

Release Note: Greenstone 2.81

Important changes in this release include, in no particular order:

wvWare updated

We have upgraded the version of wvWare included with Greenstone in the Linux binary releases to version 1.2.4. It was not possible to do the same for the MacOS binary distribution without shipping around 80Mb of additional libraries, so the MacOS binary release still comes with version 0.7.1. MacOS users may upgrade wvWare if needed by downloading it from <http://wvware.sourceforge.net/> and compiling it. Source users will also still get version 0.7.1 but may upgrade in the same way.

Stability and Bug Fixes

Some small problems in the recent 2.81rc release of Greenstone2 have been ironed out to create an overall more robust release. Some features have been restored (including realistic books and dynamic classifiers on Windows) and the LuceneWrapper.jar file has been put back into the release.

GDBM

Greenstone now comes with it's own modified version of gdbm which can auto-detect the format of database files (ldb or bdb format) and load either format on any of Greenstone's supported platforms. This achieves a better level of portability for Greenstone collections, and allows collections built on one operating system to be served from other operating systems supported by Greenstone.

Installer improvements

-
- * The new installer:
 - is now available in five languages
 - can now cope with being executed from a path with accented characters and other special characters. (There is still a known issue with chinese file paths, which can be avoided by simply running from c:\ or another path with no special characters.)
 - * The source release is now provided as a platform-independent zip / tar archive
 - * We now provide a simple binary zip / tar distribution for each platform

Vista Compatibility

-
- * When Greenstone is installed into the Program Files directory on Windows Vista, it is necessary to elevate to a higher access level to run GLI and the Greenstone server. As of this release, you can do this by launching the Start Menu items as an administrator. Just right-click the item and choose "Run as Administrator", or set "Run as Administrator" in the shortcut properties. Also, as a way to avoid having to elevate to higher access levels altogether,

the default installation directory is now inside the current user's home folder.

* Some Windows Vista users have experienced a problem when perl outputs warning messages in front of xml output passed to GLI. The error messages themselves are not fatal, but cause the xml to become invalid, resulting in a GLI crash. GLI has been modified to ignore these warning messages and thereby keep the validity of the xml.

New Start Menu icons and server image

Each item in a Greenstone Start Menu group is now has a unique icon so you can tell which is which at a glance. We have also added some links to the wiki and the Greenstone website for convenience. The old server image has also been replaced with the new GS2 Server logo.

Installer:

The Greenstone installer downloads are now created using only open source software, including Ant Installer (<http://antinstaller.sourceforge.net/>), Apache Ant (<http://ant.apache.org/>), our own search4j (based on launch4j <http://launch4j.sourceforge.net/>), and our own Greenstone release kits.

We felt it was essential to move away from our existing, closed-source installer suite as it was not in the spirit of Greenstone to rely on proprietary software, and it invited the unnecessary cost of keeping the suite up-to-date.

Using only open source software has also given us the freedom to customise the user experience of the installers. Where a feature was lacking in one of the open source packages, we have enjoyed the ability to "open the hood" and implement it.

With the development of the Greenstone release kits we are now able to generate releases of Greenstone automatically and unattended. This process takes just 20 minutes, and it has helped us a lot in the development and testing of new features of Greenstone. We have been able to set up a feedback loop, where code committed one day is included in an automatic nightly snapshot release made available for download on our website the next day. Our users have also benefited from being able to obtain up-to-date snapshot releases of Greenstone in between our official releases.

These daily snapshots are available through:

<http://www.greenstone.org/snapshots>

On Windows, the version number is included in the Start Menu which means that

multiple installations of Greenstone will have their own entries in the Start Menu. The installer no longer writes to the registry, so installation no longer requires administrator privileges.

ImageMagick is now bundled with Greenstone for binary web releases for all platforms (previously it was only provided on a CD-ROM release) and includes JPEG2000 support. The installer offers the option to install this or not. You can skip installing it if you already have ImageMagick previously installed.

Ghostscript is now bundled with Greenstone for binary web releases for Windows and Mac. The installer offers the option to install this or not. You can skip installing it if you already have Ghostscript previously installed.

