

SỬ DỤNG PHẦN MỀM NGUỒN MỞ THƯ VIỆN GREENSTONE ĐỂ XÂY DỰNG TÀI NGUYÊN HỌC TẬP*

Using the Greenstone Open Source Software to Build Learning Resources

ThS. Nguyễn Minh Hiệp

GD Thư viện ĐH Khoa học Tự nhiên

Đại học Quốc gia Thành phố Hồ Chí Minh

Nâng cao chất lượng đào tạo nguồn nhân lực để đáp ứng nhu cầu lao động xã hội ngày càng phát triển luôn luôn là châm ngôn của một trường đại học đồng thời là trách nhiệm của Đảng, Nhà nước và Nhân dân giao phó cho trường đại học đó. Để nâng cao chất lượng đào tạo nguồn nhân lực trong một trường đại học đòi hỏi phải nâng cao chất lượng giảng dạy của đội ngũ giảng viên. Để nâng cao chất lượng giảng dạy của đội ngũ giảng viên, cần phải nâng cao và cập nhật kiến thức chuyên môn của đội ngũ giảng viên đó. Ngày nay việc phát triển công nghệ thông tin (CNTT) và truyền thông đã làm cho việc phổ biến tri thức trở nên thuận tiện và nhanh chóng. Thực tế, công nghệ này đã và đang giúp cho đội ngũ giảng viên ngày càng nâng cao và cập nhật kiến thức chuyên môn. Tuy nhiên thông tin ngày càng trở nên quá tải. Trách nhiệm của những người làm công tác quản lý thông tin trong một trường đại học là làm thế nào để giúp cho đội ngũ giảng viên có được những thông tin có ý nghĩa và hữu ích để họ thật sự có thể nâng cao và cập nhật kiến thức chuyên môn. Với vai trò của người quản lý thông tin, chúng tôi nhận thức được giá trị mới của một thư viện đại học ngày nay là *"đáp ứng nhu cầu thông tin một cách có hiệu quả như thế nào từ rất nhiều nguồn thông qua công nghệ mới"*. Bài tham luận này nhằm kiến nghị một giải pháp giúp nâng cao và cập nhật kiến thức chuyên môn của đội ngũ giảng viên bằng công nghệ mới – công việc thiết thực nhất nhằm nâng cao chất lượng đào tạo nguồn nhân lực. Biện pháp thực hiện là sự *"tương tác giữa người sử dụng với thư viện để phục vụ chính người sử dụng"*.

Thông tin và tri thức

Đã từ lâu, đội ngũ giảng viên trong một trường đại học luôn cần đến thư viện là nơi có tổ chức để bảo quản tài liệu, sưu tập và để truy cập đến những thư viện khác. Ngày nay, thư viện là một mạng lưới cung cấp việc truy cập

đến thông tin được lưu giữ khắp mọi nơi, chúng ta đang ở giữa một kho tàng thông tin và tri thức đồ sộ. Chính vì thế mà đã có không ít người cho rằng thư viện và Internet là một, thậm chí cho rằng Internet là nhân tố đã làm cho thư viện bị lỗi thời và trở nên không cần thiết. Thật là một quan niệm sai lầm vì nói như thế chẳng khác gì cho rằng giày dép đã làm cho đôi chân trở nên thừa thãi!

Có sự khác nhau giữa thông tin và tri thức. Theo quan điểm của ngành thông tin - thư viện, tri thức là thông tin có ý nghĩa và hữu ích. Theo Branscomb, một nhà kinh tế tri thức cho rằng nếu thông tin được ví như bột mì thì tri thức chính là bánh mì. Thông tin trên Internet thiếu hẳn những đặc điểm quan trọng của việc sưu tầm và tổ chức thông tin; khác hẳn với thông tin của thư viện. Ngày nay cuộc cách mạng thông tin với kỹ thuật số đã đáp ứng nhu cầu chưa từng có về lưu trữ, tổ chức, và truy cập thông tin. Các thư viện đại học khắp nơi đang đứng trước ngưỡng cửa của Thư viện số. Ngành thông tin - thư viện chuyển sang giai đoạn Quản lý tri thức. Nói một cách nôm na, chúng ta không còn phục vụ bột mì nữa mà phải chuẩn bị những loại bánh mì để phục vụ người tiêu dùng – Thông tin cần được chọn lọc, tổ chức và phục vụ dưới dạng kỹ thuật số.

Bộ sưu tập thông tin

Theo Ian H. Witten, chuyên gia Thư viện số ĐH Waikato, New Zealand, Thư viện số là tập hợp những bộ sưu tập thông tin của các đối tượng số hoặc đã được số hóa có tổ chức và tập trung. Tập trung theo đề tài hay chủ đề và có tổ chức để thông tin dễ truy cập và lưu trữ theo những tiêu chuẩn chuyên biệt cung cấp hai khả năng chính:

- Phương thức truy cập, chọn lọc, hiển thị tài nguyên số (dành cho người sử dụng);
- Phương thức xây dựng, tổ chức và lưu hành (dành cho cán bộ thư viện).

Một bộ sưu tập thông tin bao gồm nhiều tài liệu dưới nhiều dạng thức khác nhau: văn bản, hình ảnh, âm thanh, hình ảnh động. Tài liệu là đơn vị căn bản từ đó sưu tập thông tin được xây dựng.

và Human Info NGO. Đây là phần mềm mã nguồn mở được cung cấp theo thỏa thuận của GNU General Public License.

Greenstone là phần mềm nguồn mở và tự do

Theo Richard Stallman, chủ tịch FSF (Free Software Foundation), Phần mềm tự do và Phần mềm nguồn mở là:

- Chương trình ta có thể chạy với bất cứ mục đích nào;
- Có thể chỉnh sửa cho phù hợp với yêu cầu của mình;
- Có thể chỉnh sửa và tái phân phối bản sao, có hoặc không có thu phí trên phần chỉnh sửa của mình.

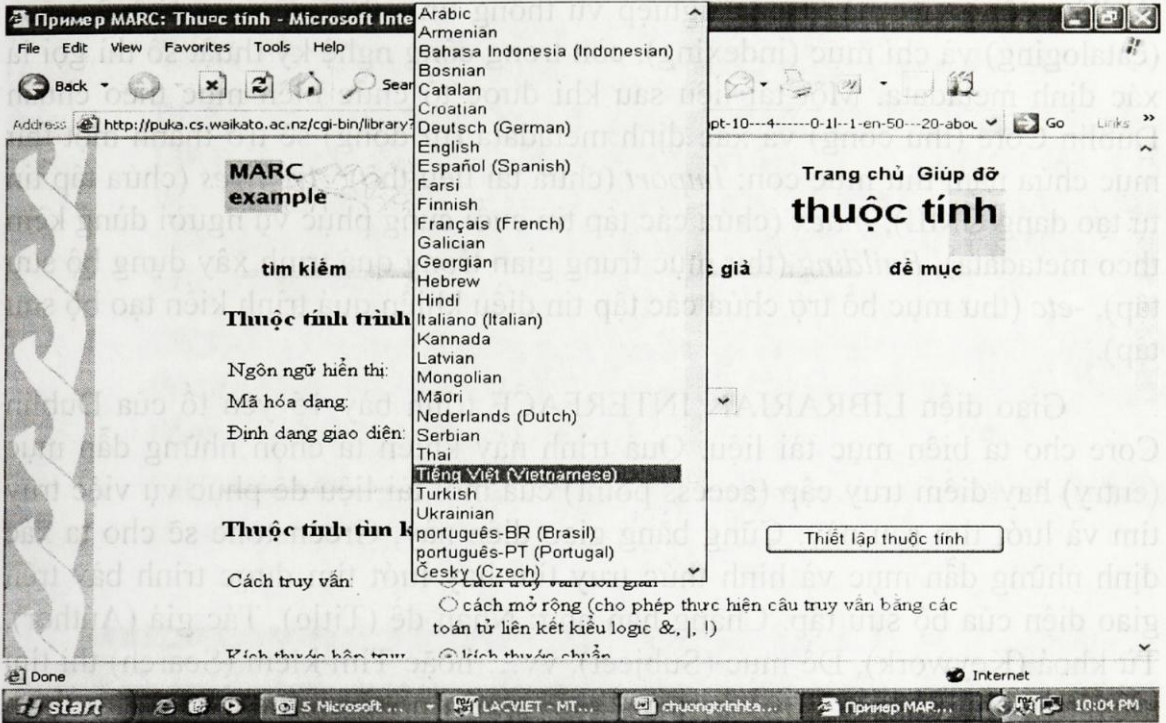
Phần mềm nguồn mở Greenstone có thể download để sử dụng miễn phí tại <http://greenstone.org>.

Greenstone là phần mềm đa ngôn ngữ.

Những bộ sưu tập thông tin được tạo ra bởi Greenstone có thể chứa một số lượng lớn những tài liệu (hàng ngàn cho đến hàng triệu) với một giao diện đồng nhất. Bộ sưu tập bao gồm từ những bài báo cho đến chương trình âm nhạc. Hiện nay, Greenstone được dùng rộng rãi trong thư viện của các trường đại học trên thế giới để hình thành Thư viện số. Phần mềm Greenstone trên CD-ROM được Liên Hiệp Quốc và những cơ quan nhân đạo khác xuất bản và phân phối cho các quốc gia đang phát triển. Greenstone được dịch sang hơn 50 ngôn ngữ bao gồm:

- o 4 ngôn ngữ nòng cốt: Anh, Pháp, Tây Ban Nha, và Nga. Ấn bản Greenstone Tiếng Pháp, Tây Ban Nha, và Nga được UNESCO bảo trì; ấn bản Tiếng Anh do Chương trình Dự án thư viện số New Zealand bảo trì. Khi chúng ta download Greenstone thì giao diện của bốn ngôn ngữ nòng cốt tự động được cài đặt; còn các ngôn ngữ khác cần có sự tham gia của người bảo trì hệ thống Greenstone.
- o Các ấn bản của các ngôn ngữ đã hoàn chỉnh bao gồm: Arabic, Armenian, Ba Lan, Bengali, Bồ Đào Nha (Brazil), Bồ Đào Nha (Portugal), Catalan, Croatian, Czech, Đức, Farsi, Galician, Georgian, Hà Lan, Hebrew, Hindi, Hy Lạp, Indonesian, Italian, Kannada, Kazakh, Kirghiz, Latvian, Maori, Mông Cổ, Nhật, Phần Lan, Serbian, Thái Lan, Thổ Nhