Chapter 9

Greenstone Digital Library Software Evaluation
http://www.greenstone.org

Introduction

Greenstone is a digital library software developed by the New Zealand Digital Library Project at the University of Waikato, New Zealand. It is a suite of software for building and distributing digital library collections that provides a way of organizing information and publishing it on the Internet and or on removable media i.e. (CD-ROM/DVD).

The aim of the Greenstone is to empower users, particularly universities, libraries and other public service institutions throughout the world, to build their own digital library collections in the field of education, science and culture. The software is distributed in cooperation with UNESCO and the Human Info NGO. UNESCO has been promoting Greenstone since 2000 and since then arranging user testing, helps with internationalization and workshops on Greenstone. The first version of Greenstone was made available during 1997 under the GNU Public License[1].

Greenstone has been very popular in developing countries as it is very easy to create collections in Greenstone as well as it has one of the important feature of exporting collections on CD-ROM, this has been found very useful where Internet connectivities are not yet available. It has helped libraries to carry whole collection on CD-ROM and the CD-ROM

collections operate on a standalone PC under Windows 3.X, 95, 98, and NT or Linux installations are there[2].

Greenstone's first version was released during 1999. For the present study current version of Greenstone i.e. version 3.0.3 which was released during 24th October 2007 was selected for the evaluation, but the installation of this version was not successful. It was overall found that the Version 3 of Greenstone is still not very stable version though it is a complete redesign and reimplementation. It was found that version 2.75 is a stable version. This version was made available during September 2007. Hence for the present evaluation Greenstone version 2.75 was selected. The following sections lists different observations after having installation of version 2.75.

9.1 Content Acquisition

9.1.1 **Which document types can be added in the software (books, reports, journal articles, lecture notes, technical reports, thesis, images, audio/video files, data set files etc.?)**

Greenstone supports to add any types of documents such as from books, reports, journal/newspaper articles, lecture notes, technical reports, thesis, images, audio/video, visual art files as well as any other documents.

9.1.2 **Does the software have capability to define the domain/scope of digital library for whom it is intended for?**

Yes, Greenstone has capability to define the domain/scope of digital library for whom it is intended for. The information about scope can be defined while creating every collection in Greenstone.

9.1.3 **If yes, what level of users it is intended for?**

Greenstone can be used to organise any type of digital document. It is not intended for a special group of users. It can be used for any type of digital library.

9.1.4 **Which digital document file formats does the system support(proprietary as well as open source)?**

Greenstone supports to upload variety of documents and their file formats. It converts submitted file formats to Greenstone supported file format plugins[3] so that the

documents added on Greenstone are converted into the Greenstone supported plugins and viewable to end users.

The following are the supported plugins in Greenstone:-

ZIP, GAP, Text, HTML, Email, PDF, RTF, Word, PS, Image, ISIS, NUL, Metadata XML, MP3, OpenDocument, LOM, BibText, Book, CONTENTdm, CSV, DB, Dspace, Excel, FOX, Favorites, GML, HB, Index, LaTex, MARC, MARCXML, METS, MediaWiki, MetadataCSV, OAI, OggVorbis, PPT, Pagedlmg, Procite, RealMedia, Refer, Rog, SRC, StructuredHTML, W3Img After importing documents into Greenstone of any of the above types Greenstone converts these documents into a simple HTML-like format known as GML (for "Greenstone Markup Language") which includes any metadata associated with the document.

9.1.5 **Does the system carry out duplicate checking while uploading the data? What happens if the same digital document is submitted multiple times?**

Yes, Greenstone supports duplicate checking. While uploading the document if the same document is selected again the system gives a message as 'document already exists'.

9.1.6 **How does the system allow to upload digital objects, whether normal digital objects or compressed digital objects?**

Greenstone supports to upload the digital objects with normal size as well as compressed digital objects.

9.1.7 **Does the software support to upload digital objects from existing URL?**

Yes, Greenstone supports to upload digital objects from existing URL.

9.1.8 **Does the system support distributed/community based acquisition?**

Greenstone allow to upload documents through its web interface where "collector" option is there and end user should have the login and password to use this option which helps then to acquire material from distributed places on the web but this option is not working satisfactorily with version 2.75.

9.1.9 **Does the software support metadata/content import? In which file format data is required for importing?**

Greenstone has importform.pl and import.pl program which helps to import metadata from OAI repository. Metadata is imported in XML format.

9.1.10 Does the system support to import bulk metadata/contents of digital objects?

Yes, Greenstone supports to import bulk metadata of digital objects from OAI repository.

9.1.11 Does the software support import of automatic metadata for the digital objects that are added into the repository?

Yes, Greenstone has limited automatic metadata recognition feature such as it identifies the date of the object, encoding, file format details, file size, language, number of pages and source details for each object that is imported into the Greenstone repository.

9.1.12 Does the software support metadata/content export? In which file format data is exported?

Yes, Greenstone supports to export metadata into METS, MARCXML format.

Greenstone supports to export data into METS format as well as can ingest documents in METS form. One more important feature of Greenstone is any collection can be exported to DSpace ready for DSpace's batch import program, and any DSpace collection can be imported to Greenstone.

9.1.13 Does the system support export of bulk metadata/contents of digital objects?

Yes, Greenstone supports to export data into MARCXML, Dspace, METS and GA format.

9.1.14 Does the system allow adding past versions of the digital document files? No

9.1.15 Does the software support its own accession number for each document that gets added into the repository?

Greenstone assigns a unique object identifier or OID to each digital document when it is added into the repository.

9.1.16 It it necessary to know whether software support weed out policy? No

9.1.17 Does the system support to delete items from the collection, move objects from one collection to other collection, cross-list objects across different collections?

Greenstone supports to delete items from the collection but does not support to move digital objects from one collection to other collection.

9.1.18 Which persistent identifier scheme does the system support?

Greenstone does not support any persistent identifier scheme.

9.1.19 Does the system's persistent URL have capability to handle an object's change in location and state in future with same identifier number? No

9.2 Content Management

9.2.1 Does the system have proper work flow in place which handles different submission processes?

Greenstone does not support different submission processes.

9.2.2 If the system supports to have proper work flow for all actions into the repository then whether submission supports following options:

9.2.2.1 System allows to create overall system administrator Yes

9.2.2.2 System allows to submit documents by users via web

Yes, Greenstone supports to allow to submit documents by users via web through "collector" utility but it does not work properly.

9.2.2.3 System allows to set up submission rules No

9.2.2.4 System can define accept/reject policy for all documents that gets added into the repository No

9.2.2.5 Allows multiple collections within same installation of the system Yes

9.2.2.6 Home page for each collection/community Yes

9.2.2.7 Allows to set different policies for different collections No

9.2.2.8 Allows to set different authorization policies for different collection/items added into the repository No

9.2.2.9 Allow to edit metadata submitted by users No

9.2.2.10 Sends email notification to users/submitters/authors No

9.2.2.11 Sends email notification to metadata reviewers No

9.2.2.12 Sends email notification to reviewer No

9.2.2.13 Sends email notification to administrators No

9.2.2.14 Allow users to review completed content No

9.2.2.15 Allow users to review uncompleted content No

9.2.2.16 Allow content administrator to review submissions No

9.2.2.17 View pending content administration tasks No

9.2.3 Does the system support to upload documents into the repository through Librarian's interface as well?

Greenstone supports to upload documents through librarian interface and through web interface which is called as "collector" and only those user's can submit items through Web UI who have login and passwords.

9.2.4 Does the software provide easy way for adding/editing of records for Librarian as well as for the remote user? Are insertion and deletion of records easy?

Greenstone's librarian's interface as well as web based "collector" is very easy for data entry. Library staff can easily add/edit records added into the repository. Basically collection building in Greenstone works with local server or it helps to build collections on remote web server with the help of "collector" if appropriate permissions are given.

9.2.5 Are the supported file formats well documented in the system? Can new file formats be added or removed by the end user?

Greenstone supported file formats are well documented in the system but new file formats cannot be added in to the software. It identifies each file format through different plugins which are available in Greenstone.

9.2.6 Does the repository software have any capabilities, such as access control lists, Internet address filters, etc., that limit who is allowed to submit items in the repository? No

9.2.7 **Does the repository software maintain audit logs that identify by whom and when documents are submitted into the repository?** Yes

9.2.8 **Does the system support knowledge organization systems such as ontologies, thesauri, classification systems and taxonomies for users while adding documents via web in the repository?**

Yes, Greenstone supports to classify each digital object that is added into the repository as well as Greenstone supports to generate authority files of all Dublin Core 15 elements.

9.2.9 **Does the system support digital objects to be grouped into organized structure (i.e. linear/hierarchical/graphical structure)?** Greenstone supports digital objects to be grouped into hierarchical structure.

9.2.10 **Does the system support to generate authority files?** Yes

9.2.11 **Does system allow digital objects to be member of multiple collections i. e. content mapping?** No

9.2.12 **Does system allow users to modify (add/delete/update) metadata/complete documents added within the repository?** Yes

9.2.13 **Does the system support to show strength of each collection?** No

9.3   Metadata Submission and Support

9.3.1 **Does the software support to add descriptive metadata (such as author, title, subject, publisher, copyright, year of publication etc.) for each digital object that gets added into the repository?**

Greenstone supports to enter all Dublin Core 15 elements which generally cover all descriptive metadata fields such as title, creator, format, date of creation , subject coverage for each document that is added into the repository as well as it supports to customize all metadata fields.

9.3.2 **Does the software support to add administrative metadata (such as technical aspects of digital documents, source of information(when and how it was created), rights management, what is file size, in what file format digital document is in, what is needed to view digital document etc.) for each digital object that gets added into the repository?**

By default system allow to enter descriptive metadata but it can be configured to enter structural metadata.

9.3.3 **Does the software support to add structural metadata (such as description, owner, data type, data deposited, version number, date of last revision, how compound objects are put together, how pages are ordered to form chapters) for each digital object that gets added into the repository?**

By default system allow to enter descriptive metadata but it can be configured to enter administrative metadata.

9.3.4 **Does the software support adding/editing/deleting metadata fields?** Yes

9.3.5 **Are help messages given in each metadata field while entering data?** Yes

9.3.6 **Does the metadata entry have at least one field mandatory? What are the mandatory fields in each software?**

No, Greenstone does not have any metadata entry field as a mandatory as Greenstone automatically extracts some of the metadata information from the document if collection is built without doing any metadata entry.

