

# Greenstone Digital Library Software

Greenstone is a suite of software for building and distributing digital library collections. It provides a new way of organizing information and publishing it on the Internet or on CD-ROM. Greenstone is produced by the New Zealand Digital Library

Greenstone is an open-source, multilingual software, issued under the terms of the GNU General Public License for building and distributing digital library collections. The aim of the Greenstone software is to empower users, particularly in universities, libraries, and other public service institutions, to build their own digital libraries. It provides a new way of organizing information and publishing it on the Internet or on CD-ROM in the form of a fully-searchable, metadata-driven digital library.

The software provides a way of organizing information and publishing it on the Web in the form of a fully-searchable, metadata-driven digital resource. It can also be run in a non-networked environment (standalone), operating from removable media such as a USB Flash Drive and DVD. Greenstone has been developed and distributed in cooperation with UNESCO and the Human Info NGO in Belgium.

**AIM:** According to Greenstone website ([www.greenstone.org](http://www.greenstone.org)), “The **aim** of the software is to empower users, particularly in universities, libraries, and other public service institutions, to build their own digital libraries. Digital libraries are radically reforming how information is disseminated and acquired in UNESCO's partner communities and institutions in the fields of education, science and culture around the world, and particularly in developing countries. We hope that this software will encourage the effective deployment of digital libraries to share information and place it in the public domain.”

**VERSIONS** - There are two major versions of the software.

Greenstone3 is a complete redesign and reimplementaion of the original digital library software we developed back in 2000 (Greenstone2). It incorporates all the

features of the previous software, and is backwards compatible. An import 'Greenstone2 collection' feature is provided in the Greenstone Librarian Interface (GLI) to assist in the transition to the new software for existing users. Greenstone3 is written in Java and takes advantage of many web technologies—such as XML Transforms (XSLT), and the Java Authentication and Authorization Service (JASS)—which have emerged since we first started developing digital library software. This makes the digital library software highly customizable and extensible in the functionality it delivers. In contrast, Greenstone2 was written in C++ and had to make heavy use of bespoke formats and syntax we devised ourselves. This was because of the lack of available web technologies at the time, meaning not only was the functionality it provided more limited, the bespoke syntax meant users were completely reliant on us providing documentation detailing how things worked.

Greenstone3 is under active development, and is recommended for download.

## FEATURES

The GSDL software is an open source software available from the New Zealand Digital Library ([www.nzdl.org](http://www.nzdl.org)) under the terms of GNU, general public library license. Greenstone CD-ROMs have been published by the United Nations and other agencies for distribution in developing countries. Some of the features of the software are:

- NT/2000) and Unix (Linux Sun OS) any of these systems can be used as a web server.
- The administration function builds in it enables the items to authorize new users to build collection, protect documents so that they can only be accessed by registered users on presentation of password.
- It builds collection with effective full-text searching and metadata-based browsing facilities. Collection containing millions of documents, up to several gigabytes can be built. Full-text searching is fast because compression is used to reduce the size of the indexes and text users can browse the list of authors, titles, date, class no., etc
- Plug Ins can be written to accommodate new document types, the collection can contain pictures, music, audio, video clips, etc. It also supports multilingual documents.

- Collection can be updated and new one brought online any time without bringing down the system.
- The Z39.50 protocol is supported for accessing external servers and for presenting Greenstone collection to external clients.

## **TECHNICAL FEATURES OF GSDL -**

- Multiplatform user friendly application
- Interoperability
- Independent librarian and user interfaces
- Supports variety of Metadata formats
- Supports variety of Document formats
- Supports multiple Languages

## **LIBRARIAN INTERFACE -** GLI supports six basic activities:

### **Downloading Files from the Internet**

The Download panel allows you to download files from the internet using a variety of protocols:

- **Web**: downloads web pages and files via HTTP and FTP.
- **MediaWiki**: downloads web pages and files via HTTP from a MediaWiki website.
- **OAI**: downloads metadata records (and optionally documents) from an OAI-PMH (Open Archives Initiative) server.
- **Z39.50**: downloads MARC records that match a particular search criterion from a Z39.50 server.
- **SRW**: downloads MARCXML records that match a particular search criterion from a Search/Retrieve Web (SRW) server.

### **Collecting Files for Your Collection**

The Gather panel is where you determine which files will be in your collection. You can add any files on your computer, and they will be copied into your collection. You can also rename and remove files from your collection here.

## Enriching Your Collection with Metadata

After adding documents to your collection, you can manually add metadata for the documents in the [Enrich panel](#). Metadata can be added for individual documents, multiple documents, and folders. You can add or remove metadata sets (by default the Dublin Core, Greenstone, and extracted metadata sets are selected).

## Configuring Your Collection

The [Design panel](#) dictates how your documents will be handled (using document plugins), and how users will interact with your collection (using search indexes, partition indexes and browsing classifiers).

### *Plugins*

[Plugins](#) tell Greenstone how to process the files in your collection. Every document must be processed by a plugin. There are two types of plugins: **Document Plugins** and **Metadata Plugins**

## Searching

When determining how users will be able to perform search queries on your collection, you have three things to consider: the search indexer, the search index(es), and the partition index(es).

For every collection, you can decide which **search indexer** to use—**MG**, **MGPP**, or **Lucene**. The search indexer is what parses and indexes the text. Basically, the indexer determines *how* search indexes will be built, and each indexer works a bit differently.

**Search Indexes** specify *which* parts of the text are searchable. You can assign any number of search indexes to a collection. You can build indexes on the full-text of the documents, on specific metadata fields (like titles or authors), and on any combination of fields. Indexes can be searched for particular words, combinations of words, or phrases, and results are ordered according to how relevant they are to the query.

**Partition indexes** can be used to split your collection into subsections for search purposes. If your collection includes documents in multiple languages, you can

create subsections based on language. You can also create partitions based on the value of any metadata field(s).

**Cross-collection search** (searching multiple collections at once), is automatically enabled for Greenstone3, and can be specified on the Format panel in Greenstone2.

## **Browsing**

**Browsing** involves lists that the user can examine: lists of authors, titles, dates, hierarchical classification structures, and so on. Users can browse interactively around lists, and hierarchical structures that are generated from the metadata that is associated with each document in the collection.

The ability to browse collections is handled by **browsing classifiers**. You choose which browsing classifiers will be created for each collection. You can create a browsing classifier for any metadata field (or combination of metadata fields) in your collection. Every classifier you create results in an additional tab on the navigation bar of your collection's website.

Configuring classifiers allows you to specify whether the documents will be displayed on one page or several (and how they will be split into sections, e.g. a specific number of documents per page or a page for each letter).

## **Producing Your Collection**

After adding in your documents, providing metadata, and configuring plugins, indexes, and classifiers, you are ready to build your collection. This is done in the Create panel. As the collection builds, information about the build (including what plugin processes each document) will be displayed.

The Create panel presents many build options (covered in-depth on the collection building page). One of the most used is the ability to schedule builds (only available in expert mode).

Any time you make changes on the **Gather**, **Enrich**, or **Design** panels, you must rebuild the collection before the changes will take effect.

## Customizing your collection's appearance

Most of your time in the Format panel, though, will likely be spent in the Format features section, where you write format statements for your collection. Format statements dictate the format of the content of individual documents pages; of documents in the list of search results; and of documents in the browsing classifiers.

## Additional Options and Features

The GLI File menu provides some important options and features. **Preferences** allows you to change the interface language; the **mode**, which effects what functions are accessible in the GLI; and connection settings (like the web path to Greenstone and proxy settings).

**File Associations** determine which application is used to open each document type when you open documents in the GLI.

## SOFTWARE REQUIREMENTS –

- OS Windows/Linux
- Apache web server/IIS
- PERL
- Java 2 Runtime Environment 'version 1.4.2\_03'
- Web browsers—Netscape Navigator or Internet Explorer
- GSDL

**STEPS FOR INSTALLATION OF GREENSTONE** - The following steps are required for installation:

1. Install the Java 2 Runtime Environment (latest version).
2. After installing J2RE, go for GSDL folder choose setup gsd1 2.70.
3. The Install Shield Wizard will begin the installation of GSDL software. Click .
4. Accept all the terms of license agreement by clicking on button.
5. Choose setup Language. English (US) is the default.
6. Choose the type of installation (Local Library).
7. Set the admin password.

Once installation is completed, to start your Greenstone system click on the Start button, open the Program menu, and select Greenstone Digital Library. This brings you a dialogue box: just click Enter Library. This automatically starts your Internet browser and loads the Greenstone Digital Library home page.

## **ADVANTAGES OF GSDL**

- It is based on FOSS platform and has active community supporting it.
- It is Multi-platform application and can run on various operating system platforms, including Windows (any version), Linux, Sun Solaris, and Mac OSX. It is available in both binary (executable) and source code form for the Windows (all versions), Linux, and Mac OS X operating systems and in source code form for other operating systems (Unix).
- A Greenstone Collection can be served on the World Wide Web or it can be exported to a CD-ROM and accessed from the CD-ROM or local hard disc without the need for Internet connectivity.
  - Greenstone can build indexes from full text documents and also metadata associated with these documents. It supports creation of indexes for various metadata fields, either automatically extracted or manually assigned.
- It uses Perl-scripting, MG(PP) or Lucene for indexing, Apache (or built-in webserver), XML, which are proven technologies
  - Greenstone lets you build collections of multimedia documents such as audio, video, and pictures accompanied by textual description or metadata to allow searching and browsing.
  - UNICODE compliant facilitating building, searching and browsing documents in any Unicode-compliant language.
  - Separate modules are available for different uses: – JAVA-based interface for management – Web-browser based access to collections – CLI client : remote collection building
- Multi-metadata (with editor)
- Practical GLI interface for editing/managing GSDL.
- Plug-ins for most document formats also available as well as for crosswalks for ISIS, Dspace, e-mails, MARC, MARCXML.

**RAKHI SINGH**