
Greenstone: Overview of features, capabilities, applications
Course material prepared by
Greenstone Digital Library Project
University of Waikato, New Zealand
and
National Centre for Science Information,
Indian Institute of Science, Bangalore


Agenda
What is digital library software?
What does Greenstone do?
Example Greenstone collections
Collection building
Documentation and help
Getting into details …

What is a "digital library"?

Collection of digital objects (text, video, audio) along with methods for access and retrieval, [user] and for selection, organization, and maintenance [lib]

- **Traditional user/librarian distinction is blurred**
- **Computers make information active**
- **Kitchens for knowledge preparation**
- **WWW ≠ DL!—organization, selectivity**
- **Nice Web site ≠ DL!—import new documents easily**

Data	*smallest discernible difference (change of state)*
Information	*difference between data and your expectations*
Knowledge	*accumulation of your set of expectations*
Wisdom	*the value attached to knowledge*





What are digital libraries for?

- Knowledge/content management
  - Manage and access internal information assets
- Scholarly communication, education, research
  - E-journals, e-prints, e-books, data sets, e-learning
- Access to cultural collections
  - Cultural, heritage, historical & special collections, museums, biodiversity
- E-governance
  - Improved access to government policies, plans, procedures, rules and regulations
- Archiving and preservation
- Many more …

DL Software: Key requirements

- Document types (book, journal article, lecture …)
- Document formats (text, PDF, Word, PS, …)
- Content acquisition (online and offline)
  - Metadata description, content tagging
  - Content uploading
- Indexing and retrieval
  - Structured/ full text indexing
  - Automatic metadata extraction
- Storage
  - Data compression
  - Efficient storage for metadata
  - Efficient location of metadata and documents
- Access and delivery
  - Structured search, browse, hierarchical browsing
  - CD-ROM distribution

DL Software: More requirements

- Scaling up – for large collections
- Multilingual support
- Access management and security
- Usage monitoring and reporting
- Standards compliance
  - XML, Dublin Core, Unicode
- Interoperation
  - OAI, Z39.50 compliance, MARC, CDS/ISIS, …

DL Software: Alternatives

- What are our expectations?
- Develop local web-based application?
- Adopt LA solution?
- Commercial DL solution?
- Adopt open source software?
  - Greenstone
  - Eprints
  - DSpace
  - (CDS/ISIS, Koha)



Agenda

- What is digital library software?
- What does Greenstone do?
- Example Greenstone collections
- Collection building
- Documentation and help
- Getting into details …

Example

Humanity Development Library

for sustainable development and basic human needs

- 160,000 pages
- 30,000 images
- 800 books
- 430 magazines
- 340 kg
- US$20,000
- CD-ROM
- US$6
- Win3.1x upward
- Stand-alone
- *and* intranet server
- Web browser user interface

Global Help Project, Antwerp (+ UN agencies)

Accessing a Greenstone collection

Be adventurous—Click liberally!


Book icon: read the corresponding book


Bookshelf: look at books on that subject


Open folder and view contents


View this document


Click on this to close the book

Reading the documents

Once you have found a document, you can



Text: expand or contract



Table of contents: expand/contract



Highlight search terms, or not

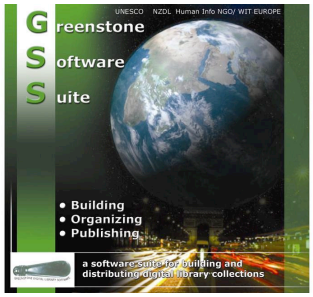


Open this page in a new window

What is the Greenstone software?

- Software suite for building, maintaining, and distributing digital library collections
- Comprehensive, open-source
- Developed by New Zealand Digital Library Project at the University of Waikato
- Distribution and promotion partners:
  - UNESCO
  - Human Info NGO, Belgium
  - NCSI, Bangalore; UCT, Cape Town; Dakar, Senegal; Almaty, Kazakhstan; …
  - You!

UNESCO: Distributing Greenstone DL software



Sustainable development

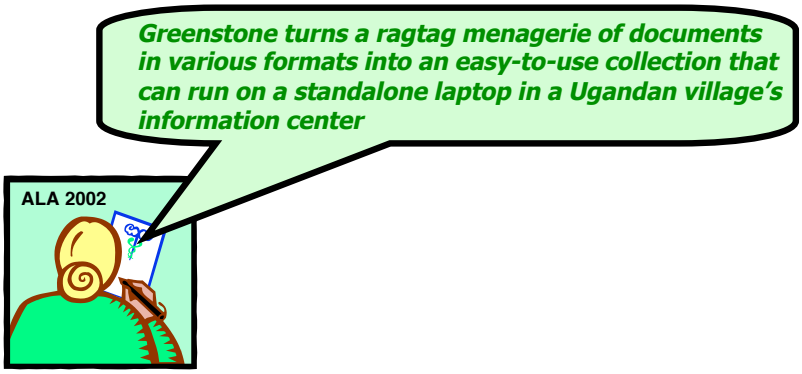
"Give a man a fish, feed him for a day
Teach a man to fish, feed him for life"

Greenstone software on CD-ROM

- **GNU licensed**
- **Fully documented … in English/French/Spanish/Russian**
- **Language interfaces … Arabic Chinese Czech … Thai Turkish**
- **Unix/Windows/Mac OS-X**
- **Trivial to install**
- **GUI interface for gathering, enriching, building …**
- **Serve collections on Web or write them to CD-ROM**
- **Document formats: HTML, Word, PDF, PS, plain text, e-mail**
- **Metadata formats: XML, DC, OAI, MARC, …**

download from *http://greenstone.org*

What we wanted



What we wanted

- "Collections" of digital material
- Individualized, depending on metadata etc
- Up to several Gb of text …
- … + associated images, movies, whatever
- Fully searchable
- Served on WWW, or published on CD-ROM
- Multi-platform (Unix + all Windows + Mac)
- Multi-format documents *and* metadata
- Multi-lingual: documents *and* interfaces
- Multimedia
- Metadata: standard *and* non-standard

Getting information

- Available at: www.greenstone.org
  - Software
  - Demo collections
  - FAQ
  - Examples
  - Documented example collections
- Documentation:
  - Installer's Guide, User's Guide, Developer's Guide, From Paper to Collection
- Mailing lists:
  - Greenstone User's List
  - Greenstone Developer's List

Greenstone, Libraries and Open Access

"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."