9.3.7 **Does the software support Unicode character set for metadata entry?** Yes

9.3.8 **Does the software allow to create and manage a variety of different metadata schemas to meet different needs across the subject disciplines ?** Yes

9.3.9 **Can the metadata fields or formats be customized?** Yes

9.3.10 **What type of Metadata Schemas are supported by the software default?**

Greenstone supports following metadata sets:-

- New Zealand Government Locator Service Metadata Standard version 2.1 (nzgls)

- RFC 1807 Metadata Element Set, version TR-v.2.1 (rfc1807)

- Qualified Dublin Core Metadata Element Set, version 1.1:Reference Description (dc)

- Development Library Subset Example Metadata (dls)

- Greenstone Metadata Set (gs)

- Australian Government Locator Service Metadata Element Set version 1.3 (agls)

9.3.11 Does the software have default metadata entry templates? Yes

9.3.12 Are the metadata formats documented? Yes

9.3.13 Does the software support different interface for metadata entry and whether customized data gets added into the system? Yes

9.3.14 How does the software verify that submitted objects and metadata are correct?

There are no ways by which software can verify submitted objects and metadata are correct or not.

9.3.15 Are there automated checks of the metadata, such as to verifying that a date entered into a field is really a date string? No

9.3.16 Does the repository software have any means to verify that objects or metadata have not been tampered with, such as checksums or digitally signed checksums? No

9.3.17 Does the software support real time updating and indexing of accepted contents?

No, every time new document is added into the repository, the collection has to be rebuild then only it is available through web interface. Greenstone has now introduced 'minimum rebuild' concept which helps to index only those documents which are added recently into the Greenstone collection.

9.3.18 Does the software allow for metadata extensibility and complexity?

Yes, Greenstone supports to create customized metadata set for all the documents that are entered into the software. Hence it is possible to have extensible metadata for different collections in Greenstone.

9.3.19 Does the software support metadata versioning? What happens to old versions of metadata or content objects when they are replaced by new versions or modified in some way? No

9.3.20 **Does the repository software support provenance metadata? What is included in the provenance?** Yes

9.3.21 **Does the software support metadata crosswalk?** Yes

9.3.22 **Does the software support "thesaurus" building? Does it specify which thesaurus standard is used? Which subject heading DDC, DC or LC is used for thesaurus building?** No

9.3.23 **Does software have ontology support?** No

9.3.24 **How does the repository verify file types, for example, if JPEG images is submitted to the repository, does the software verify that it is really a valid JPEG, or does it just blindly take the submitters word?**

Yes, Greenstone's supported plugin will identify the uploaded file and select appropriate plugin to attach to the file.

9.3.25 **While exporting metadata whether software support checksums?** No

9.3.26 **It is necessary to know whether the software support RAP protocol?** Yes

9.4 Classification

9.4.1 **Does the system allow digital objects to be grouped into classification/subjects?**

Yes the software supports to enter class number. It uses Dublin Core metadata field where subject and keyword and class number can be added.

9.4.2 **Which classification system does the software support?**

System supports to allow to enter any classification numbers.

9.4.3 **Does it allow to group documents as per the class number ?**

Yes, the documents are grouped as per the class numbers.

9.4.4 **It is also necessary to know whether system supports hierarchical browsing of collections through class number.** Yes

9.5   Information Search & Retrieval

9.5.1   Browsing

Greenstone supports to browse all collections using all the metadata sets that are supported by Greenstone and that are used while collection building. Greenstone supports rich browsing facility such as manually linking parts of the documents together and building explicit indexes and tables of contents. Greenstone supports browsing by subjects, publications by titles a-z, publications by organization, publications by how to listing and browsing collections chronologically.

9.5.1.1 **Author/Creator/Contributor** Yes

9.5.1.2 **Title of the document/Title of the book/Title of the article etc.** Yes

9.5.1.3 **Issue Date/Date of Publication** Yes

9.5.1.4 **Collection** Yes

9.5.1.5 **Communities** No

9.5.1.6 **Subject browsing** Yes

9.5.1.7 **Publisher wise browsing** Yes

9.5.1.8 **Table of Contents browsing** Yes

9.5.1.9 **Multi-dimensional Browsing** No

9.5.2   Searching

9.5.2.1 **Searching content via different metadata fields such as author, title, subject, publisher etc.** Yes

9.5.2.2 **Full text searching** Yes

9.5.2.3 **Boolean (AND, OR, NOT) searching** Yes

9.5.2.4 **Basic search** Yes

9.5.2.5 Advanced search and retrieval feature Yes

9.5.2.6 Does the system allow to search the database putting boolean operators such as !, |, &, ~ , +, −, etc. as well as "must contain", "should contain" "may contain", "not contain", as operators in place of AND, OR, and AND NOT Boolean operators respectively. Yes

9.5.2.7 Truncation/Wild card searching: Yes

9.5.2.8 Exact words/phrases searching: Yes

9.5.2.9 Proximity searching Yes

9.5.2.10 Stemming Yes

9.5.2.11 Fuzzy searching Yes

9.5.2.12 Phonetic searching No

9.5.2.13 Case sensitive or case insensitive while searching Yes

9.5.2.14 System should support term weighting Yes

9.5.2.15 System should support search history option , reuse of query, query save option. Yes

9.5.2.16 Boosting the term Yes

9.5.2.17 Range searching Yes

9.5.2.18 Use of thesaurus or permuted index for searching No

9.5.2.19 Expand search Yes

9.5.2.20 Lateral searching No

9.5.2.21 Search support for special collections Yes

9.5.2.22 Searching via table of content and classification codes Yes

9.5.2.23 Ability to browse subject/author authority files Yes

9.5.2.24 PIC variant searching (plurals, international spelling variants, and compound words) No

9.5.2.25 Support multilingual search and retrieval Yes

9.5.2.26 **Allow refinement of searches with qualifiers** Yes

9.5.2.27 **Ability to group: prior to search, user can delimit pending results' listings, and group results by item characteristics (type, format)** Yes

9.5.2.28 **Spell checker support** No

9.5.2.29 **Refine searches: Once results are retrieved, users have the ability to focus and reduce number of items retrieved** Yes

Greenstone also support a special search feature such as search within sections of a document or search with section headings, paragraphs.