Importing and Plugin Changes:

Plugin Restructuring: The plugins have been restructured. This has been done mostly for coding efficiency. The most noticeable change is that plugins have been renamed. xxxPlug now becomes xxxPlugin, and some have also had their names modified or expanded to be clearer. For example, DBPlug is now DatabasePlugin, PPTPlug is PowerPointPlugin. The complete list of new plugins can be seen at <http://wiki.greenstone.org/wiki/index.php/Plugins>. Old collections should still build okay in GLI, which will map the old plugin names to the new ones. Command line building should also work, but the configuration file won't be altered to use the new names.

A lot of work has been done trying to get Greenstone to work properly with non-ascii filenames. Two versions of the filename are stored, [Source] for displaying the filename, and [SourceFile] for linking to the file from a web page.

The -smart_block option (most commonly seen with HTMLPlugin) has been deprecated. Instead we have a first pass through the documents to find out which ones should be blocked.

MARCPlugin and MARCXMLPlugin can now map metadata to Qualified Dublin Core as well as Dublin Core.

MARCXMLPlugin now assigns metadata from metadata.xml files.

EMAILPlugin now saves binary attachments as binary files on Windows

ImagePlugin and PagedImagePlugin now offer the -cache_generated_images option to prevent thumbnails and screenview images being recreated each import.

OAIPlugin saves the metadata as Dublin Core rather than moving it all to the ex metadata set. This means that it won't show up in the GLI unless the files are exploded. Exploding is now available for OAI records.

ProCitePlugin no longer uses the pc namespace. Metadata is stored using the field names set in the workform definition, rather than using Field1Name and Field1Value etc..

-reversesort option to import.pl to be used with -sortmeta option to sort in reverse order.

Tidied up OID generation. OIDtype and OIDmetadata options are now available for each plugin as well as globally from import.pl. Hashing usually occurs on the original file, but some plugins specify that hashing should be done on the Greenstone XML document.

Building and Indexing Changes:

Improvements to Lucene incremental building. Lucene document identifiers now match Greenstone identifiers, rather than 1, 2, 3, thanks to DL Consulting. This improves search efficiency, and is another step towards true incremental building. Rebuilding with no new documents shouldn't break the index.

MGPP and Lucene building tidied up. 'allfields' now means combined searching over all specified indexes (not all document metadata) for both mgpp and lucene. 'metadata' will index all metadata, but no longer reindexes metadata that has already been specified. MGPP indexing over combined fields now works properly.

New collection configuration file option: infodbtype. Values include gdbm (the default), gdbm-txtgz, sqlite, mssql. This specifies what database system to use for the collection metadata database. Currently this needs to be added to collect.cfg by hand, as it is not available in GLI yet.

gdbm: the default, and is what has always been used previously. This is platform dependent. It transfers between Windows and Linux, but not to/from a Mac.

gdbm-txtgz: A zipped text version of the database. At runtime, the first time this collection is accessed, it will be unpacked and converted to the appropriate GDBM database using txt2db. This is a good choice if you are creating a collection to be used on another operating system, especially if one or more of those operating systems is a Mac.

sqlite: Platform independent database using SQLite. The new dynamic classifiers can be used with this database. Thanks to DL Consulting.

mssql: Windows specific database using MS SQL Server. See http://wiki.greenstone.org/wiki/index.php/Using_MSSQL_for_Collection_Database

for information about how to set this up. The new dynamic classifiers can be used with this database. Thanks to DL Consulting.

Depositor:

Collection editing using the Depositor is now enabled by default for the admin user.

The metadata fields offered for each document can now easily be customised from the GLI.

By default, the depositor will import and build each document into the collection when it is added. It can now be easily set up to just "deposit" the item without rebuilding.

See http://wiki.greenstone.org/wiki/index.php/How_to_use_the_depositor for more information.

GLI

A new right-click file option for binary files that go through a conversion process to be imported, such as PDF, Word and PowerPoint files. "Replace source document with HTML" will convert the original file to HTML and replace it in the collection. This means that the converted HTML (which may be unattractive) can now be modified in the collection. This is available for Remote GLI too.

A new right-click file option on the collection tree background to refresh the file view. This is useful if you have manually added files to the collection outside of GLI.

MetadataXMLPlugin has been moved 'below the line' in the Plugins panel so cannot be removed in GLI. If you are using GLI to add metadata, then you need this plugin.

The Export As option now only exports one collection at a time. Collections can be exported as GreenstoneMETS, FedoraMETS, MARCXML and DSpace archive.

'NavigationBar pulldown' added to the list of format options - this makes the navigation 'bar' a drop down list instead of a bar across the page.

Lots of work done on making the Download panel work better, including getting download processes to terminate when they are cancelled or when GLI is closed while they are still running.

New panel under the Format tab, "Depositor Metadata" which allows customisation of which metadata fields should be offered in the depositor for that collection.

CJK segmentation option has been added to GLI. This doesn't do proper word segmentation, but adds a space in between each CJK character. The character ranges have been expanded to work with Japanese and Korean. It is applied to metadata as well as document text.

Another new feature in Greenstone v2.81 is the ability to schedule the automatic rebuilding of a greenstone collection using the GLI. There's a new panel under the Create tab allowing the scheduling of automatic collection rebuilding. Scheduling will require some configuration in order for it to work.

The following link has more information on how to do this for each platform:

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

Greenstone Runtime:

New dynamic classifiers (if the collection uses sqlite or MSSQL as the collection database). These are generated at runtime so do not require the collection to be rebuilt for classifier changes. Another step towards true incremental building. See

http://wiki.greenstone.org/wiki/index.php/More_about_classifiers#How_do_I_use_dynamic_classifiers.3F

for more information. Thanks to DL Consulting.

Large code restructuring. Code is split between common-src, build-src and runtime-src. Makes it easier to use code in other projects, such as Greenstone 3 and running Greenstone on an iPhone/iPod Touch.

External links now default to going straight to the web page, rather than showing a warning page first.

Improvements to the oaiserver, thanks to DL Consulting. Resumption token support has been added, and it now validates against online validation tools.

Greenstone now compiles for Linux on 64-bit platforms. All programs are compiled natively for 64-bit, except for those that rely on mg/mgpp, which still need to be 32-bit compiled. As part of the configure/make process, the correct flags are automatically set, so from the developers point of view you compile up Greenstone on a 64-bit machine the exact same way you do for any other Unix-based system.

Greenstone should now compile with Microsoft Visual Studio 2005, thanks to DL Consulting.

FLI - Fedora Librarian Interface

The Fedora Librarian Interface (FLI) is very much like GLI in that it allows you to gather together your documents. It then uses Greenstone's ability to convert documents into the FedoraMETS format so that these can be ingested (put into) a Fedora repository. The process of building a collection using FLI exports documents into a repository maintained by a Fedora server rather than one maintained by Greenstone. In order to use FLI, you will need to have Fedora installed. More information on how to install Fedora and set it up to work with FLI, and how to run FLI itself, can be found in the documents:

1. Installing Fedora (2.2.1/3.0) and Fedora-related information
<http://wiki.greenstone.org/wiki/index.php/InstallingFedora>

2. Installing Fedora Generic Search (to provide full-text indexing and searching capabilities for a Fedora repository)
<http://wiki.greenstone.org/wiki/index.php/InstallingFedoraGSearch>

3. Running FLI, the Fedora Librarian Interface
<http://wiki.greenstone.org/wiki/index.php/RunningFLI>

Translations:

Thanks to:

John Rose, for help with English GLI help, French translations for GLI and Greenstone.

Maxime Rouast for Greenstone French translations

Celine Guimbertaud for GLI French translations

Yohannes Mulugeta and Abiyot Bayou for Greenstone Amharic translations

Kamal Salih for GLI Arabic translations

Gerhard Riesthuis for Greenstone Dutch translations

Mohan Raj Pradhan for GLI Nepali translations.

Diego Spano for translating the installer's interface into Spanish.

Xiaofeng Yu for translating the installer's interface into Mandarin.

Doris Jung for translating the installer's interface into German.

...and many other minor improvements and bug fixes

We want to ensure that Greenstone works well for you. Please report any problems to the Greenstone mailing list
greenstone-users@list.scms.waikato.ac.nz.