

THƯ VIỆN ĐIỆN TỬ GREENSTONE HƯỚNG DẪN SỬ DỤNG

Tác giả: Lan H.Witten và Stefan Boddie

*Khoa Công Nghệ Thông Tin trường
Đại học Waikato, New Zealand*

Greenstone là một bộ phần mềm hỗ trợ việc xây dựng và phân phối các bộ sưu tập Thư viện số. Nó cung cấp một phương thức mới trong việc tổ chức và xuất bản thông tin trên Internet hoặc trên CD-ROM. Greenstone là kết quả của dự án Thư viện số tại trường đại học Waikato, New Zealand (New Zealand Digital Library Project), đã được triển khai và phân phối với sự hợp tác của hai tổ chức UNESCO và Human Info NGO. Đây là một phần mềm có mã nguồn mở tại địa chỉ <http://greenstone.com> , trong mục GNU General Public License.

Chúng tôi muốn đảm bảo rằng phần mềm này sẽ đáp ứng tốt nhu cầu của bạn. Vì vậy, nếu có bất kì vấn đề nào liên quan đến phần mềm này, đề nghị các bạn liên lạc với greenstone@cs.waikato.ac.nz

Hướng dẫn sử dụng:

Tài liệu này hướng dẫn chi tiết cách sử dụng Greenstone để truy cập và xây dựng các bộ sưu tập Thư viện số.

Phần 1: Trình bày tổng quát về các tính năng của phần mềm

Phần 2: Hướng dẫn sử dụng các bộ sưu tập Greenstone. Giao diện này rất dễ sử dụng – Cách học Greenstone nhanh nhất là thực hành – Trong phần này bao gồm luôn những thông tin hướng dẫn trực tuyến cho một bộ sưu tập bình thường.

Phần 3: Hướng dẫn bạn tự xây dựng các bộ sưu tập Thư viện bằng công cụ Greenstone Collector. Công cụ này bao gồm các trang Web hướng dẫn bạn từng bước xây dựng một bộ sưu tập.

Phần 4: Giới thiệu tính năng Quản lý, cho phép người quản lý hệ thống có thể theo dõi được diễn tiến công việc và kiểm soát được những người đang làm công việc thiết kế các bộ sưu tập.

Phân phụ lục: Liệt kê các tính năng của phần mềm Greenstone và cung cấp một bảng chú giải thuật ngữ được sử dụng trong toàn bộ tài liệu này.

Các tập tài liệu trong Bộ phần mềm Greenstone

Bộ phần mềm này bao gồm 4 tập tài liệu:

- Hướng dẫn cài đặt
- Hướng dẫn sử dụng
- Hướng dẫn phát triển
- Từ tài liệu bằng giấy đưa lên bộ sưu tập.

Những thành viên tham gia dự án phần mềm Greenstone

Greenstone là sự hợp tác của nhiều người trong đó Rodger McNab và Stefan Boddie là hai nhân vật chủ yếu trong việc thiết kế và phát triển phần mềm này. Ngoài ra còn có sự đóng góp của các tác giả sau: David Bainbridge, George Buchanan, Hong chen, Elke Duncker, Carl Gutwin, Geoff Holmes, John McPherson, Craig Nevill-Manning, Gordon Paynter, Bernhard Pfahringer, Todd Reed, Bill Rogers và Stuart Yeates. Những thành viên khác trong dự án Thư viện số tham gia phần Thiết kế hệ thống là: Mark Apperley, Sally Jo Cunningham, Steve Jones, Te Taka Keegan, Michel Loots, Malika Mahoui và Lloyd Smith.

Chúng tôi cũng chân thành cảm ơn tất cả những ai đã góp sức vào việc xây dựng MG, GDBM, WGET, WV, PDF2HTML, PERL trong phiên bản này

1.	TỔNG QUAN VỀ GREENSTONE	4
1.1.	Các bộ tài liệu	4
1.2.	Tìm kiếm thông tin.....	5
1.3.	Định dạng dữ liệu.....	5
1.4.	Các tài liệu Đa phương tiện và Đa ngôn ngữ.....	5
1.5.	Chức năng phân phối của phần mềm.....	6
2.	SỬ DỤNG BỘ PHẦN MỀM GREENSTONE	7
2.1.	Cài đặt phần mềm bằng CD-ROM.....	7
2.2.	Tìm kiếm thông tin.....	8
2.3.	Thay đổi thuộc tính	14
3.	TẬP HỢP CÁC TÀI LIỆU	17
3.1.	Logging in.....	17
3.2.	Cấu trúc Dialog	18
3.3.	Thu thập thông tin.....	19
3.4.	Dữ liệu nguồn	20
3.5.	Cấu hình Bộ sưu tập.....	24
3.6.	Xây dựng Bộ sưu tập	25
3.7.	Trình bày bộ sưu tập	26
3.8.	Làm việc với bộ sưu tập đã tồn tại.....	26
3.9.	Định dạng tài liệu.....	27
4.	Chức năng quản trị.....	30
4.1.	File cấu hình.....	33
4.2.	Logs.....	33
4.3.	Quản lý người dùng	34
4.4.	Thông số kỹ thuật.....	34

1. TỔNG QUAN VỀ GREENSTONE

Greenstone là một hệ thống hoàn chỉnh dùng để xây dựng và trình bày các bộ sưu tập gồm có hàng ngàn, hàng triệu tài liệu bằng chữ, hình ảnh, băng tiếng hay băng hình.

1.1. Các bộ tài liệu

Một Thư viện số thông thường được xây dựng bằng phần mềm Greenstone sẽ chứa được nhiều bộ sưu tập, được sắp xếp riêng lẻ thông qua sự giống nhau nổi bật, được duy trì dễ dàng. Ngoài ra bộ tài liệu còn có thể được bổ sung và tự động tái tạo lại.

Có nhiều cách để tìm kiếm thông tin trong bộ sưu tập Greenstone. Ví dụ, bạn có thể *tim bằng từ khoá* (là từ xuất hiện trong một đoạn văn bản hay một phần trong tài liệu cần tìm). Bạn có thể *trình duyệt tài liệu theo tiêu đề* bằng cách nhấp chuột vào quyển sách. Bạn cũng có thể *trình duyệt tài liệu theo chủ đề*. Các chủ đề được sắp xếp theo dạng kệ sách, bạn chỉ cần nhấp chuột vào kệ sách để tìm các quyển sách ở trong đó. Nhiều khi các tài liệu còn có cả Bảng mục lục: Bạn có thể nhấp chuột vào một chương hoặc một thư mục nhỏ để mở ra xem, để mở rộng cả bảng mục lục, hay mở toàn bộ tài liệu trong cửa sổ trình duyệt của bạn (Giúp ích cho việc in ấn). Trang web Thư viện số New Zealand (nzdl.org) cung cấp rất nhiều bộ sưu tập mẫu.

Trên mỗi trang đầu của bộ sưu tập đều trình bày mục đích sử dụng và hình trang bìa của từng quyển sách, và những chỉ dẫn cách sắp xếp bộ tài liệu. Hầu hết các bộ sưu tập đều có thể được truy cập bằng cả hai cách: *Tim kiếm* và *trình duyệt*. Khi tìm kiếm, phần mềm Greenstone sẽ tìm toàn bộ nội dung của tất cả các văn bản trong bộ sưu tập. Trong hầu hết các bộ sưu tập, người sử dụng có thể chọn sử dụng các chỉ mục được xây dựng từ những phần khác nhau của tài liệu. Một số bộ sưu tập có chỉ mục tất cả tài liệu, chỉ mục từng đoạn, và chỉ mục các tựa sách. Có thể tìm các chỉ mục này theo từ hay cụm từ khóa. Bằng cách này, bạn có thể tìm được tất cả các văn bản có chứa các từ khóa nào đó (các từ khoá này có thể được phân bố rải rác trong văn bản), hoặc tất cả những đoạn văn có chứa các từ này (các từ này phải cùng xuất hiện trên cùng một đoạn văn), hoặc tất cả các sưu tập có các tiêu đề chứa các từ này (các từ này phải cùng xuất hiện trên tựa của văn bản). Còn có cả các chỉ mục khác, ví dụ chỉ mục theo từng phần hay theo những đề mục nhỏ. Trong quá trình trình duyệt, người sử dụng có thể nghiên cứu danh mục các tác giả, danh mục các tiêu đề, danh mục các ngày, cấu trúc phân loại theo từng lớp, v.v... Các bộ sưu tập khác nhau có thể có các cách trình duyệt khác nhau.

1.2. Tìm kiếm thông tin

Phần mềm Greenstone xây dựng những chỉ mục toàn phần theo nội dung văn bản - nghĩa là những chỉ mục giúp bạn tìm kiếm theo bất cứ từ nào trong toàn bộ nội dung tài liệu. Chỉ mục có thể được dùng để tìm kiếm theo các từ khoá, hay các cụm từ, và kết quả sẽ được sắp xếp theo thứ tự dựa trên sự phù hợp đối với câu truy vấn.

Trong hầu hết các bộ tài liệu, các dữ liệu mô tả như Tác giả, Tiêu đề, ngày tháng, từ khoá, v.v.. đều đi kèm với từng tài liệu. Dữ liệu này được gọi là dữ liệu khoá (metadata). Nhiều bộ sưu tập chứa cả chỉ mục nội dung của một số loại dữ liệu khoá. Ví dụ, một số bộ sưu tập, người dùng có thể tìm kiếm dựa trên chỉ mục về Tên tài liệu.

Người dùng có thể trình duyệt một cách trực tiếp dựa trên danh mục và cấu trúc phân lớp được thiết lập từ dữ liệu khoá có liên quan đến các tài liệu trong bộ sưu tập. Dữ liệu khoá là dữ liệu thô của việc tìm kiếm. Nó phải được trực tiếp cung cấp hoặc được tự động rút ra từ chính các tài liệu. Các bộ sưu tập khác nhau có thể có những tính năng tìm kiếm và trình duyệt khác nhau. Chỉ mục tìm kiếm và trình duyệt được thiết lập trong quá trình xây dựng bộ sưu tập dựa trên các thông tin trong tập tin cấu hình.

Greenstone tự động tạo ra các chỉ mục từ các tài liệu và tập tin hỗ trợ: Không một quá trình nào phải làm bằng tay. Nếu các tài liệu mới có cùng một định dạng, chúng sẽ tự động được xếp vào chung một bộ sưu tập. Trên thực tế, trên nhiều bộ sưu tập, việc này được tiến hành bởi các quá trình được kích hoạt thường xuyên để theo dõi tài liệu mới, và cập nhập lại các chỉ mục - Tất cả đều được thực hiện một cách tự động.

1.3. Định dạng dữ liệu

Tài liệu nguồn có nhiều dạng khác nhau, và được chuyển sang định dạng chuẩn XML để sử dụng plugins. Các plugins được phân phát với Greenstone để hỗ trợ các văn bản thô, HTML, WORD, PDF, Usenet và E-mail. Còn plugins khác có thể được viết cho những kiểu dữ liệu khác (để sử dụng chúng, bạn cần phải đọc phần *Hướng dẫn phát triển Phần mềm Greenstone*). Để xây dựng cấu trúc trình duyệt từ dữ liệu khoá, ta sẽ tương tự sử dụng một cách phân loại theo lớp. Các phân loại này sẽ tạo ra các chỉ mục trình duyệt khác nhau: liệt kê dạng thanh cuộn, dạng mẫu tự Alphabet, dạng ngày tháng, hay một dạng cấu trúc lớp tùy ý. Một lần nữa, các lập trình viên trên Greenstone có thể tạo ra cấu trúc trình duyệt mới.

1.4. Các tài liệu Đa phương tiện và Đa ngôn ngữ

Bộ sưu tập có thể chứa chữ, hình ảnh tĩnh, hình ảnh động, âm thanh. Những tài liệu không thuộc dạng chữ được liên kết với những tài liệu dạng chữ hay những tài liệu mô tả dạng chữ (ví dụ như chú thích hình ảnh) để hỗ trợ việc tìm kiếm và trình duyệt nội dung

Unicode, phong chữ chuẩn trên Thế giới dùng để trình bày nội dung tài liệu, cũng được sử dụng trong Greenstone. Điều này cho phép bất cứ ngôn ngữ nào cũng đều được xử lý và trình bày theo một kiểu thống nhất. Các bộ sưu tập được đã được xây dựng có chứa các kiểu chữ Ả rập, Trung Quốc, Anh, Pháp, Māori và Tây Ban Nha. Chương trình tự động nhận ra ngôn ngữ mẫu trong bộ sưu tập và giao diện được trình bày theo những ngôn ngữ sẵn có trên.

1.5. Chức năng phân phối của phần mềm

Các bộ sưu tập được truy cập thông qua Internet, dưới dạng các ấn phẩm, hoặc bằng đĩa CD-ROM tự cài đặt. Chương trình nén được sử dụng để nén nội dung tài liệu và chỉ mục. Một giao thức Corba sẽ hỗ trợ cho các bộ sưu tập được phân phối và giao diện truy vấn..

Thư viện số New Zealand(nzdl.org) cung cấp các bộ sưu tập mẫu bao gồm: Tập tài liệu về lịch sử, thông tin về con người và sự phát triển, báo cáo kỹ thuật và tiểu sử, các tác phẩm văn học và tạp chí.

Là phần mềm mã nguồn mở, Greenstone dễ mở rộng và hưởng lợi từ các modules truy cập nội dung, quản lý cơ sở dữ liệu, và lấy nội dung từ các loại định dạng khác nhau của sự cho phép của GNU. Chỉ với sự hợp tác quốc tế thì phần mềm này mới có thể trở thành một phần mềm hoàn thiện, đáp ứng được những yêu cầu ngày càng đa dạng và phong phú của người sử dụng.

2. SỬ DỤNG BỘ PHẦN MỀM GREENSTONE

Bộ phần mềm Greenstone được thiết kế với các chức năng dễ sử dụng. Bộ sưu tập trên Web hay CD-ROM đều có cùng giao diện. Việc cài đặt phần mềm Greenstone từ CD-ROM hoàn toàn dễ dàng tích hợp với cả hệ điều hành Windows hay Linux. Việc cài đặt được tiến hành qua một chương trình cài đặt chuẩn và các tập tin binary được dịch trước. Một bộ sưu tập có thể được cài đặt để sử dụng trên chính máy tính được cài hoặc khi máy tính này được kết nối vào mạng, phần mềm này sẽ tự động cho phép tất cả các máy trên mạng truy cập cùng một bộ sưu tập.

Phần tiếp theo mô tả cách cài đặt phần mềm bằng CD-ROM. Sau đó chúng ta sẽ xem xét đến tính năng tìm kiếm và trình duyệt trên một bộ sưu tập Greenstone mẫu, đó là chương trình “Demo” được cung cấp cùng với phần mềm. Những bộ sưu tập khác cũng có các tính năng tương tự. Nếu sử dụng được một cái, bạn cũng sẽ sử dụng được những cái khác. Phần tiếp theo sẽ chỉ dẫn dùng trang Thuộc tính để thay đổi giao diện theo ý muốn.

2.1. Cài đặt phần mềm bằng CD-ROM

Bản thân phần mềm Thư viện số Greenstone là từ CD-ROM, và bạn hay người quản lý hệ thống có lẽ đã cài đặt chương trình vào hệ thống của mình, dựa theo phần *Hướng dẫn cài đặt*. Nếu đã cài đặt rồi, bạn nên bỏ qua phần này.

Một số bộ sưu tập Greenstone đến từ một CD-ROM có sẵn các chương trình để chạy bộ sưu tập. Khi sử dụng, bạn chỉ cần để nó vào ổ đĩa CD-ROM trên bất kỳ máy tính nào dùng Windows. Thông thường, (nếu chức năng tự động “Autorun” đã được kích hoạt trên máy của bạn), một cửa sổ sẽ xuất hiện và mời bạn cài đặt phần mềm này. Nếu không có chức năng tự động (Nếu sử dụng hệ điều hành Windows bạn nhấp chuột vào biểu tượng *My computer* trên Desktop) chương trình cài đặt sẽ xuất hiện, sau đó double-click vào đó, tìm file *setup.exe* và nhấp chuột vào đó thì quá trình cài đặt được tiến hành và bạn sẽ được hướng dẫn từng bước, trong quá trình cài đặt chương trình sẽ luôn xuất hiện các câu hỏi và hầu hết các câu trả lời là ”yes”.

Khi quá trình cài đặt hoàn thành, bạn sẽ tìm thấy Thư viện trong phần Programs với tên của chương trình đã cài đặt như (“Development Library” hay “United Nations University”).

Khi chương trình đã được cài đặt, nếu chức năng “Autorun” đã được kích hoạt, Thư viện sẽ tự động khởi động mỗi khi để CD-ROM vào ổ.

2.2. Tìm kiếm thông tin

Để sử dụng phần mềm này, cách dễ nhất là chạy thử các chương trình. Đừng lo ngại - bạn chẳng làm hỏng chương trình đâu. Hãy tự do nhấp chuột vào Thư viện: hầu hết các hình ảnh của từng chương trình sẽ hiện ra. Nếu giữ yên con chuột phía trên một hình nào đó, thông thường sẽ có một thông báo nhỏ báo với bạn chương trình này hoạt động như thế nào.

Hãy thử nghiệm bằng cách chọn những từ quen thuộc như “the” và “and” để tìm kiếm

Hệ thống Thư viện số Greenstone được tập hợp từ nhiều tài liệu riêng lẻ, ví dụ như các tài liệu kỹ thuật về tin học, về các tác phẩm học, Các tài liệu của Tổ chức Lương thực và Nông nghiệp của Liên Hiệp Quốc (FAOs), các tạp chí chuyên đề...v.v.. Thư viện số có một trang chủ cho phép bạn truy cập. Thêm vào đó, mỗi bộ sưu tập có một trang thông tin riêng nhằm cung cấp cho bạn các thông tin về cách sắp xếp và quản lý các tài liệu con. Để quay trở lại trang Thông tin này, bạn chỉ cần nhấp chuột vào biểu tượng “collection” ở góc trên bên trái cạnh các trang tìm kiếm và trình duyệt.

Hình 1 sẽ trình bày phần “demo” là phần rất nhỏ trong tập tài liệu “Development Library collection”, chúng ta sẽ dùng nó như là một ví dụ để mô tả những cách tìm kiếm thông tin khác nhau (Nếu không tìm thấy chương trình *Demo* bạn có thể sử dụng luôn chương trình Development Library, chúng cũng tương tự nhau). Trước tiên phải kích hoạt chương trình bằng cách nhấp chuột vào các biểu tượng. Có nhiều biểu tượng xuất hiện trên dòng đầu ở hầu hết mỗi trang; Bảng 1 sẽ trình bày các ý nghĩa của biểu tượng

Hình 1:
Sử dụng chương
trình Demo



Các thanh phía dưới “search.. subjects..titles a-z..organization.. how to..” là các công cụ trình duyệt và tìm kiếm, nút ngoài cùng bên phải là nút tìm kiếm, 4 nút còn lại là những điều kiện tìm kiếm khác nhau ứng với các tài liệu khác nhau.

Cách tìm kiếm thông tin

Bảng 2 trình bày 5 cách tìm kiếm thông tin trong phần Demo.

Bạn có thể *tìm kiếm bằng từ khóa* mà đã xuất hiện trong phần nội dung tài liệu từ trang “tìm kiếm”. Phần này cũng giống như trang “about” đã được trình bày ở hình 1 (ngoại trừ việc nó không chứa phần *About this collection*). Trang tìm kiếm có thể có thể được truy xuất từ những trang khác bằng cách nhấn nút *search*). Bạn có thể truy cập tên xuất bản theo từng chủ đề (*Publications by subject*) bằng cách nhấn nút *subjects*. Sau đó một bảng danh mục các chủ đề hiện ra, và được trình bày theo dạng kệ sách, có thể mở rộng ra bằng cách nhấp chuột vào các mục con tiếp theo. Bạn cũng có thể truy cập bằng Danh sách liệt kê *How to* bằng cách nhấn nút *How to*. Phần này sẽ là một bảng danh mục gợi ý bằng câu hỏi “*How to*”. Xem Hình 1 (PV).

Bảng 1 Trình bày ý nghĩa của các biểu tượng (icons) trên dòng đầu của mỗi trang

greenstone
demo

Phần này đưa bạn đến trang “about”.

HOME	Phần này dẫn bạn quay về trang chủ, nơi mà bạn có thể chọn lại tài liệu cần tìm.
HELP	Phần này giúp bạn tìm hiểu nội dung đang đọc.
PREFERENCES	Phần này cho phép bạn chọn lựa giao diện và các hộp hội thoại tìm kiếm theo ý muốn và sử dụng phần đã chọn trong suốt thời gian tìm kiếm.

Bảng 2 Trình bày ý nghĩa của các biểu tượng (icons) trên thanh tìm kiếm.

search	Tìm kiếm bằng từ khóa
subjects	Truy cập sách xuất bản theo Chủ đề
titles a-z	Truy cập sách xuất bản theo Tên sách
organisations	Truy cập sách xuất bản theo Tên các Tổ chức thế giới
how to	Truy cập sách xuất bản theo bảng trợ giúp “how to”

Cách đọc hiểu tài liệu

Trong phần Demo, bạn có thể biết được đây là quyển sách cần tìm nhờ vào ảnh bìa của quyển sách (Hình 2). Bên cạnh ảnh bìa là bảng mục lục: Khi bạn truy cập đến phần nào thì phần đó sẽ được in đậm, trong trường hợp này là phần *Introduction and summary* – Phần 1. Bảng này sẽ được mở rộng, bạn có thể nhấp chuột vào các thư mục để đóng mở chúng. Nhấp chuột vào biểu tượng mở sách để đóng sách và ngược lại.

Phần bên dưới sẽ là nội dung của phần hiện hành (“The international demand for tropical butterflies... Trong phần ví dụ, nội dung được trình bày ngay dưới mỗi bức hình minh họa). Khi đọc đến những dòng cuối bạn có thể sang phần tiếp theo để đọc hoặc trở về đầu trang.

Phía dưới bức hình có 4 nút.

Nút detach có tác dụng mở ra cửa sổ mới. (rất tiện khi mở nhiều quyển sách cùng một lúc để đọc hoặc so sánh). Nếu bạn đang đọc quyển sách được tìm bằng nút “Search” thì mục search sẽ được tô sáng, còn nút no highlighting sẽ ẩn.