9.5.3   Sorting

The system should support sorting results by

9.5.3.1 **Author** Yes

9.5.3.2 **Title** Yes

9.5.3.3 **Issue Date** Yes

9.5.3.4 **Relevance** Yes

Greenstone supports to sort records as per the end users' requirements.

9.5.4 **Which search engine tool kit is used for searching the database?**

Greenstone uses Lucene search engine tool kit for searching. Indexing in Greenstone is done with "Managing Gigabytes (MG) indexing tool kit. MG provides document level indexes, and compression of the source documents. MGPP extends MG to also provide word-level indexes, enabling proximity and fielded searching .

9.5.5 **How is the recall and precision ratio while searching the documents?**

Since the system supports searching via metadata fields as well as full text searching the recall is high via word searching and precision is high if the metadata fields are used for searching the database.

9.5.6 How is the speed of the retrieval of the documents?

Speed of the retrieval of the documents is very fast in Greenstone.

9.5.7 Does the system provide help for the users while searching digital documents? Yes, but with very less search help feature.

9.5.8 How the results are displayed to the users? Whether users can download contents after searching? How are results displayed such as metadata, filetype, file size?

Search results gets highlighted in Greenstone. End user cannot download contents after searching. Only titles are displayed while showing search results.

9.5.9 While viewing the results does the system provide access control? Yes

9.5.10 Does system support streaming audio/video option? Or it allows only to download documents? Whether browser can support any other document viewing? Does the end user needs to have a plugins and software for viewing particular documents?

If streaming server is set up Greenstone supports streaming audio/video option. There are couple of open source streaming server's available on Internet with few modifications in Greeenstone it is possible to set up streaming server with Greenstone.

9.5.11 Does the system support advanced features such as my binders, virtual bag, shelving etc.

Greenstone has only one advanced feature such as my binders.

9.5.12 Whether system supports unified search? (Unified search is the support for finding items that are scattered among a distributed collection of information sources or services, typically involving sending queries to a number of servers and then merging the results to present in an integrated, consistent, coordinated format)? No

9.5.13 Whether system supports federated search?(Federated search allow users to search multiple digital libraries with a single query from a single user interface. The user enters a search query in the portal's interface's search box and the query is sent to every other digital library software)?

Greenstone has Z3 9.50 server hence it supports to search through other Greenstone repositories which are available on network either through OAI-PMH or through Z3 9.50 server.

9.6 Access Control, Privacy and Management

9.6.1 **How users are managed in the software?**

Greenstone supports to add different users through web interface. This tool is called as 'collector'[3]. Creating users were explored on 'collector' but while uploading documents it did not work satisfactorily.

9.6.2 **Does the system support to add/edit/delete user profiles?** Yes

9.6.3 **Does the system support to define different roles for different users?** No

9.6.4 **Does the system keep detailed information of each user that registers in the repository such as name, email, phone number, address etc.?** No

9.6.5 **How the permissions are given to users who have registered into the repository?**

This option is not working properly with 2.75 version.

9.6.6 **Does the system have provision to identify active and inactive members and take necessary action against inactive members?** No

9.6.7 **User Roles**

9.6.7.1 **Does system allow to create different groups for different users** No

9.6.7.2 **Does system allow to create different groups of users** No

9.6.7.3 **Does system allow definition of different user groups** No

9.6.7.4 **Does system limit access by role** No

9.6.7.5 **Does system allow collection to be customized for each role/group** No

9.6.8 **Does the system allow access limits at different levels?**

    (a) File/object level No

    (b) Collection level Yes

    (c) User/Group level No

9.6.9 **How passwords are administered in the system?** This option not working properly.

9.6.10 **Does the system assign passwords for each user?**

Greenstone allows to create passwords for different users through its WEB Interface but it is not functioning properly in version 2.75

9.6.11 **Does the system allow users to select passwords?** This function is not working properly.

9.6.12 **Does the system have mechanism to retrieve forgotten password?** No

9.6.13 **How is the level of access for Inside/outside users(such as outside users can access only metadata and abstract level access and internal members can access full texts as well as metadata)?**

Greenstone shows all collections along with their metadata if they are meant for public access.

9.6.14 **Does the system have any mechanism to provide access to the digital library via**

(a) IP source address filtering No

(b) Proxy filtering No

(c) Credential-based access No

9.6.15 **Are access denials flagged in any special manner by the software? Does the software keeps access logs and how accessible are the logs to either human or machine processing and interpretation?**

Greenstone keeps event and access logs in the Greenstone installation and they are accessible to the system administrator.

9.6.16 **Does the system support access to the collection for a specific period?** No

9.7 Authentication and Authorization

Authentication and authorization is not supported in Greenstone hence none of the criteria for authentication and authorization were considered for evaluation in Greenstone.

In Greenstone there is mention of incorporating an authentication scheme that can be explored to control access to certain facilities[4]. It is used to restrict people who are allowed to carry certain actions, it also allows documents to be protected on an individual basis so that they can only be accessed by registered users on presentation of password but however when this feature was explored it did not work successfully.

9.8   Interoperability

9.8.1 **It is necessary to know whether the software supports Z3 9.50 protocol and OAI-PMH protocol?**

Greenstone supports Z3 9.50 protocol for both accessing the external servers and for presenting Greenstone's collections to external clients as well as it supports OAI-PMH.

9.8.2 **Which harvesting software is used by the Digital Library Software for metadata harvesting?**

Greenstone uses its own OAI software called as oaiserver.

9.8.3 **SRU/SRW support** Yes

9.8.4 **Does the software have openURL support?** No

9.9   Ease of deployment of each software

9.9.1   Software Installation

9.9.1.1 **Packaging and installation steps**

While evaluating Greenstone latest version of Greenstone i.e. Version 3.03 was selected for evaluation but it was found that version 3.03 is still not stable version hence the earlier version of Greenstone i.e. Version 2.75 was selected for installation. Greenstone installation is very easy. There are no installation steps for Greenstone. Its installation file takes care of everything and installs software without any difficulty.

9.9.1.2 Automatic installation script

Greenstone has automatic installation script. Among all the software considered for evaluation Greenstone found to be the simplest software for installation. The automatic script takes care of finding allied software and completes the installation.

9.9.1.3 Disk space used 111 MB

9.9.1.4 Time required to install each software

Greenstone is very trivial to install it hardly takes any time.

9.9.1.5 Documentation help available and useful for installation

Greenstone wiki page has help available on installation of the software.

9.9.1.6 Environments needed to set for installation of each software

Greenstone requires JRE environments to be set.

9.9.1.7 Operating systems (Linux/Windows/Solaris/MacOS/GNU/AIX) on which software can be installed

Greenstone can be installed on Operating System: All 32-bit MS Windows (95/98/NT/ 2000/XP), All POSIX (Linux/BSD/UNIX-like OSes), OS X

9.9.1.8 Programming languages used

Greenstone is developed using C++ and JAVA and Perl

9.9.1.9 Database supported at back end

Greenstone 2.75 uses GDBM database i.e. Gnu Database Manager which contains an entry for each document, with its OID, its internal MG document number, and metadata.

9.9.1.10 Web Server used

Greenstone version 2.75 uses Apache/IIS web server.

9.9.1.11 Java servlet engine needed

Greenstone version 2.75 uses JRE version.

9.9.2 System Support/Maintenance

9.9.2.1 Documentation/manuals

Greenstone has its documentation available at http:// wiki.greenstone.org /wiki/ index.php/ Manual

9.9.2.2 **Mailing lists/discussion forums**

Greenstone has two mailing lists one is Greenstone-users archive and other is Greenstone development archive. greenstone-users @lists.scms.waikato.ac.nz & greenstone-devel @ lists.scms.waikato.ac.nz

9.9.2.3 **Wiki pages**

http://wiki.greenstone.org/wiki/index.php/greenstoneWiki

9.9.2.4 **Help desk support**

http://sourceforge.net/forum/forum.php?forum_id=38148

9.9.2.5 **Ease of system administration (ability to configure for different uses)** No

9.9.3 Hardware/Software Requirements

9.9.3.1 **What are the minimum hardware requirements for the repository software?**

Greenstone requires minimum P III or PIV machine for installation.

9.9.3.2 **How actively is the software maintained by its owners or contributors?**

Since 1997, Greenstone is continuously bringing out its new versions. A major change is made in Greenstone version 2 and Greenstone version 3 is available as a beta testing.

9.9.3.3 **Does the software have procedures in place to monitor and receive notifications when hardware technology changes are needed?** No

9.9.3.4 **Does the software creators have a process to stay current with the latest operating system security fixes?** Yes

9.9.4 Security

9.9.4.1 **Does the system have security methods such as:**

    (a) Data encryption No

    (b) Digital signatures No

9.9.4.2 **Server security (Does software has secure database connection?)** No

9.9.4.3 **Ability to restrict access repository at item level (For e.g. to view metadata but not content)**

Greenstone can restrict access at collection level only.

9.9.5 System Level Security

9.9.5.1 **Fixity - such as checksums or digests** No

9.9.5.2 **Management of user accounts and rights to specified locations within the DL** No

9.9.5.3 **Logging and auditing events**

Yes, Greenstone supports to log access and event logs in its installation as a txt file.

9.9.5.4 **Does the software support any security if OS doesn't have any firewall implemented?** No

9.9.6 General features related to technical aspect

9.9.6.1 **Does the software have any back end maintenance?** No

9.9.6.2 **Does the software have RSS support?** No

9.9.6.3 **Does the software have easy mechanisms to upgrade the software from old versions to new versions?**

Yes, it is easy to upgrade Greenstone from old version to new version. Greenstone version 3 also has a backward compatibility support.

9.9.6.4 Does the software support migration of data from one repository software to another repository software?

Yes, Greenstone supports to migrate data from DSpace repository. It has DSpace plugin which helps to understand DSpace data and imports that data into Greenstone [5]. It also has CDS-ISIS plugin which helps to bring CDS-ISIS data into Greentone.

9.9.6.5 Does the system support any ranking algorithms for the documents added in the repository? No

9.9.6.6 Are there any known issues/bugs in the system?

One of the main drawback of Greenstone is for every item added build command is required to be run otherwise those documents are not available to the end users. If there are millions of documents build process will take long time like a day or two.

The "collector" option available for web interface does not work properly.

There is no "log out" option on web user interface if a user logs in to the system. This is one of the major bug in Greenstone.

9.9.6.7 Whether a novice user can easily publish content ? Yes

9.9.6.8 How often user needs to use database schema for making any change?

Greenstone database schema is not easily accessible to end user.

9.9.6.9 Whether user can reorganize database? No

9.9.6.10 Does system support automated content acquisition, harvesting and automatic metadata generation, automatic subject indexing/classification?

Yes, Greenstone supports automatic metadata generation if no metadata is entered into the system.

9.9.6.11 Does the system support to store metadata records separately from the actual contents?

No, metadata and actual contents are stored in the same directory. Metadata is stored in XML format.

9.9.6.12 Does the software support storing varied file formats? Yes

9.9.6.13 Does the system keep the original file's name, size and created date?

No, Greenstone changes name of the file while storing it in Greenstone.

9.9.6.14 While using the software if there are any errors occur, does the system allow to submit the error report? No

9.9.6.15 How many developers are working for the software?

There are 9 developers working for the development of the software.

9.9.6.16 How many users are using the software/strength of community?

As with all open source projects, the user base for Greenstone is unknown. It is distributed on SourceForge, a leading distribution centre for open source software. Witten[6] found that its user base hails from 70 countries and the reader's interface has been translated into 45 languages. Downloads from SourceForge have risen from a steady (for many years) 4,500 times a month to 6,500 over the last two years.

9.9.6.17 Does the software provide its history on the web page for evidence of sustainability and vitality?

Yes, Greenstone has also been facing a great crunch of funding. In New Zealand funding for open source efforts is not available thought the worth of the open source is widely recognized[7].

9.9.6.18 Total number of downloads of the software?

Total number of downloads of version 2.75 was not available on SourceForge site.

9.9.6.19 Is the software really under Open Source License terms and conditions?

Yes, Greenstone is under GNU General Public License (GPL).

9.9.6.20 Are there any costs involved for using OSS-DL? No

9.9.6.21 Does the repository software have a clear policy regarding software upgrades? Yes

9.9.6.22 Whether the software has a long-term leadership and consulting services? Yes

9.9.6.23 After releasing the software whether it was tested and whether it has integrated new features or not?

Yes, software is constantly bringing out new features.

9.9.6.24 **How is the mailing list used by the users?**

Greenstone has two mailing lists and users are extensively using these mailing lists. Since 2002 Greenstone's mailing list is extensively used. The following table (Table 9.1) shows number of requests sent on Greenstone's user's mailing list as observed on 1st April 2008.

Table 9.1: Greenstone User's Mailing List Data

Year	Number of requests sent on Greenstone's User's mailing list
2003	683
2004	899
2005	1015
2006	1054
2007	1079
2008,1st April	226

9.9.6.25 **Whether software has conducted any workshops, training sessions on local, regional, national and international levels?**

Yes, Greenstone is a popular software and many workshops have been conducted on Greentone at National and International level. In India IIM, Kozhikode is identified as a Greenstone support team for South Asia.

9.9.6.26 **Whether software is built according to open standards?** Yes

9.9.7 Architecture of the Software

9.9.7.1 **How is the architecture and design of the software: whether it is simple or complex?**

The key components in Greenstone architecture are the receptionist and the collection server. The receptionist provides the user interface, dispatching requests to the

appropriate collection server or servers, and aggregating results for display back to the user. The collection servers manage the contents of collections.

The Greenstone protocol depends on the server configuration. In the default, single-server configuration, the receptionist and collection servers are combined into a single executable, and the protocol is simply function calls. This is called the null protocol. However, there is also a Corba[8]-based protocol which allows for a truly distributed environment. Greenstone has a Java-based client that can utilize the Corba protocol.

In addition, third-party systems have been developed that use the Corba protocol.

Another significant part of the Greenstone architecture are plug-ins and classifiers. Plug-ins parse and extract metadata from imported documents. Greenstone supports a large selection of standard plug-ins for processing different file types. Custom plug-ins can also be written. Classifiers create the Greenstone browse indexes, and Greenstone comes with a number of predefined classifiers, such as for hierarchical or list-based browsing.
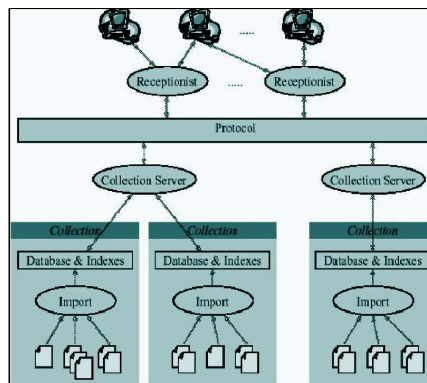

Figure 9.1: Greenstone Architecture

9.9.7.2 Does the architecture support separation between different local parts and put into different machines (e.g. centralised/distributed database, relational/object-oriented database management system, different components of the directory to distribute to different machines) and the transport model (protocols for communication between the system and the user interface or between system components). No

9.9.7.3 Does system support grid architecture? How grids can help to disaster recovery of data?

No Information about grid architecture was found.

9.9.8 **Scalability: It is necessary to know whether software is scalable in terms of storage and retrieval?**

Greenstone can support large collections up to several gigabytes of text, associated images, movies etc. But yet it is not tested by anybody about having millions of documents on Greenstone.

9.9.9 **Extensibility: Extensible, to allow new elements (collections, data types, services, etc.) to be easily added to the digital library.**

Yes, Greenstone supports extensibility and which can be obtained through plugins available in Greenstone as well as classifiers[9]. These modules can be used to enhance the extensibility.

9.9.10 **Storage: Does the system support data compression ? For e.g., in case of multimedia collections the size of the files (for e.g. video/audio files) would be very large, hence system should support data compression/decompression techniques. Does the system support to store compressed file formats?**

Greenstone supports to store compressed file formats to ensure best use of storage.

9.9.11 **Backup/Restore facility**

9.9.11.1 **Does software support backup function, which contains metadata associated with access controls, customized files, as well as full texts of the documents added in the repository?** No

9.9.11.2 **Does the software explicitly require any particular backup strategy, or does it just rely on system-level backup plans, like periodic disk backups to tape?**

There is no backup strategy in Greenstone. It relies on system-level back up plans such as just copying the "collect" folder which contains all collections added into the repository.

9.9.11.3 **In the event of a disaster, what recovery plans are supported by the software?** No

9.10 User Friendly Interface

9.10.1 **Whether end user can easily customize the user interface "look and feel"? Whether the end user can change the header, theme, footer, overall web interface of the**

system?

Yes, end user can easily customize the user interface "look and feel. Greenstone allows to change header, footer and overall web interface of the system. Greenstone also has one of the unique feature of creating a text only interface for the visually impaired users.

9.10.2 **Can user interfaces be localized in their language?** Yes, currently Greenstone reader's interface is available in Arabic, Armenian, Bengali, Catalan, Chinese (both simplified and traditional), Croatian, Czech, Dari, Dutch, English, Farsi, Finnish, French, Gaelic, Galician, Georgian, German, Greek, Hebrew, Hindi, Hungarian, Indonesian, Italian, Japanese, Kannada, Kazakh, Kirghiz, Latvian, Maori, Marathi, Mongolian, Polish, Portuguese (both European and Brazilian versions), Pushto, Romanian, Russian, Serbian, Slovak, Spanish, Thai, Turkish, Ukrainian, Vietnamese Librarian's interface is available in Arabic, English, French, Marathi, Spanish, Romanian, Russian Chinese (simplified), Latvian, Vietnamese in progress.

9.10.3 **How is user interface for the experienced or non experienced users?**

Greenstone's user interface is easy to use for experienced and non experienced users.

9.10.4 **How is the user interface designed for the use of librarian or for the use of users?**

The user interface for librarian is different and for users it is different.

9.10.5 **Whether system provide web based interface for all its functionalities?** No

9.10.6 **Does the system provide online help?** Yes

9.10.7 **Does the system provide multilingual access support?** Yes

9.10.8 **Does the system allow users to submit feedback on system problems?** No

9.10.9 **Does the system notify user's current state of the task?** No

9.11 Usability

9.11.1 **What user skill levels are expected to use the web interface?**

Greenstone's Librarian's interface is very easy to use as well as web interface is very easy to use which has helped Greenstone to get more acceptability all

over the world.

9.11.2 **Is the user able to navigate and find his/her ways to information quickly and easily?**
Yes

9.11.3 **How are the help features provided with the software? Does help feature allow a user to feel comfortable within a digital library? Does the help feature provide general as well as technical answers to the problems?**

Greenstone help feature is not very extensively provided. It gives a general outline of help but does not cover all aspects as required by a newcomer if he/she does not have any knowledge of searching the digital library. Greenstone does not provide any text on error messages.

9.11.4 **Does the user need to know which player or viewer is necessary to view multimedia or other digital objects from the system?**

No, the web browser automatically picks up the relevant player if it is installed by the end user.

9.11.5 **Does the user interface provide facility to give user opinion on the web as well as does the software support user to fill up online user feedback form?** No

9.11.6 **Does the system support to show usage statistics such as:**

    i. Usage patterns No

    ii. Use of materials Yes

    iii. Usage statistics Yes

    iv. Who uses what, when Yes

    v. For what reasons/decisions No

9.11.7 **Transaction Log Analysis**

9.11.7.1 **Does the system keep error log files?**

Yes, Greenstone keeps usage logs, error logs as well as initialization logs [10]. These files are kept in /Greenstone/etc directory.

9.11.7.2 **Does the system keep user logs?** No

9.11.7.3 **Does the system keep track of query's submitted to the system? Does it store any query log?** No

9.12 Copyright/Policy Issues

9.12.1 **Does the repository software have any means to manage, store, or enforce contract agreements for all the digital documents that are added into the repository?** No

9.12.2 **Does the software have any capabilities, such as access control lists, Internet address filters, etc., that can be used to enforce copyright or access restrictions? Can different restrictions be applied to different objects in the repository?** No

9.12.3 **If repository ingests digital content with unclear ownership/rights, does it have policies addressing liability and challenges to those rights?** No

9.13 Advanced Features

9.13.1 **Whether software has long term leadership and consulting services?** Yes, Greenstone has long term leadership and consulting services. Though, actual number of users of Greenstone are not known, there are many digital libraries which are using Greenstone and are available on local servers of the organization. Greenstone has been supported by UNESCO and University of Waikato hence there will be support for Greenstone in future.

9.13.2 **Does the system provide citation data?** No

9.13.3 **Does the software have defined road map for the future?** Yes, Greenstone has a defined roadmap for the future developments and details about the future developments are available at http://trac.greenstone.org/wiki/TracRoadmap

9.13.4 **Does the software support indexing of all documents that are added in the repository in a distributed way, without central control?** No

9.13.5 **Virtual Collection Support: Does the software support to generate virtual collections /special collections across several content providers?** No

9.13.6 **Load Balancing: Does the software has any mention of load balancing? There is no mention of load balancing on Greenstone's documentation or Wiki page.**

9.13.7 **Visualization: Does the software support visualization feature?** No

9.13.8 Personalization: Does the system support Personalization feature? No

9.13.9 Translation Service: Does the system provide translation service? No

9.13.10 Community Services: Does the software support to provide community services such as allowing members of the digital library community to exchange ideas, make announcements, write reviews etc.? No

9.13.11 Page Ranking: Does the software support page ranking such as supported by commercial digital libraries? No

9.14 Digital Preservation

Greenstone is not oriented towards long-term preservation hence none of the evaluation criteria related to digital preservation were valid and considered for the present evaluation.
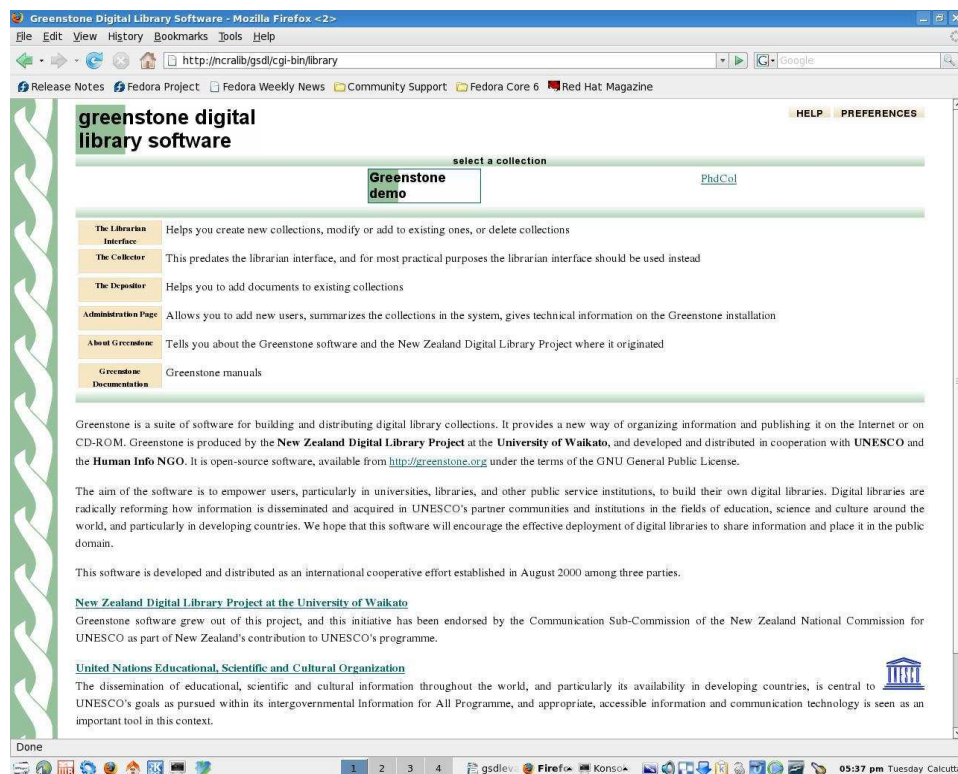

Figure 9.2: Greenstone Installation Screen

Bibliography

[1] Witten, Ian H. & Bainbridge, D. (2007). A retrospective look at Greenstone: Lessons from the first decade. *JCDL'07*, p. 147-156.

[2] Witten, I. H., McNab, R. J., Boddie, S. J., & Bainbridge, D. (2000). Greenstone: a comprehensive open-source digital library software system. *Digital Libraries*, p.113-121.

[3] Witten, Ian H. , Bainbridge, D., Paynter, G. & Boddie, S. (2002). The Greenstone plugin architecture. *JCDL'02*, July 13-17, p.285-286.

[4] Bainbridge, D., Thompson, J., Witten, I. H. (2003). Assembling and enriching digital library collection. *Digital Libraries Proceedings IEEE*, p.323-333.

[5] Witten, I. H. [et al.] (2005). A bridge between Greenstone and Dspace. *D-Lib Magazine, 11(9)*, September 2005, http://www.dlib.org/dlib/september05/witten/09witten.html. (Browsed on 10th January 2008).

[6] Witten, I. H., ibid., p.147.

[7] Witten, I. H., ibid., p.153.

[8] http://en.wikipedia.org/wiki/CORBA. (Browse on 10th April 2008)

[9] Witten, Ian H., Bainbridge, D., & Boddie, S. J. (2001). Power to the people: End-user building of digital library collections. *JCDL'01*, June 24-28, p.94-103.

[10] Buchanan, G., Bainbridge, D. & Don, J. J. (2005). A new framework for building digital library collections. *JCDL'05*, June 7-11, p.23-31.