—from www.greenstone.org

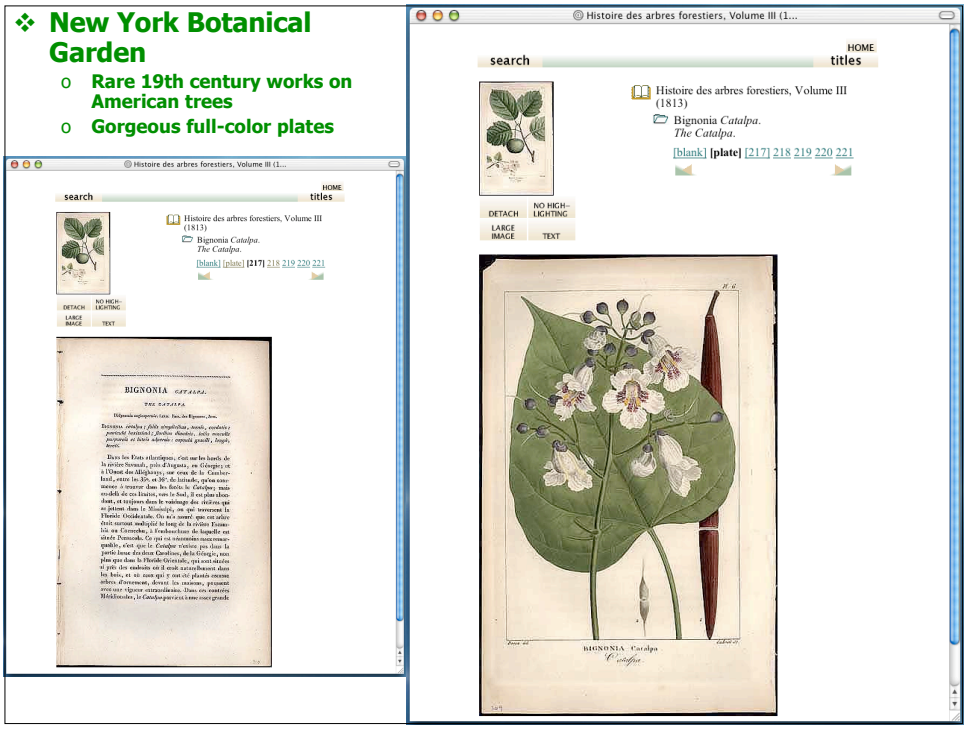


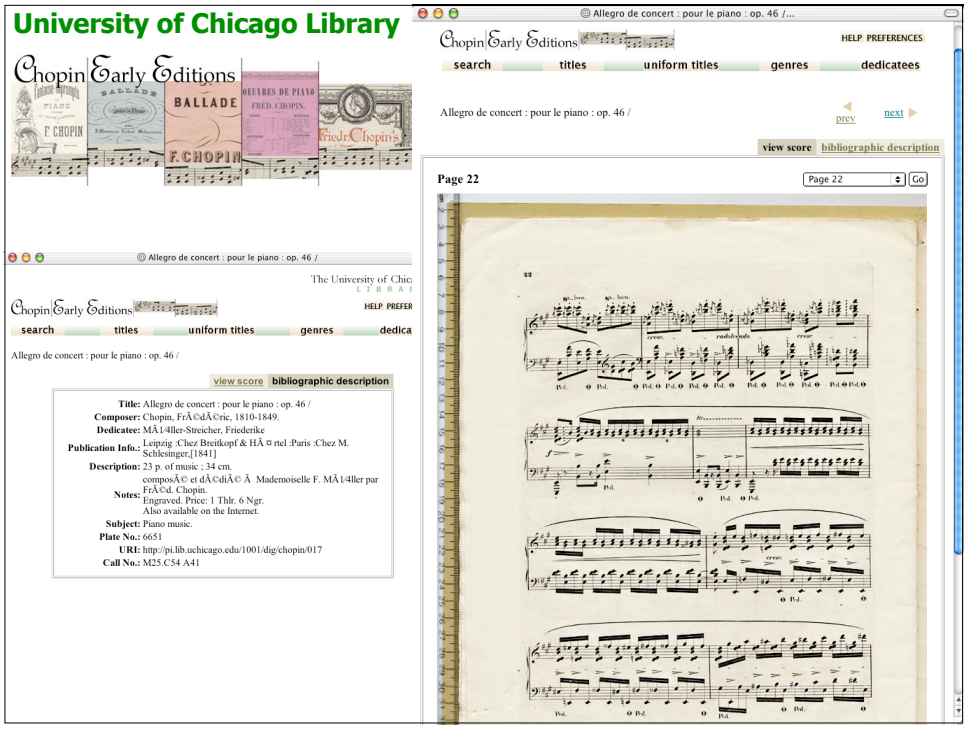
Agenda

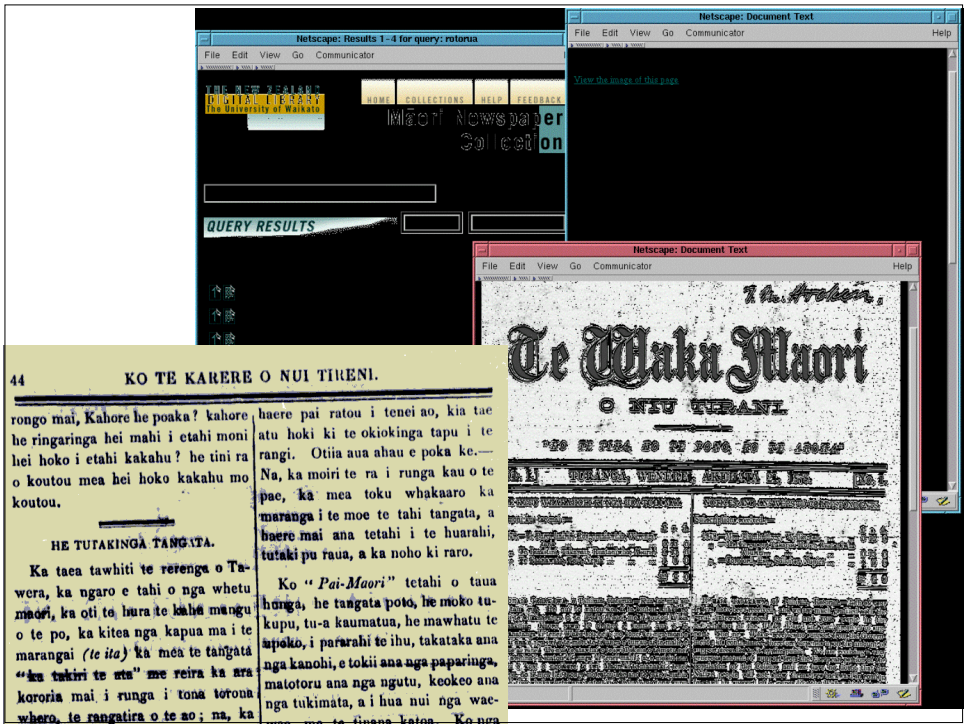
- What is digital library software?
- What does Greenstone do?
- Example Greenstone collections
- Collection building
- Documentation and help
- Getting into details …

