (in GLI) & not built-in
- Consistency & Precision required upon entry of data
  - Case sensitivity, effect of space characters, abbreviations & acronyms
- Collection locking
  - Each collection may only be open (for editing) by one person at a time otherwise work done may be lost
- General Purpose DL software
  - Can be used in many DL applications
References


Rajasekharan K (2007), “Building up a Digital Library with Greenstone”, Kerala Institute of Local Administration


http://www.greenstone.org
QUESTIONS?
Thank you!