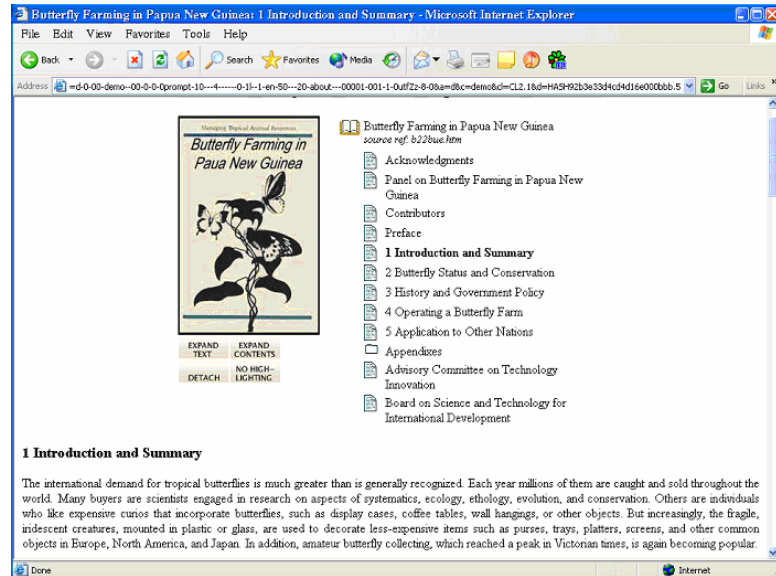
Nút expand text sẽ mở rộng toàn bộ nội dung thư mục hiện hành hoặc toàn bộ nội dung cuốn sách hiện hành.

Nút expand contents sẽ mở rộng toàn bộ nội dung của mục lục để bạn thấy tên tất cả các chương và mục nhỏ.

Trong một số tập tài liệu, phần nội dung không cần phải trình bày theo cấu trúc phân cấp. Với trường hợp này sẽ không có bảng mục lục khi bạn xem phần nội dung tài liệu - chỉ xuất hiện duy nhất phần nội dung. Một số trường

hợp thì nội dung được chia thành nhiều trang, và bạn có thể đọc lần lượt từng trang hoặc nhảy sang từ trang này sang trang khác.










Hình 2
Một quyển sách trong phần Demo.



Ý nghĩa các biểu tượng

Trong khi tìm kiếm tài liệu bạn sẽ gặp những đề mục sau trong bảng 3.

Bảng 3 Các biểu tượng (icons) bạn sẽ gặp trong lúc tìm kiếm.

-  Nhấp chuột vào biểu tượng này để đọc sách
-  Nhấp chuột vào để xem từng Đề tài
-  Trình bày nội dung tài liệu
-  Mở Thư mục và trình bày nội dung
-  Nhấp chuột vào đây để đóng sách
-  Nhấp chuột vào đây để đóng Thư mục
-  Sang trang kế tiếp
-  Trở về trang trước
-  Mở trang này trong cửa sổ mới.

EXPAND
CONTENTS

Mở rộng bảng mục lục

DETACH

Trình bày toàn bộ nội dung văn bản

HIGH-
LIGHTING

Tô sáng đề mục tìm kiếm

Cách tìm bằng từ khoá

Từ trang tìm kiếm bạn thực hiện từng bước sau:

- Xác định rõ bạn cần tìm cái gì: trong bộ sưu tập demo, bạn có thể tìm tên sách hoặc toàn bộ nội dung quyển sách.
- Chọn cách tìm kiếm bằng toàn bộ từ khóa hay chỉ một số từ khóa
- Gõ những từ khóa vào trong hộp hội thoại truy vấn.
- Nhấp chuột nút begin search để tìm.

Khi bạn thực hiện câu truy vấn, tên của hai mươi tài liệu phù hợp sẽ được hiện lên. Ở cuối mỗi trang sẽ có nút cho phép bạn chuyển sang trang kế chứa 20 tài liệu kế tiếp, cứ thế bạn có thể chuyển sang trang kế hay trở về trang đầu tiên, và cứ thế tiếp tục di chuyển tới, lui các trang kết quả. Tuy nhiên kết quả được trình bày từ 100 mẫu tin trở lại mới được xem là chính xác. Bạn có thể thay đổi chỉ số này trong mục *preferences* trên phần đầu của trang.

Nhấp chuột vào tên của bất cứ tài liệu nào hay icon nhỏ bên cạnh đó để mở tài liệu ra xem. Các biểu tượng có thể trình bày một cuốn sách, một thư mục hay một trang: Biểu tượng là hình cuốn sách nếu như bạn chọn phần tìm kiếm là sách và ngược lại là sẽ là thư mục hay từng trang nếu như phần tìm kiếm có các phụ mục.

Đề mục tìm kiếm

Bất cứ khi nào bạn gõ vào hộp hội thoại tìm kiếm một cụm từ, thì cụm từ đó được gọi là “Từ khóa”. Các từ này chỉ chứa các kí tự alphabet hoặc các chữ số. Các từ phải được cách ra bằng khoảng trắng. Dấu chấm câu, chấm phẩy, gạch ngang..v.v.. được coi như khoảng trắng và do đó bị xem như không có. Bạn không thể tìm kiếm với các từ khóa chứa dấu chấm câu. Ví dụ:

Ta có câu truy vấn sau (query)

Argo-forestry in the Pacific Islands: Systems for Sustainability 1993

Sẽ tương đương với

Argo forestry in the Pacific Islands Systems for Sustainability 1993

Loại câu truy vấn

Có 2 loại truy vấn khác nhau.

- Loại truy vấn thứ nhất tìm các tài liệu (chương hoặc tiêu đề) chứa tất cả các từ khóa. Những tài liệu nào thỏa điều kiện tìm kiếm sẽ được trình bày.
- Loại thứ hai tìm các tài liệu chứa một số từ trong một loạt từ khóa. Các tài liệu được sắp xếp dựa theo tính phù hợp của chúng đối với điều kiện tìm kiếm. Tính phù hợp được xác định dựa trên các tiêu chuẩn sau:

- Tài liệu nào chứa càng nhiều các từ tìm kiếm trong hộp hội thoại thì càng phù hợp
- Các từ khóa không thông dụng quan trọng hơn các từ thông dụng
- Tài liệu có nội dung ngắn quan trọng hơn

Người sử dụng có thể chọn những cách tìm kiếm mà họ thích - Chọn cả một câu hoặc ngay cả một đoạn. Nếu bạn chỉ dùng một từ khóa thì việc chọn cách tìm kiếm là Tìm tất cả hay chỉ tìm một vài từ không quan trọng, ngoại trừ việc các kết quả được sắp thứ tự theo số lần xuất hiện của từ khóa trong mỗi tài liệu khi chọn cách thứ hai

Phạm vi câu truy vấn

Trong hầu hết các Bộ sưu tập, bạn có thể lựa chọn các danh mục tìm kiếm khác nhau. Tìm kiếm theo danh mục Tác giả, Tên sách hoặc theo từng Chương, từng Đoạn. Nói chung, dù bạn dùng cách nào, bạn cũng sẽ nhận được một văn bản phù hợp điều kiện tìm kiếm

Nếu tài liệu là sách thì sẽ được mở ra ở nơi thích hợp.

Chức năng tìm kiếm cấp cao

Chương trình này còn có thêm chức năng tìm kiếm cấp cao. Phần này được kích hoạt từ trang Preferences bằng cách nhấn nút Thuộc tính ở góc trên bên phải màn hình, xem ở phần 2.2 hình 1 bên trên

Khả năng phân biệt chữ hoa, thường và bỏ qua chữ cuối

Khi đã xác định danh mục tìm kiếm, bạn có thể chọn chức năng phân biệt giữa chữ hia và chữ thường cũng như chức năng loại bỏ phần đuôi của từ.

Bên dưới mục lựa chọn thuộc tính tìm kiếm trên trang Preferences, bạn sẽ thấy hai nút có tên sau ignore case differences và upper/lower case must match dùng để điều khiển chức năng có phân biệt chữ hoa hay chữ thường. Hai nút có tên ignore word endings và whole word must match dùng để điều khiển chức năng tìm kiếm chọn hay bỏ qua từ cuối.

Ví dụ: nếu bật nút *ignore case differences* và nút *ignore word endings* thì câu truy vấn sau

African buiding.

Sẽ tương đương với

africa builds.

vì chữ in hoa “A” sẽ được đổi thành chữ “a” thường trong từ “African” và chữ “n” và “ing” sẽ được tách khỏi hai từ “African” và “buiding”, tương tự “s” cũng được tách khỏi chữ “builds”.

Nói chung sự phân biệt giữa chữ viết hoa/viết thường hoặc các chữ cuối nên được bỏ qua nếu như câu truy vấn phải là tên riêng hoặc từ viết tắt của một nhóm từ (NATO, UNESCO..v..v).

Cụm từ tìm kiếm

Nếu câu truy vấn bao gồm cả cụm từ nằm trong ngoặc kép, chỉ những tài liệu nào chứa chính xác cả cụm từ này mới được chấp nhận.

Nếu muốn sử dụng phương pháp này, bạn cần phải hiểu được cách làm việc của nó. Quá trình tìm kiếm được tiến hành như sau: Quét để lấy ra từ giống như trong đề mục tìm kiếm (post-retrieval scan). Trước tiên câu truy vấn sẽ được cấp phát ở dạng chữ thường, sau đó chương trình sẽ quét để loại ra những từ không có trong tài liệu đang tìm và cuối cùng tài liệu nào thỏa điều kiện, tức là có chứa các từ trong đề mục tìm kiếm sẽ được trả ra.

Trong suốt quá trình quét và tìm kiếm (post-retrieval scan), các cụm từ tìm kiếm sẽ được kiểm tra bao gồm luôn cả dấu chấm câu, chấm hỏi...v.v..

Ví dụ: Câu truy vấn sau
“what's a post-retrieval scan?”

Trước tiên việc tìm kiếm sẽ tìm tất cả những tài liệu có liên quan đến những từ như

what s a post retrieval scan

Sau đó sẽ tìm đến những tài liệu có liên quan đến cả cụm từ
post-retrieval scan?

Nếu bật nút *ignore case differences* ở trang *Thuộc tính*, quá trình tìm kiếm sẽ không phân biệt chữ hoa, thường, bỏ qua từ cuối)

Cách thực hiện câu truy vấn cấp cao

Trong phần tìm kiếm cấp cao được mô tả chi tiết ở trên, loại câu truy vấn kiểu Chân trị, bao gồm những từ được liên kết bằng các toán tử logic & (and), | (or), ! (not). Trường hợp giữa các từ tìm kiếm không có toán tử liên kết thì sẽ được hiểu là phép liên kết & (and), do đó nếu câu truy vấn không chứa bất cứ toán tử nào thì tài liệu được trả về sẽ chứa tất cả các từ khóa.

Bạn phải dùng các toán tử “ & , | , ! “ thay cho phép “*and, or, not*” và dùng dấu ngoặc đơn để nhóm các từ tìm kiếm lại.

Sử dụng chức năng tìm kiếm *History*

Khi chuyển sang chức năng tìm kiếm *History*, trên trang Preferences sẽ trình bày một số kết quả tìm kiếm trước đó cùng với bảng tóm tắt số lượng kết quả thu được. Nhấp chuột vào nút bên cạnh một trong những kết quả tìm kiếm trước đó để copy phần nội dung vào hộp tìm kiếm. Điều này tạo sự thuận lợi trong việc chỉnh sửa câu truy vấn trước đó.

2.3. Thay đổi thuộc tính

Khi nhấp chuột vào nút *Thuộc tính* ở trên phần đầu trang bạn có thể thay đổi một số thuộc tính của giao diện cho phù hợp với yêu cầu của mình. Các

thuộc tính tùy thuộc vào bộ tài liệu, như một ví dụ ở Hình 3. Sau khi thay đổi thuộc tính tìm kiếm, bạn nên nhấn nút *set Preferences*. Bạn nên nhấp chuột vào một trong những nút trên thanh access ở đầu trang sau khi cài đặt thuộc tính, thay vì dùng nút “back”, vì khi nhấn nút này trạng thái sẽ quay trở lại như khi chưa cài đặt.

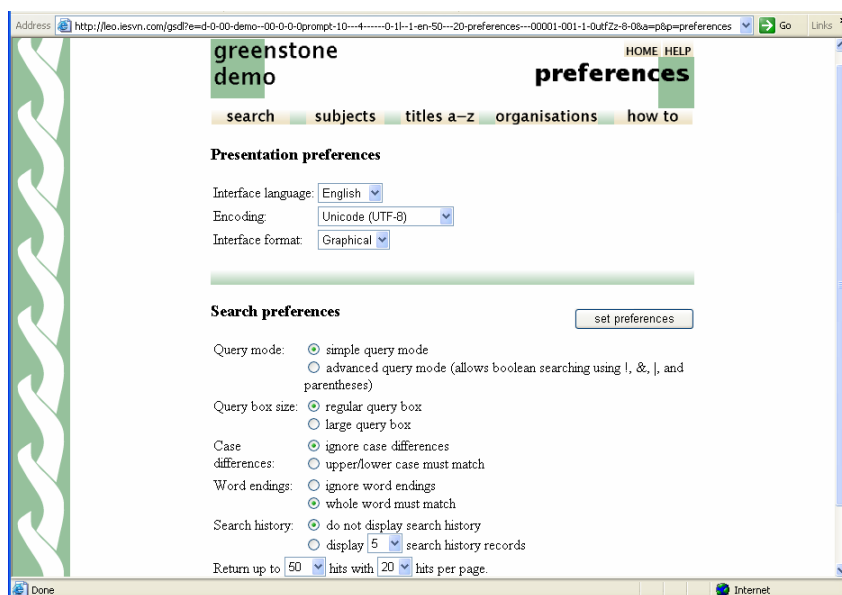
Thuộc tính Bộ tài liệu

Một số Bộ tài liệu bao gồm nhiều bộ tài liệu nhỏ hơn. Mà ta có thể tìm kiếm chúng một cách độc lập hoặc coi chúng như là một. Khi đó, bạn có thể chọn tìm kiếm luôn thư mục con trên trang *Thuộc tính*.

Thuộc tính ngôn ngữ

Mỗi Bộ tài liệu đều có ngôn ngữ thể hiện riêng, nhưng bạn vẫn có thể chuyển sang ngôn ngữ khác nếu thích. Bạn có thể thay đổi kiểu mã hoá mà phần mềm Greenstone đã sử dụng để hiển thị kết quả ra ngoài. Khi chuyển sang những biểu đồ mã hoá khác nhau có thể cho kết quả tìm kiếm tốt hơn. Tất cả các bộ tài liệu đều cho phép bạn chuyển từ giao diện đồ họa dạng chuẩn thành dạng văn bản. Điều này đặc biệt hữu ích cho người sử dụng khi các giao diện đồ họa trên màn hình không được rõ, người dùng có thể điều chỉnh phông chữ màn hình lớn hơn hoặc chuyển sang dạng âm thanh

Hình 3:
Trang thuộc tính



Thuộc tính trình bày

Tùy thuộc vào từng bộ tài liệu mà có thể các khả năng thay đổi thuộc tính trình bày là khác nhau. Các bộ sưu tập được trình trên trang Web sẽ cho phép bạn tắt thanh điều khiển của Greenstone trên đầu mỗi trang tài liệu, để khi bạn tìm đến đúng trang web cần tìm dưới dạng không có thanh điều khiển Greenstone, bạn phải bấm vào nút “back” để quay trở lại trạng thái tìm kiếm ban đầu.

Bộ tài liệu này cũng cho phép bạn ngăn chặn những cảnh báo khi nhấp chuột vào liên kết thoát khỏi bộ tài liệu của Thư viện và kết nối vào chính trang web đó. Trên một số bộ sưu tập web, bạn có thể quyết định có nên kết nối vào chính trang Web hay là chỉ sử dụng bản sao của trang web đó của Thư viện.

Thuộc tính tìm kiếm

Xem hình 3 để nhìn thấy giao diện trang thuộc tính tìm kiếm, ta thấy cặp nút đầu tiên dùng để chuyển đổi sang chế độ Tìm kiếm cấp cao như đã trình bày ở trên, cho phép bạn thực hiện câu truy vấn bằng các toán tử liên kết kiểu Logic AND (&), OR (|), NOT (!).

Cặp nút tiếp theo dùng để thay đổi kích cỡ hộp truy vấn, với chức năng này bạn có thể điền vào hộp tìm kiếm cả một đoạn tài liệu giúp cho quá trình kiểm tra nhanh hơn. Hai cặp nút tiếp sẽ xác định nội dung tìm kiếm thông qua việc điều chỉnh chức năng “*case differences*” nghĩa là tìm kiếm có quan tâm đến chữ viết hoa hay không và chức năng “*word ending*” có bỏ qua phần hậu tố của từ hay không.

Bạn cũng có thể bật hoặc tắt chức năng tìm kiếm “History” đã được miêu tả chi tiết phía trên. Cuối cùng là chức năng điều khiển số lượng kết quả đưa về trên mỗi trang, và số lượng kết quả tối đa trả ra cho một lần tìm kiếm.

3. TẬP HỢP CÁC TÀI LIỆU

Thu thập tài liệu là tính năng của phần mềm Greenstone sẽ giúp bạn tạo ra bộ tài liệu mới, sửa chữa, bổ sung hay xóa bỏ bộ tài liệu đã có.

Bạn sẽ được hướng dẫn cách tập hợp tài liệu thông qua các trang web và yêu cầu những thông tin cần thiết từ những trang web đó. Ngoài ra bạn cũng có thể xây dựng các bộ tài liệu bằng các dòng lệnh. Đọc hướng dẫn chi tiết trong phần đầu của tài liệu Hướng dẫn phát triển phần mềm.

Việc xây dựng và phân phối bộ tài liệu còn có những nhiệm vụ riêng mà bạn cần phải nắm rõ trước khi tiến hành quá trình tổng hợp tài liệu. Có những tài liệu phát hành có bản quyền: Bạn có thể truy cập tài liệu này không có nghĩa là bạn có thể để cho những người khác tự do xem tài liệu. Có những vấn đề mang tính xã hội: Các bộ sưu tập phải phù hợp với thuần phong mỹ tục ở nơi phát hành. Có những vấn đề chủng tộc, những thứ đơn giản là chúng ta không thể phát hành rộng rãi. Hãy thận trọng với sức mạnh của thông tin và sử dụng chúng một cách đúng đắn

Để tiến hành việc thu thập tài liệu, nhấp chuột vào nút liên kết trang chủ của Thư viện. Cấu trúc của một tập tài liệu được quyết định trong quá trình xây dựng, bao gồm các việc như: Định dạng tài liệu nguồn, cách trình bày dữ liệu lên màn hình, dữ liệu mô tả chương trình (metadata), , danh mục tìm kiếm toàn phần, các trình bày kết quả. Khi đã xong thì việc thêm vào tài liệu mới có cùng định dạng với tài liệu đã có là hoàn toàn dễ dàng, cùng dữ liệu mô tả chương trình.

Một bộ tài liệu khi được xây dựng có những chức năng cơ bản sau:

1. Tạo một bộ tài liệu mới có cùng cấu trúc với bộ tài liệu đang tồn tại;
2. Tạo một bộ tài liệu mới có cấu trúc khác với bộ tài liệu đang tồn tại;
3. Thêm dữ liệu mới vào bộ tài liệu đang tồn tại;
4. Sửa đổi cấu trúc bộ tài liệu đã có;
5. Xóa bộ tài liệu ;
6. Ghi bộ sưu tập đã có lên đĩa CD-ROM tự khởi động

Ở hình 4 sẽ cho thấy chức năng tập hợp tài liệu được sử dụng để tạo ra bộ tài liệu mới. Trường hợp này là tập hợp các file HTML được lưu trữ nội bộ. Đầu tiên bạn sẽ quyết định xem sẽ làm việc với bộ tài liệu đã có hay tạo một cái mới. Bạn sẽ phải chọn 1 trong 2 nút tùy chọn bên trên và từ 3 đến 6 tùy chọn nữa trong phần tiếp theo ở hình 4a, để tạo ra bộ tài liệu mới.

3.1. Logging in

Cần phải thực hiện quá trình log in (nhập username và password) trước khi tiến hành quá trình tập hợp tài liệu. Nói chung người ta thường sử dụng các trình duyệt web để truy cập vào hệ thống xây dựng bộ tài liệu từ máy tính, hoặc trên máy chủ. Dĩ nhiên không phải bất cứ người nào cũng có thể tham gia vào việc xây dựng hệ thống Thư viện số này (Điều này hoàn toàn hợp lý). Vì thế Thư viện Greenstone còn có một hệ thống bảo mật, buộc những ai muốn tham gia vào quá trình thiết kế Thư viện điều phải Log in vào

mạng. Một hệ thống trung tâm sẽ đáp ứng cho những người có nhu cầu xây dựng Thư viện và sử dụng dịch vụ đó để tạo thêm những bộ tài liệu khác. Tương tự như vậy, nếu như bạn đang sử dụng phần mềm Greestone trên máy tính, bạn cũng có thể xây dựng một tập tài liệu sử dụng nội bộ, nhưng cần thiết phải qua quá trình log in bởi vì một số người sử dụng phần mềm Greenstone trên máy tính của bạn không được tham gia vào công việc thiết kế xây dựng nếu không được phép.

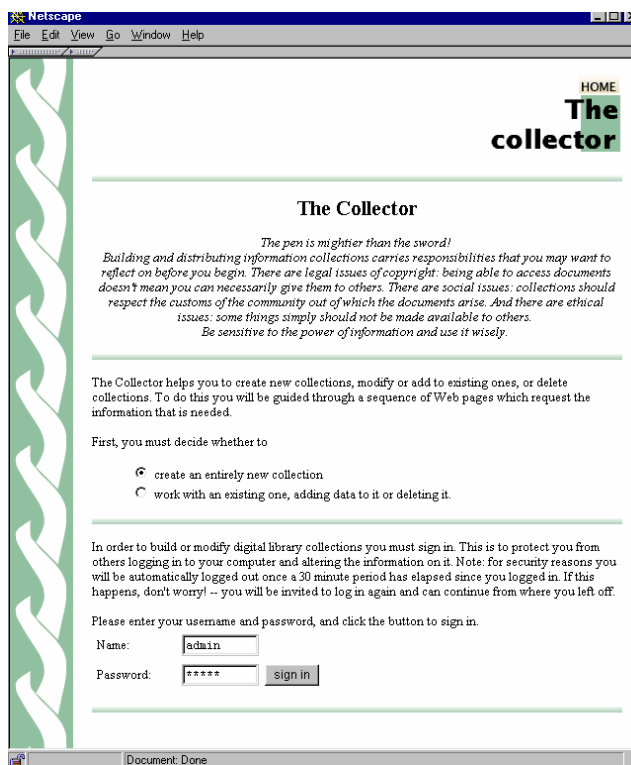
3.2. Cấu trúc Dialog

Sau khi hoàn thành quá trình log in, bạn sẽ phải thực hiện trình tự các bước sau để xây dựng bộ sưu tập.

1. Thông tin bộ sưu tập
2. Dữ liệu nguồn
3. Cấu hình bộ sưu tập
4. Xây dựng bộ sưu tập
5. Trình bày bộ sưu tập

Trước tiên là xác định tên và các thông tin có liên quan đến bộ sưu tập. Ở bước hai chỉ ra nơi để lấy dữ liệu bộ sưu tập. Bước thứ ba bạn có thể điều chỉnh cấu hình bộ sưu tập, rất tiện dụng khi bạn đã sử dụng thành thạo Greenstone. Bước thứ tư thực hiện toàn bộ công việc. Trong suốt quá trình xây dựng bộ sưu tập hệ thống sẽ tạo các chỉ mục và thu thập những thông tin cần thiết để bộ sưu tập có thể hoạt động. Bước thứ năm là trình bày bộ tài liệu vừa tạo ra.

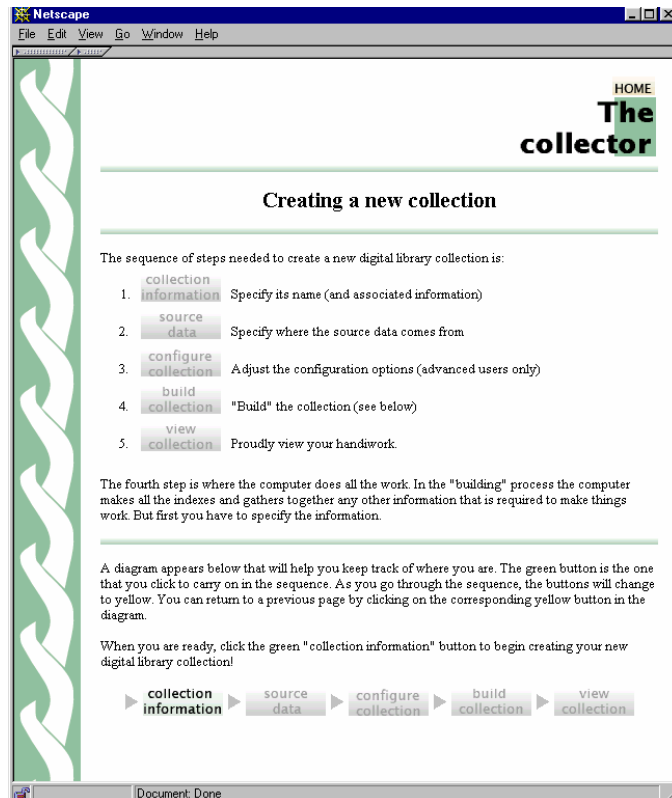
Hình 4a: Sử dụng chức năng tập hợp tài liệu để xây dựng bộ sưu tập mới



Cả năm bước này đều được lần lượt trình bày tại các nút màu xám ở phía dưới màn hình trong hình 4b và cả ở những trang khác. Điều này giúp cho

người sử dụng biết được họ đang thực hiện giai đoạn nào của quá trình. Những nút này khi được nhấp chuột vào sẽ hiện màu xanh (Thông tin bộ sưu tập (*Collection information*) ở hình 4b), còn những nút có màu xám thì không được kích hoạt – Hình 4b. Những nút màu vàng nghĩa là bạn đã thực hiện xong giai đoạn đó, người dùng có thể quay trở lại bước trước đó bằng cách nhấp chuột vào nút vào vàng tương ứng trong sơ đồ. Quá trình này được tạo thành một mô hình sau đó phần “wizards” sẽ hướng dẫn người dùng thực hiện từng bước trong việc cài đặt phần mềm mới.

Hình 4b (Tiếp theo)



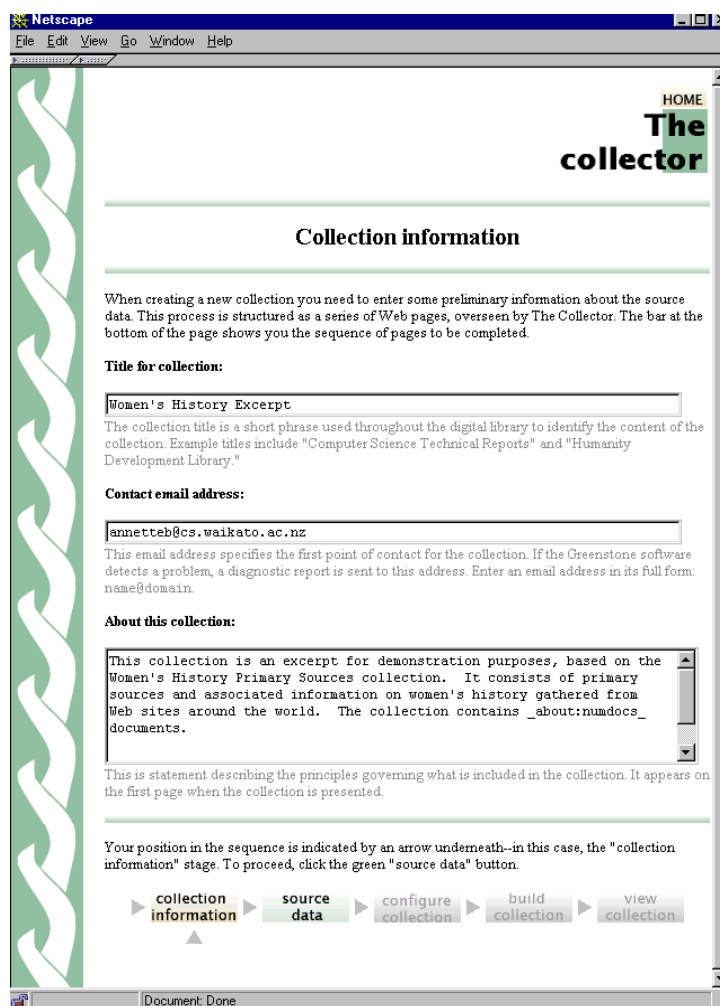
3.3. Thu thập thông tin

Bước tiếp theo là việc thu thập thông tin – Hình 4c. Khi tạo bộ sưu tập mới bạn cần phải nhập những thông tin sau:

- Tên bộ sưu tập
- Địa chỉ liên lạc (Email)
- Bảng tóm tắt mô tả bộ tài liệu

Tên bộ sưu tập là cụm từ ngắn gọn, súc tích được sử dụng trong Thư viện để nhận dạng nội dung bộ sưu tập. Địa chỉ E-mail được cung cấp để người dùng liên lạc với bạn khi gặp sự cố với bộ sưu tập. Khi gặp sự cố, một báo cáo sự cố sẽ được gửi đến địa chỉ này. Cuối cùng là bảng tóm tắt nội dung, miêu tả những chức năng chính của bộ sưu tập. Nó xuất hiện dưới tiêu đề About this collection trên trang đầu khi bộ sưu tập được hiển thị.

Hình 4 (Tiếp theo)

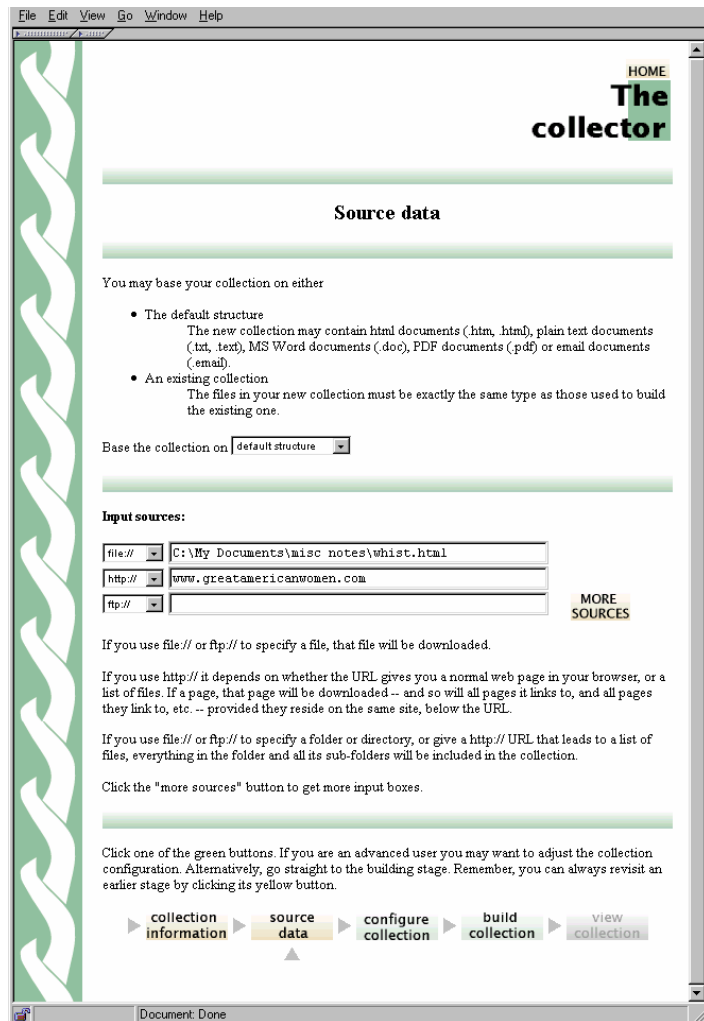


Vị trí hiện tại của người dùng cho biết người dùng đang ở giai đoạn nào của quá trình tạo bộ sưu tập, là những trang cuối của màn hình – Hình 4c – Giai đoạn thu thập thông tin. Người dùng chuyển sang giai đoạn tiếp theo bằng cách nhấp chuột vào nút *source data* màu xanh – Hình 4d

3.4. Dữ liệu nguồn

Hình 4d – người dùng chỉ rõ dữ liệu nguồn có chứa bộ sưu tập. Bạn cũng có thể xây dựng dựa trên cấu trúc mặc định đã được cung cấp hoặc dựa trên cấu trúc bộ sưu tập đã có.

Hình 4d (Tiếp theo)



Nếu bạn chọn cấu trúc mặc định, bộ sưu tập mới sẽ chứa các tài liệu dạng HTML (Là những file có đuôi chấm .htm, html), hoặc tài liệu dạng văn bản thuần túy như (Những file có đuôi .txt, .text), Dạng Word (.doc), Dạng PDF (.pdf), or dạng E-mail (.email). Những thông tin về định dạng các loại tài liệu khác, các bạn có thể tham khảo phần 3.9 dưới đây.

Còn nếu bạn xây dựng dựa trên cấu trúc bộ sưu tập đã có, những file mới này phải có định dạng hoàn toàn giống như các file của bộ sưu tập đã có.

Lưu ý rằng một số bộ sưu tập có file nguồn không thuộc dạng chuẩn, trong khi một số bộ sưu tập khác sử dụng metadata, được xác định trong file hỗ trợ. Nếu dữ liệu nguồn của bạn thiếu những thông tin này thì một số tính năng trình duyệt có thể hoạt động sai. Nếu bạn chạy bản sao của chương trình **Demo collection** có thể những nút *subjects*, *organization*, và *how to* sẽ không hoạt động.

Có 3 hộp được xây dựng để chỉ ra đường dẫn đến dữ liệu nguồn. Nếu bạn cần thêm hộp hội thoại để ghi thêm đường dẫn thì nhấp chuột vào nút “more sources.”

Có ba loại đường dẫn:

- Tên thư mục trên hệ thống chủ (server) Greenstone (Bắt đầu với “file://”)
- Địa chỉ bắt đầu với <http://> đối với những file được download từ trang web
- Địa chỉ bắt đầu với <ftp://> đối với những file được download sử dụng thuộc tính ẩn FTP.

Nếu bạn sử dụng *file://* hay *ftp://* để xác định file, file đó sẽ được download.

Nếu bạn sử dụng <http://>, điều đó tùy thuộc vào việc URL là một trang web thông thường để sử dụng, hoặc là một danh sách các file. Nếu là một trang thì trang đó sẽ được tải về - vì thế tải tất cả các trang được liên kết từ trang này – Tất cả các thư mục và các thư mục con của nó và vì thế chúng ở cùng dưới một trang URL

Nếu sử dụng *file://* hoặc *ftp://* để dẫn đến một thư mục, hoặc cho một đường dẫn <http://> URL dẫn đến một danh sách các file, thì tất cả các danh mục và các danh mục con đều chứa tải về bộ sưu tập.

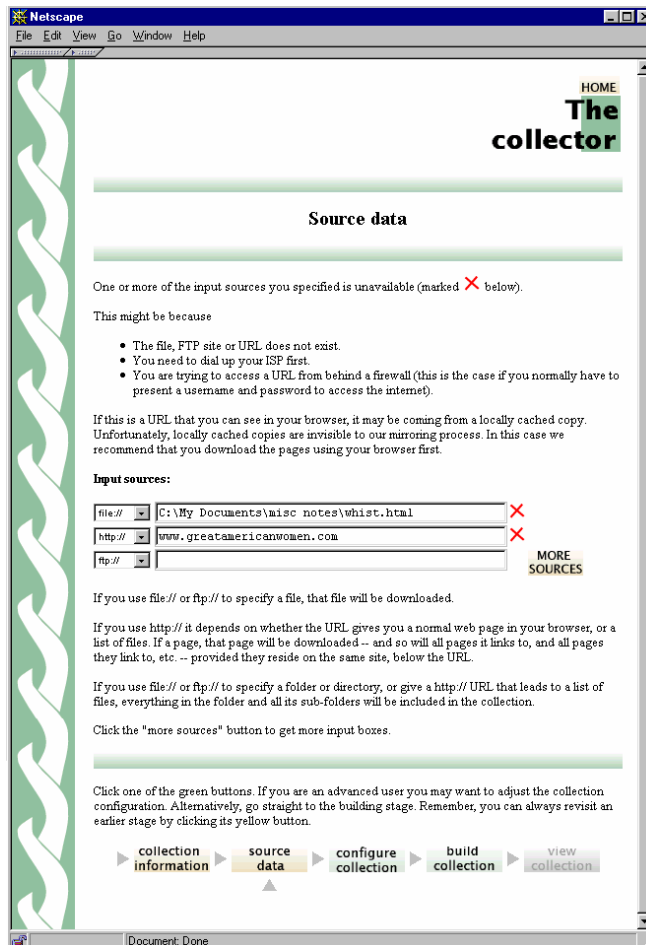
Bạn có thể đưa ra nhiều hơn một đường dẫn.

(Hình 4d) Một bộ sưu tập mới sẽ chứa các tài liệu được lấy từ hệ thống các file nội bộ cũng như điều khiển các trang web sẽ được nhân bản trong quá trình xây dựng.

Khi bạn nhấp chuột vào nút *configure collection* để chuyển sang trang kế tiếp, Chức năng Collector sẽ kiểm tra tất cả các nguồn dữ liệu đầu vào mà bạn chỉ đến đường dẫn. Việc kiểm tra này sẽ mất vài giây hoặc ngay cả vài phút, nếu càng chỉ đến nhiều đường dẫn. Nếu một trong các đường dẫn được chỉ định là không hợp lệ thì một trang web tương tự hình 4e sẽ được trình bày, các đường dẫn không hợp lệ sẽ được đánh dấu x (Trong hình này là cả hai đường dẫn).

Các đường dẫn không hợp lệ là vì:

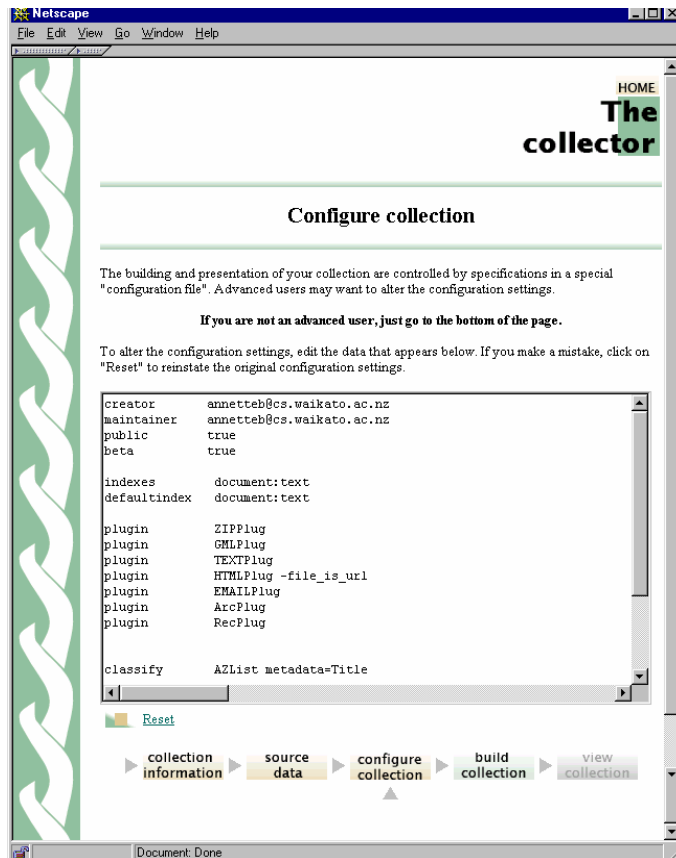
Hình 4 (Tiếp theo)



- Các file, các trang FPT, URL không tồn tại
- Bạn cần phải kết nối vào ISP trước
- Bạn đang cố truy cập vào Internet mà không qua firewall

Ở trường hợp cuối là khó hiểu nhất. Thông thường khi truy cập vào mạng bạn phải nhập username và password nhưng thỉnh thoảng bạn cũng có thể kết nối thẳng vào mạng mà không cần nhập username và password, nhưng URL sẽ thông báo là không hợp lệ. Rất có thể là trang được trình bày là trang được truy xuất từ các bản sao trong mạng nội bộ. Nhưng các bản sao này thì collector lại không nhìn thấy. Trong trường hợp này - Hình 4f, collector giới thiệu những trang mà bạn đang tải về.

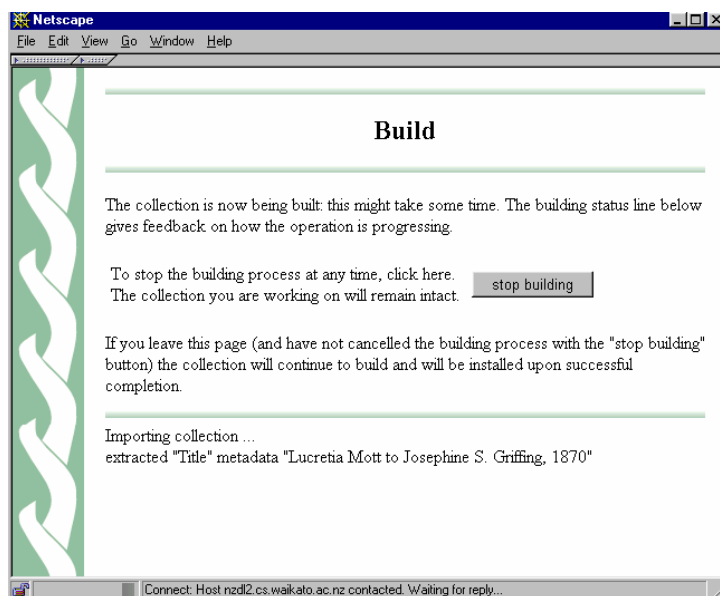
Hình 4f



3.5. Cấu hình Bộ sưu tập

(Hình f) Cấu trúc và cách trình bày tất cả các bộ sưu tập được xác định trong một file cấu hình đặc biệt (hình trên). Người dùng được phân quyền có thể sử dụng trang này để sửa đổi cấu hình, quá trình này có liên quan trực tiếp đến giai đoạn cuối. Lưu ý rằng bước này hoàn toàn có thể bỏ qua.

Hình 4 (g)



Trong phần ví dụ của chúng ta, người dùng có thể tạo một sự thay đổi nhỏ để mặc định file cấu hình bao gồm `flag file_is_url` với file đính kèm HTTP. Flag này tạo ra metadata URL được chèn vào mỗi tài liệu dựa trên quy ước tên file được chọn bởi quá trình theo dõi. Metadata này được sử dụng trong bộ sưu tập cho phép người đọc tham khảo được nguồn dữ liệu gốc hơn là bản sao nội bộ.

3.6. Xây dựng Bộ sưu tập

Hình 4g cho thấy giai đoạn xây dựng bộ sưu tập. Cho đến lúc này mọi công việc chỉ được ghi vào một tập tin tạm. Quá trình xây dựng bộ sưu tập là nơi công việc thật diễn ra.

Trong quá trình xây dựng cả danh mục tìm kiếm và trình duyệt đều được xây dựng theo chỉ dẫn trong file cấu hình bộ sưu tập. Quá trình xây dựng mất khoảng từ vài phút đến hàng giờ, tùy thuộc vào kích cỡ bộ sưu tập và tốc độ của máy tính. Một số bộ sưu tập có kích thước lớn có thể mất một ngày hoặc hơn mới xây dựng xong.

Khi bạn đạt đến giai đoạn này, cứ năm giây một lần, một dòng trạng thái ở dòng cuối của trang màn hình sẽ cho biết quá trình hoạt động đang diễn ra như thế nào. Ví dụ, tin nhắn hình trong hình 4f cho biết tên của dữ liệu đang được lấy ra từ một tập tin nguồn.

Thông báo xuất hiện khi mà các tập tin nguồn hoặc URL nhập vào không tồn tại, hoặc có tồn tại mà không có plugin, hoặc có plugin nhưng không tìm được file liên kết, chẳng hạn như file được đặt trong tài liệu HTML. Mục đích là khi bạn giữ cho cửa sổ này luôn mở, bạn có thể theo dõi quá trình hoạt động. Bất cứ lỗi làm gián đoạn quá trình sẽ được lưu lại ở đây.

Bạn có thể dừng quá trình xây dựng này bất cứ lúc nào bằng cách nhấp chuột vào nút *Dừng xây dựng* - Hình 4g. Nếu bạn rời khỏi trang web này mà không nhấn vào nút này thì quá trình vẫn tiếp diễn và bộ sưu tập mới sẽ được cài đặt sau khi quá trình xây dựng bộ sưu tập cũ hoàn thành.

3.7. Trình bày bộ sưu tập

Khi bộ sưu tập đã được xây dựng và cài đặt xong, dãy nút ở phía dưới hình 4b, sẽ xuất hiện ở phía dưới Hình 4g, với nút view collection được kích hoạt để liên kết người dùng với bộ sưu tập mới xây dựng thành công.

Cuối cùng, có một chức năng là gửi E-mail đến địa chỉ liên lạc của bộ sưu tập và đến người quản trị hệ thống bất cứ khi nào bộ sưu tập được tạo ra (hoặc sửa đổi). Điều này cho phép những người có trách nhiệm kiểm tra khi có những thay đổi, và theo dõi những việc đang xảy ra với hệ thống. Các điều kiện mặc định là không hoạt động nhưng có thể được chỉnh cho kích hoạt bằng thay đổi các file cấu hình main.cfg (xem phần *Greenstone Digital Library Developer's Guide*. Phần 4).

3.8. Làm việc với bộ sưu tập đã tồn tại

Khi bạn sử dụng chức năng *collector* bạn phải xác định được rằng bạn muốn tạo ra một bộ sưu tập mới hoàn toàn hay làm việc với bộ sưu tập đã tồn tại, thêm hay xóa bộ sưu tập. Bằng cách tạo ra các cấu trúc tìm kiếm và trình duyệt tự động, việc thêm thông tin vào bộ sưu tập đã tồn tại trên Greenstone diễn ra rất dễ dàng. Bởi vì không có liên kết nào được nhập vào bằng tay, khi một bộ tài liệu mới có cùng một dạng, hệ thống sẽ tự nhận biết và trộn vào chung với bộ sưu tập đã có.

Để làm việc với bộ sưu tập có sẵn, trước tiên bạn phải chọn bộ sưu tập từ danh mục bộ sưu tập được cung cấp. Một số bộ sưu tập được đánh dấu là không thể sửa đổi thì sẽ không xuất hiện trong danh mục liệt kê này. Với bộ sưu tập, bạn có thể.

- Thêm dữ liệu mới và tái tạo lại bộ sưu tập
- Chỉnh sửa file cấu hình của bộ sưu tập
- Xóa toàn bộ bộ sưu tập
- Truy xuất bộ sưu tập ra CD-ROM

Thêm dữ liệu mới

Tất cả những file bạn chỉ định sẽ được thêm vào bộ sưu tập, do đó bạn phải chắc rằng nó chưa tồn tại trong bộ sưu tập, nếu không bộ sưu tập sẽ chứa hai file giống hệt nhau. Các file này được nhận dạng bằng đường dẫn đầy đủ,

những trang web hay địa chỉ web tuyệt đối. Bạn phải xác định rõ những thư mục hay những file như khi bạn xây dựng bộ sưu tập mới.

Nếu khi thêm dữ liệu mới vào bộ sưu tập mà vì một số nguyên nhân có thể làm cho quá trình này không thực hiện được, có thể là do bộ sưu tập cũ chưa thể thay đổi

Chỉnh sửa file cấu hình

Những người sử dụng thành thạo có thể chỉnh sửa tập tin cấu hình của bộ sưu tập, như là họ làm khi xây dựng bộ sưu tập mới. Phần 4 dưới đây sẽ giải thích rõ hơn về cách cài đặt cấu hình.

Xoá bộ sưu tập

Bạn sẽ được thông báo để xác nhận xem có thật sự muốn xoá bộ sưu tập không. Một khi đã xoá thì Greenstone không thể hồi phục lại.

Truy xuất bộ sưu tập

Bạn có thể xuất bộ sưu tập ra ngoài dưới một đĩa CD-ROM tự động cài đặt và sử dụng của Greenstone dành cho Windows. Vì các phần mềm thương mại để tạo các CDs tự cài đặt rất tốn kém, công cụ này bao gồm cả một phần cài đặt riêng.

Khi bạn truy xuất bộ sưu tập, hộp hội thoại sẽ thông báo cho bạn biết tên thư mục được xuất đến. Toàn bộ nội dung của thư mục sẽ được ghi lên CD-ROM bằng các công cụ ghi chép CD-writing thông thường. Vì cấu hình của Windows rất đa dạng, Greenstone rất khó kiểm tra và sửa lỗi cho tất cả các trường hợp. Nếu bạn bị lỗi và bạn có được một bộ phần mềm thương mại dùng để cài đặt (ví dụ như. InstallShield), bạn có thể sử dụng nó để tạo ra đĩa CD-ROMs từ những thông tin mà bộ sưu tập cung cấp. Thư mục truy xuất đến được đề cập phía trên chứa 4 file liên quan đến quá trình cài đặt và 3 thư mục con chứa toàn bộ phần mềm và bộ sưu tập. Bạn hãy xóa bốn file này và sử dụng InstallShield để tạo một ảnh CD-ROM dùng trong việc cài đặt những thư mục và tạo một biểu tượng dẫn đến chương trình *gsdl\server.exe*.

3.9. Định dạng tài liệu

khi xây dựng bộ sưu tập, phần mềm Greenstone xử lý các định dạng khác nhau của tài liệu gốc bằng dùng một plugin viết riêng cho dạng tài liệu đó. Các plugin được ghi rõ trong file cấu hình bộ sưu tập. Nói chung phần mềm Greenstone sử dụng tên file để quy định thuộc tính của tài liệu ví dụ như *foo.txt* được quy định là file văn bản, còn *foo.html* là dạng file HTML, và *foo.doc* dạng file Word.

Đây là bảng tóm tắt các file đính kèm (plugins) có sẵn dùng để định dạng các tài liệu. Để biết thêm chi tiết về các plugins này và các plugins khác ít thông dụng hơn, các bạn hãy tham khảo phần tài liệu tên *Greenstone DigitalLibrary Developer's Guide*.

TEXTPlug (*.txt, *.text)

TEXTPlug là dạng file văn bản thuần túy. Nó thêm tiêu đề vào phần metadata dựa trên dòng đầu của file.

HTMLPlug (*.htm, *.html; also .shtml, .shm, .asp, .php, .cgi)

HTMLPlug xử lý các file HTML. Nó tách *title* metadata dựa trên phần đuôi <title>; các metadata khác trình bày cách sử dụng cú pháp metatag của HTML cũng được tách ra. Plugin này cung cấp nhiều chọn lựa khác nhau, tham khảo thêm trong tài liệu *Greenstone Digital Library Developer's Guide*

WORDPlug (*.doc)

WORDPlug lấy các tài liệu dạng Microsoft Word. Có nhiều tài liệu dạng Word khác nhau và ngay cả các chương trình của Microsoft gặp trục trặc khi chuyển đổi. Greenstone sử dụng các chương trình độc lập để chuyển đổi các file dạng Word sang dạng HTML. Đối với một số tài liệu thuộc các phiên bản trước đó, hệ thống phải sử dụng đến thuật toán phân tích giản đơn để tìm tất cả các chuỗi ký tự trong file nguồn

PDFPlug (*.pdf)

PDFPlug nhận tài liệu dạng PDF (Adobe's Portable Document Format). Giống như WORDPlug, nó sử dụng một chương trình độc lập, trong trường hợp này là chương trình *pdftohtml* để chuyển đổi file dạng PDF sang dạng HTML.

Đối với WORDPlug, theo mặc định thì bộ sưu tập sẽ trình bày file tương tự dạng HTML, khi người dùng nhấp chuột vào biểu tượng *document*; tuy nhiên chuỗi ký tự định dạng trong file cấu hình bộ sưu tập có được điều chỉnh để cho phép người dùng truy cập vào file gốc PDF (các bạn nên sử dụng cách này). Giống như trên, chỉ cần thay thế các tag <link> ... </link> bằng <srclink> ... </srclink>

Chương trình *pdftohtml* có lỗi trên một số file PDF. Chuyện gì sẽ xảy ra khi mà quá trình chuyển đổi diễn ra trong một thời gian dài, và thường xuất hiện thông báo lỗi liên quan đến quá trình chuyển đổi trên màn hình. Nếu có xảy ra lỗi, thì giải pháp duy nhất là loại bỏ tài liệu bị lỗi ra khỏi bộ sưu tập. Thêm vào đó, PDFPlug không xử lý được các tập tin PDF bị mã hóa.

PSPlug (*.ps)

PSPlug lấy tài liệu dưới dạng PostScript. Nó làm việc tốt nhất nếu một chương trình *ps2ascii* của Linux đã được cài đặt sẵn. Chương trình này có sẵn trên hầu hết các cài đặt Linux, nhưng không phải Windows. Nếu chương trình này không có sẵn, PSPlug phải dùng đến thuật toán phân tích văn bản đơn giản.

EMAILPlug (*.email)

EMAILPlug xử lý các tập tin chứa E-mail, và xử lý những định dạng email thông thường được dùng trên Netscape, Eudora, and Unix. Mỗi tài liệu nguồn được kiểm tra xem nó có chứa một hoặc nhiều email khác gộp chung

lại hay không, nếu có, phần nội dung của email sẽ được xử lý. Plugin sẽ thu thập metadata về *Subject*, *To*, *From*, và *Date*.

Tuy nhiên, plugin này chưa xử lý được các email mã hóa dưới dạng MIME. Trên lý thuyết là có thể nhưng trên thực tế, dạng mã hóa này hơi lạ so với các dạng khác.

ZIPPlug (.gz, .z, .tgz, .taz, .bz, .zip, .tar)

Plugin ZIPPlug xử lý các file dưới dạng nén hoặc lưu trữ: gzip (*.gz*, *.z*, *.tgz*, *.taz*), bzip (*.bz*), zip (*.zip* *.jar*), và tar (*.tar*). Nó giải nén bằng các chương trình *gunzip*, *bunzip*, *unzip*, và *tar* theo chuẩn Linux utilities. ZIPPlug không hoạt động trên hệ điều hành Windows.

4. CHỨC NĂNG QUẢN TRỊ

Các phần mềm Greenstone đều có tính năng quản trị. Để sử dụng tính năng này, bạn hãy nhấp chuột vào mục *administration* ở trang đầu tiên.

Trang đầu tiên sẽ trình bày thông tin của mỗi bộ sưu tập trên hệ thống. Lưu ý rằng, có một số bộ sưu tập có tính cá nhân sẽ không trình bày trên “home page” và mang những tên riêng hoặc là tên đầy đủ, nếu là trang công cộng mà không được trình bày hoặc là không chạy, nhấp chuột vào bảng tóm tắt của bộ sưu tập (cột đầu tiên trong - Hình 5) sẽ cho ta biết một số thông tin về bộ sưu tập lấy từ file cấu hình của bộ sưu tập hoặc từ các cấu trúc khác bên trong của bộ sưu tập. Nếu bộ sưu tập đều có tính công cộng và hoạt động được, khi nhấp chuột vào cột *fullname* (cột thứ 2 – Hình 5) bạn sẽ có được thông tin chi tiết về bộ sưu tập cần tìm.

Bộ sưu tập mà chúng ta xây dựng trong phần 3.2–3.6 có tên là *wohiex*, và tên đầy đủ là *Women's History Excerpt*, ở dòng cuối của hình 5. Hình 6 sẽ trình bày thông tin chi tiết về bộ sưu tập khi chúng ta nhấp chuột vào cột *fullname*. Phần đầu tiên sẽ cho biết thông tin về file cấu hình và kích thước bộ sưu tập (khoảng 1000 tài liệu, 1 triệu từ, chiếm trên 6Mb). Phần kế sẽ chứa thông tin nội bộ liên quan đến giao thức liên lạc đến các bộ sưu tập. Ví dụ với chức năng *filter options* dùng cho “*QueryFilter*” sẽ trình bày các hộp tùy chọn và các giá trị có thể được dùng khi truy vấn đến bộ sưu tập.

Chức năng quản trị sẽ trình bày các thông tin về quá trình cài đặt và cho phép được sửa đổi. Nó giúp cho quá trình kiểm tra lỗi nhanh chóng và dễ dàng hơn trong việc ghi lại những lỗi trong máy. Nó có thể phân quyền cho người dùng để cho phép người khác xây dựng tiếp bộ sưu tập và thêm thông tin mới vào bộ sưu tập sẵn có. Tất cả những tính năng này đều được truy cập trực tiếp từ danh mục liệt kê ở bên trái trong hình 5.

Hình 5 Tính năng quản trị của Greenstone

Administration

Maintenance and administration services available include:

- view on-line logs
- create, maintain and update collections
- access technical information such as CGI arguments

These services are accessed using the side navigation bar on the lefthand side of the page.

Collection Status

Collections will only appear as "running" if their build.cfg files exist, are readable, contain a valid builddate field (i.e. > 0), and are in the collection's index directory (i.e. NOT the building directory).

click *abbrev.* for information on a collection
click *collection* to view a collection

abbrev.	collection	public?	running?
acrodemo	acrodemo	yes	yes
bibdemo	greenstone demo	yes	no
bylw	bylw	yes	no
cnrub	cnrub	no	yes
csbib	Computer Science Bibliographies	yes	yes
csbib.old	Computer Science Bibliographies	yes	no
date	Date	yes	yes
demo	greenstone demo	yes	yes
dmail	email plugin demo	yes	yes
election	The Election Collection	no	yes
fao.org	www.fao.org	yes	yes
fi1998	FAO on the Internet (1998)	yes	yes
folktale	folktales: language extraction demo	yes	yes
forestry	www.fao.org	yes	yes
gsdldocs	Greenstone Source and Documentation	yes	yes
hcibib2	hcibib2	no	yes
hcibib4	HCI Bibliography 4	yes	yes
hemuka	He Muka	no	yes
hemuka2	He Muka	no	yes
iantes	ian's test	yes	no
knowbase	Knowbase	yes	yes
localweb	localweb	yes	no
mhl	Medical and Health Library	yes	yes
niupepa	Niupepa: Maori Newspapers	yes	yes
niupepa_places	Niupepa: Maori Newspapers	yes	yes
ohst	Hamilton Public Library Youth Oral History Collection	yes	yes
rweg	rweg	yes	yes
schoolj	The New Zealand School Journal	no	yes
scms	Search Computing, Mathematics and Statistics	yes	yes
snew	snew	no	no
tang	tang	yes	no
tescol	test collection	yes	no
testword	Word plugin demo	yes	yes
tidbits	TidBITS	yes	yes
unu	United Nations University	yes	yes
whist	Women's History Primary Source Documents	yes	no
whistbuildingimages	whistbuildingimages	no	no
whohex	Women's History Excerpt	yes	yes
wordtest	Word plugin demo	yes	yes
wordtest.tar.gz	wordtest.tar.gz	no	no

Hình 6: Thông tin về bộ sưu tập
Women's History Excerpt

The screenshot shows the Greenstone interface for a collection named 'wohsex'. The left sidebar contains navigation links for admin, configuration files, logs, user management, and technical information. The main content area is titled 'Collection info' and displays various metadata fields. Below the metadata, there are three sections for filter options: 'BrowseFilter', 'NullFilter', and 'QueryFilter', each with a table of configuration parameters.

Collection info

Collection info

collection name "wohsex"
 host ""
 port "0"
 is public? true
 is beta? true
 build date "978487241"
 interface
 languages
 collection .documenttext documents
 metadata This collection is an excerpt for demonstration purposes, based on the Women's History Primary Sources collection. It consists of primary sources and associated information on women's history gathered from Web sites around the world. The collection contains _aboutname&oc=documents&ln collectionname Women's History Excerpt iconcollection

format info
building info
 number of documents "1073"
 number of sections "1073"
 number of words "1038463"
 number of bytes "6294492"
 preferred receptionist ""

Filter options for "BrowseFilter"

option name	type	repeatable	default value	valid values
"EndResults"	integer	one per query	"-1"	"-1", "10000"
"ParentNode"	string	one per query	""	
"StartResults"	integer	one per query	"1"	"1", "10000"

Filter options for "NullFilter"

option name	type	repeatable	default value	valid values

Filter options for "QueryFilter"

option name	type	repeatable	default value	valid values
"Casefold"	boolean	one per term	"true"	"false", "true"
"CombineQuery"	enumerated	one per query	"and"	"and", "or", "not"
"EndResults"	integer	one per query	"10"	"-1", "1000"
"Index"	enumerated	one per term	"dix"	"dix"
"Language"	enumerated	one per term	""	
"MatchMode"	enumerated	one per query	"some"	"some", "all"
"Maxdocs"	integer	one per query	"200"	"-1", "1000"
"PhraseMatch"	enumerated	one per query	"some_phrases"	"all_phrases", "some_phrases", "all_docs"
"QueryType"	enumerated	one per query	"ranked"	"boolean", "ranked"
"StartResults"	integer	one per query	"1"	"1", "1000"
"Stem"	boolean	one per term	"false"	"false", "true"
"Subcollection"	enumerated	one per term	""	
"Term"	string	one per term	""	

4.1. File cấu hình

Có hai file cấu hình điều khiển hoạt động của Greenstone là file định vị *gsdlsite.cfg* và file cấu hình chính *main.cfg*.

File *gsdlsite.cfg* dùng để định vị trí phần mềm Greenstone được cài đặt. Nó được thiết kế để tách riêng với các phần khác. Ví dụ tên của thư mục cài đặt phần mềm, địa chỉ trang web của hệ thống và tính năng *fastcgi* của hệ thống có được sử dụng hay không. Các dữ liệu trong file này được mô tả trong ***Greenstone Digital Library Installation Guide***

File *main.cfg* chứa những thông tin chung cho tất cả các bộ sưu tập Greenstone, bao gồm địa chỉ E-mail của người bảo trì hệ thống, các trang trạng thái hoặc collector có được kích hoạt hay không, danh mục người dùng có được lưu trữ hay không, chứa năng “cookies” có được dùng để xác định người dùng hay không.

4.2. Logs

Có ba loại log: user log, error log và các giá trị ban đầu. Hai danh mục cuối cùng liên quan đến người bảo trì phần mềm.

Mọi hoạt động của người dùng – Bất cứ ai truy cập vào trang nào cũng đều được ghi lại bởi phần mềm Greenstone, mặc dù không ghi lại tên người sử dụng. Quá trình đăng nhập được có thể được kích hoạt hay không được kích hoạt tùy được viết bằng các dòng lệnh sau:

```
logcgiargs true
usecookies true
```

Trong file hệ thống chính, cả hai chọn lựa trên đều được mặc định là sai (false) vì thế không có quá trình đăng nhập nào được cho phép trừ khi được người quản trị mạng đặt lại thuộc tính. Dòng *logcgiargs* là dòng quyết định quá trình logging được kích hoạt hay không. Bởi vì quá trình xây dựng *usecookies* là một mã khoá nhận dạng được đánh dấu cho mỗi người dùng, mọi hoạt động của người dùng đều được theo dõi thông qua file log này.

Mỗi dòng là mỗi mẫu tin người dùng đã truy cập trang web – Ngay cả những trang được tạo ra để duyệt các file log! Nó có chứa một địa chỉ IP của máy người dùng, (b) Thời gian trong ngoặc vuông, (c) các tham số CGI trong ngoặc đơn, và (d) tên loại trình duyệt web (Netscape được gọi là “Mozilla”). Đây là một dòng mẫu, được cắt ra và chú giải cho lệnh của người dùng:

```
/fast-cgi-bin/niupepalibrary
(a) its-www1.massey.ac.nz
(b) [Thu Dec 07 23:47:00 NZDT 2000]
(c) (a=p, b=0, bcp=, beu=, c=niupepa, cc=, ccp=0, ccs=0, cl=, cm=,
cq2=, d=, e=, er=, f=0, fc=1, gc=0, gg=text, gt=0, h=, h2=, hl=1,
hp=, il=l, j=, j2=, k=1, ky=, l=en, m=50, n=, n2=, o=20, p=home,
pw=, q=, q2=, r=1, s=0, sp=frameset, t=1, ua=, uan=, ug=,
uma=listusers, umc=, umnpw1=, umnpw2=, umpw=, umug=, umun=, umus=,
un=, us=invalid, v=0, w=w, x=0, z=130.123.128.4-950647871)
(d) “Mozilla/4.08 [en] (Win95; I ;Nav)”
```

Đối số CGI, “z”, là một mã nhận dạng hoặc “cookie” phát sinh ra bởi trình duyệt người dùng: Nó bao gồm số IP của người dùng được theo dõi dựa trên thời gian họ truy cập vào Thư viện số lần đầu. File log *usage.txt* được thay thế trong thư mục *etc* trong cấu trúc file của Thư viện Greenstone (xem trong *The Greenstone Digital Library Developer’s Guide*). Khi quá trình logging được kích hoạt, mỗi hành động của người dùng đều được ghi ghi lại. Tuy nhiên chỉ 100 dòng cuối cùng của file log được hiển thị bởi *link usage* trong Hình 5.

4.3. Quản lý người dùng

Greenstone sử dụng chế độ bảo mật để kiểm soát việc sử dụng một số tính năng. Hiện chức năng này chỉ được dùng để kiểm soát những người được phép sử dụng Collector và một số chức năng quản lý. Đối với bộ sưu tập đặc biệt, vban có thể yêu cầu xác định người dùng trước khi cung cấp thông tin cho họ - ví dụ như, các bộ tài liệu có thể được bảo vệ cho từng một cá nhân, và chỉ những ai có password mới sử dụng được chúng. (Tuy nhiên, không có bộ sưu tập hiện hành nào sử dụng chức năng này).

Việc xác định người dùng được thực hiện bằng việc xác định Tên (user name) và mật khẩu (password) như trong hình 4a.

Từ trang administration, bạn có thể liệt kê, thêm hoặc bớt người sử dụng. Khả năng này dĩ nhiên là phải được quản lý: chỉ những người được có quyền quản lí mới có thể thêm tên người dùng mới, và những nhóm khác nhau với các quyền khác nhau. Hiện tại chỉ có những nhóm như: “administrator” và “colbuilder”. Thành viên của nhóm đầu tiên thì được phép thêm hay di chuyển người dùng. Nhóm thứ hai thì được phép truy cập theo các tính năng kể trên để xây dựng bộ sưu tập mới và xoá bộ sưu tập đã tồn tại.

Khi phần mềm Greenstone đã được cài đặt, chỉ những người dùng được gọi là *admin* thì mới thuộc cả hai nhóm trên. Password cho người sử dụng này được thiết lập trong suốt quá trình cài đặt. Người dùng có thể tạo ra tên và password mới cho những người dùng thuộc nhóm *colbuilder*, là nhóm có khả năng cho phép người dùng khác có thể xây dựng bộ sưu tập. Thông tin người dùng được lưu vào trong hai cơ sở dữ liệu để trong cấu trúc file Greenstone (xem the *Greenstone Digital Library Developer’s Guide*).

4.4. Thông số kỹ thuật

Các liên kết ở phía dưới dòng thông số kỹ thuật cho biết thêm những thông số cài đặt chương trình. Các liên kết *chung* cho phép truy cập đến các thông số kỹ thuật, bao gồm các thư mục mà thông tin được lưu trữ trong đó. Bảng danh mục các *protocol* ứng với mỗi loại *protocol* có khả năng hoạt động, thông tin về mỗi bộ sưu tập được cung cấp bởi protocol.

Cuối cùng, mã giao diện người dùng được gọi là “receptionist” dùng *actions* để kết nối với nhu cầu của người dùng. Những hành động này tương ứng với đối số CGI được gắn nhãn *a*. Ví dụ: nếu *a=status* thì receptionist gọi cho hiện lên trang trạng thái (*status page*). Một đề mục trong danh mục cho phép

truy cập vào danh mục vào các hành động được hỗ trợ bởi file hệ thống, và những chỉ dẫn khác để các đối số liên quan đến hoạt động xảy ra.