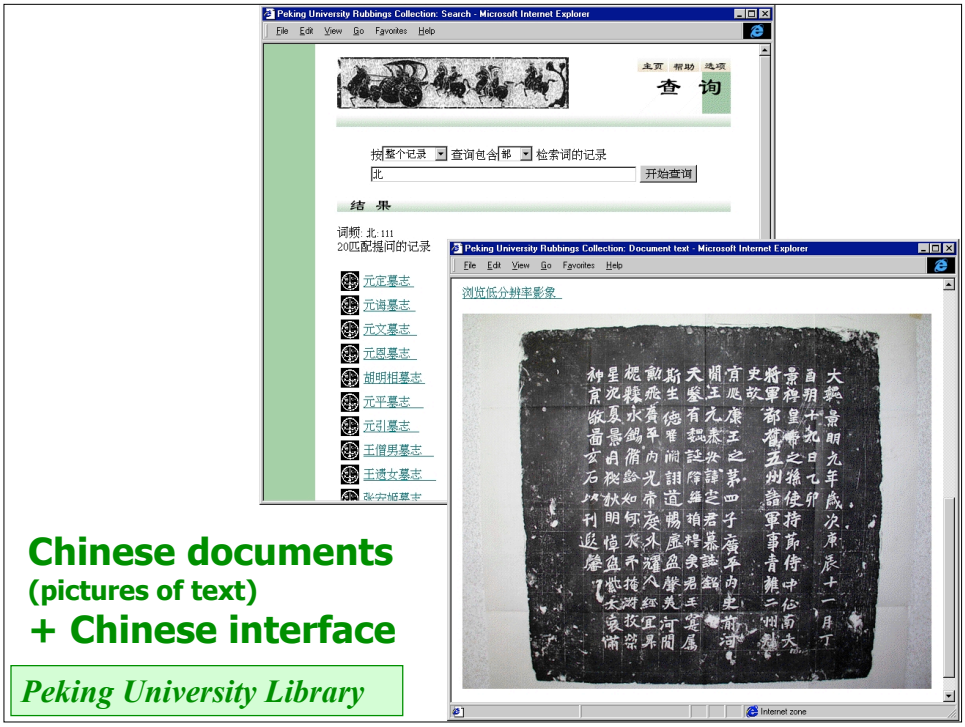
Example Greenstone collections

- Rapid growth in use
- International – China, Germany, India, UK, USA, Russia, etc.
- Increasing activity on Greenstone mailing list
- Promotion by UNESCO – "deployment of DL's for sharing public domain information"
- Wide variety of DL collections have been developed in several languages
  - historical, educational, cultural, and research


New York Botanical Garden
Rare 19th century works on American trees
Gorgeous full-color plates
Histoire des arbres forestiers, Volume III (1...
search
titles
HOME
Histoire des arbres forestiers, Volume III (1813)
Bignonia Catalpa.
The Catalpa.
[blank] [plate] [217] 218 219 220 221
DETACH
NO HIGHLIGHTING
LARGE IMAGE
TEXT
BIGNONIA CATALPA.
THE CATALPA.


University of Chicago Library
Chopin Early Editions
BALLADE
F. CHOPIN
Allegro de concert : pour le piano : op. 46 /
search
titles
uniform titles
genres
dedicatees
HELP PREFERENCES
prev
next
view score
bibliographic description
Page 22
Go
Title: Allegro de concert : pour le piano : op. 46 /
Composer: Chopin, FrÃ©dÃ©ric, 1810-1849.
Dedicatee: MÃ¼ller-Streicher, Friederike
Publication Info.: Leipzig :Chez Breitkopf & HÃ ¤ rtel :Paris :Chez M. Schlesinger,[1841]
Description: 23 p. of music ; 34 cm.
Notes: composÃ© et dÃ©diÃ© Ã Mademoiselle F. MÃ¼ller par FrÃ©d. Chopin.
Engraved. Price: 1 Thlr. 6 Ngr.
Also available on the Internet.
Subject: Piano music.
Plate No.: 6651
URI: http://pi.lib.uchicago.edu/1001/dig/chopin/017
Call No.: M25.C54 A41


Netscape: Results 1 - 4 for query: rotorua
Netscape: Document Text
QUERY RESULTS
Te Waka Maori
44 KO TE KARERE O NUI TIRENI.
HE TUTAKINGA TANGATA.

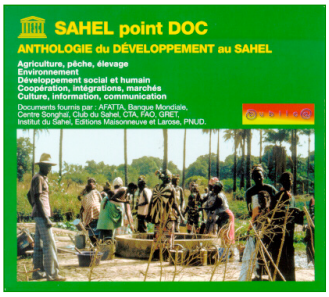

Peking University Rubbings Collection: Search - Microsoft Internet Explorer
查 询
Peking University Rubbings Collection: Document text - Microsoft Internet Explorer
浏览低分辨率影象
Chinese documents
(pictures of text)
+ Chinese interface
Peking University Library


Chinese
(Chinese & English interfaces)
Classic Chinese literature


Arabic
(Arabic & English interfaces)
Famous mosques

French

UNESCO, Paris

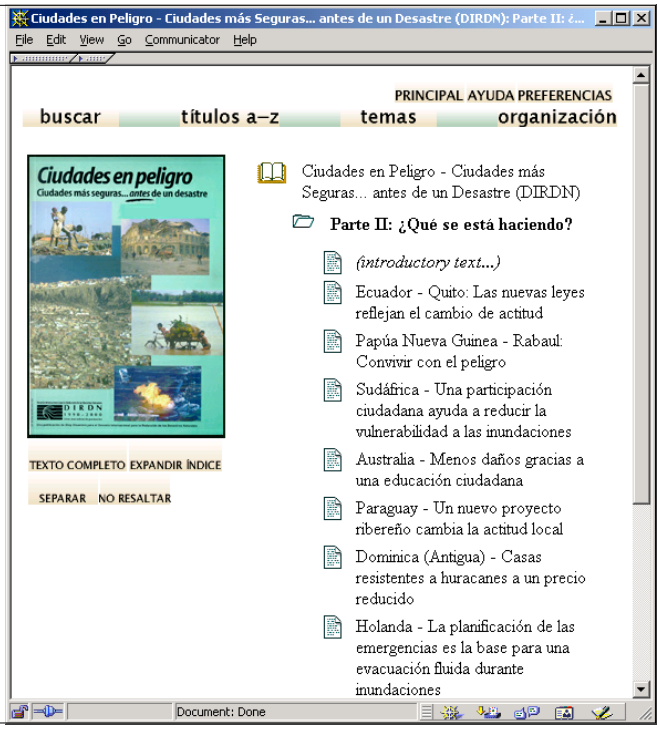




Spanish

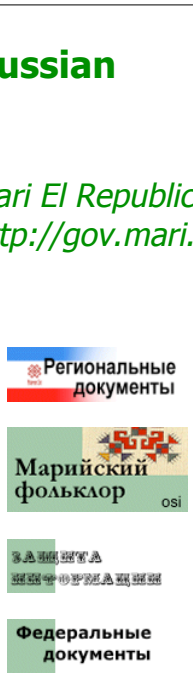
PAHO, WHO

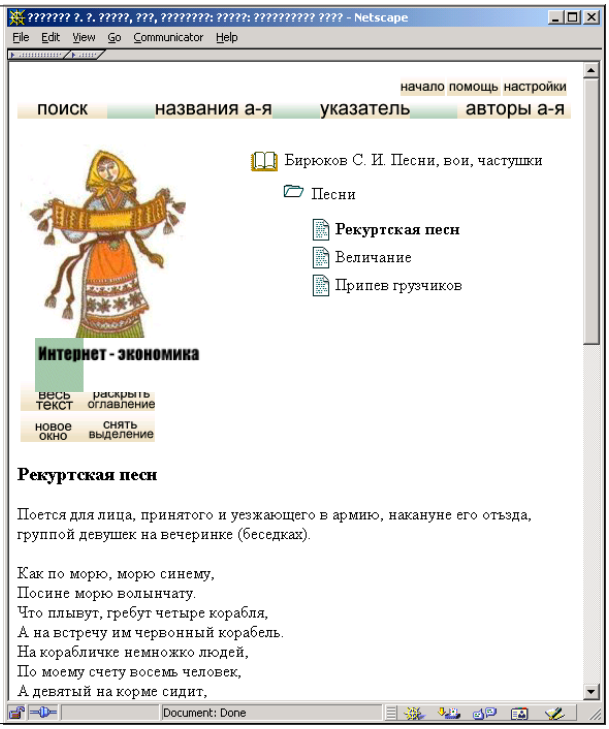




Russian

Mari El Republic
http://gov.mari.ru/gsdl





Example Greenstone collections

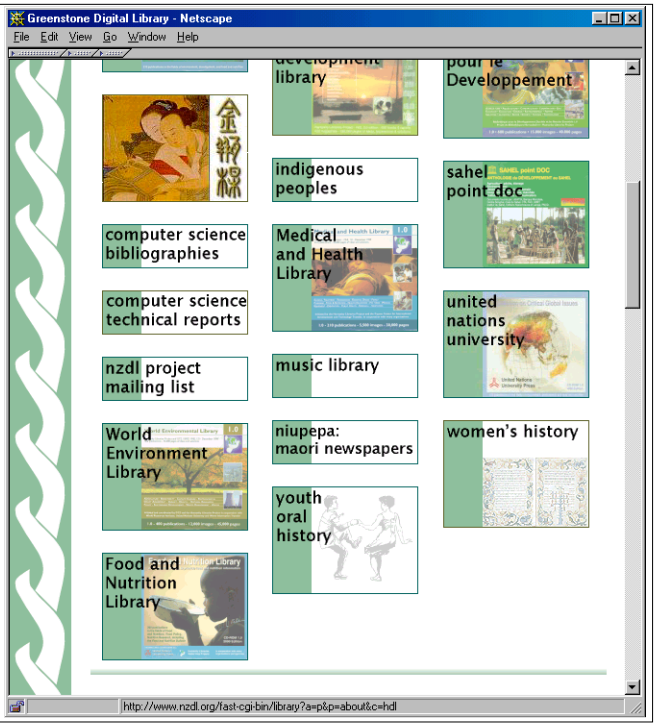
- For brief description and links to DLs built using Greenstone:
  - "*Examples of practical digital libraries: Collections built internationally using Greenstone*", By Ian Witten. D-Lib Magazine, March 2003. (www.dlib.org)
- www.greenstone.org
- New Zealand DL Project (www.nzdl.org)

Collections: on the Web

nzdl.org

(demo, not service)





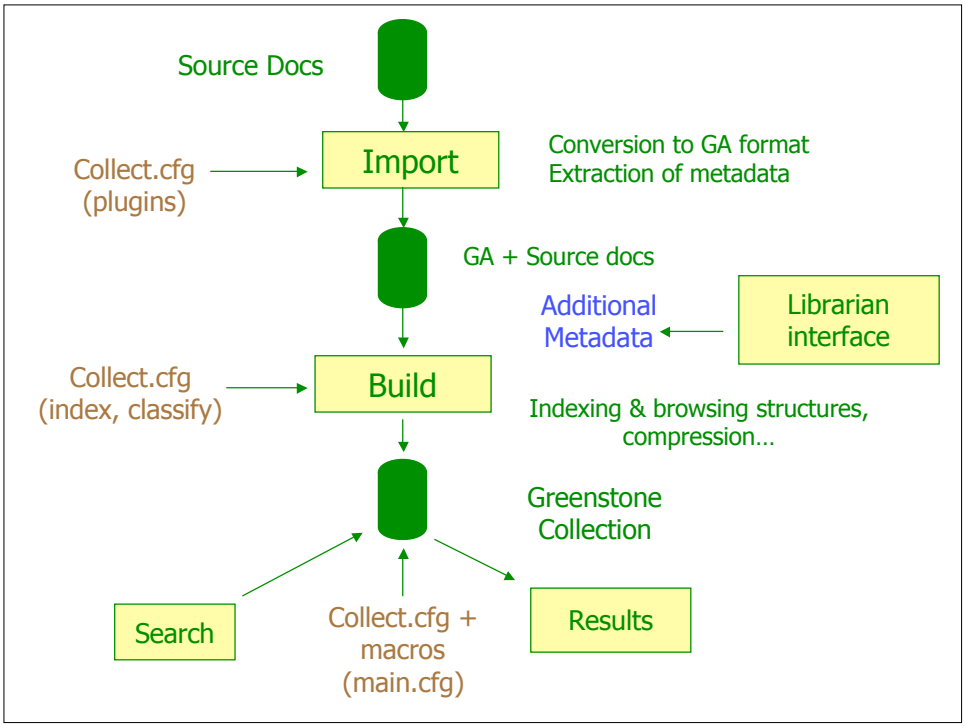


Agenda

- What is digital library software?
- What does Greenstone do?
- Example Greenstone collections
- Collection building
- Documentation and help
- Getting into details …

Collection Building

- Input: a set of source documents
- Greenstone "imports" these collections and converts them to GA format (XML/HTML)
- Greenstone "builds" indexes and browsing structures using the GA files
- Collection configuration file determines content conversion, extraction and building of indexes and browsing structures
  - indexes, classifiers, plugins
- Presentation of search/browse results and collection interface is determined by "format" strings and "macros"



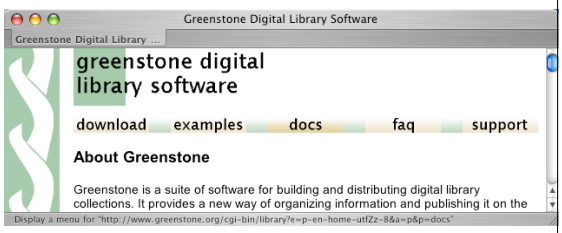
Documentation and Help

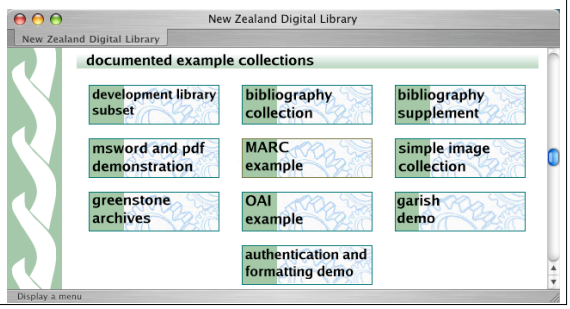
Manuals on the CD-ROM (gsdl/docs)

- *From Paper To Collection* (paper.pdf, 30pp)
  Scanners and scanning, OCR, 3 examples – from 1,000 to 100,000 pages, Creating an electronic collection
- *Installer's Guide* (install.pdf, 36pp)
  Versions of Greenstone, installation procedure, Greenstone collections, setting up the web server, configuring your site, personalizing your installation
- *User's Guide* (user.pdf, 90pp)
  Overview of Greenstone, using Greenstone collections, the collector, administration, software features, glossary of terms
- *Developer's Guide* (develop.pdf, 113pp)
  Understanding the collection building process, getting the most out of your collections, the Greenstone runtime systems, configuring your Greenstone site

Documentation and Help

- greenstone.org
  - Examples
  - support (+ join mailing list)
  - Configuration files for nzdl.org collections
- FAQ
  - General info section
- nzdl.org
  - Documented example collections





Getting into details...

Remaining sessions:

1. (this session)
2. Greenstone installation
3. Librarian interface: Building collections
4. Librarian interface: Adding and using metadata
5. Librarian interface: Advanced features
6. Customizing your collection, creating a CD-ROM
7. How building works; bibliography collections
8. Multilingual support; interface languages; Niupepa

Questions?
Comments?
Discussion?
Feedback form!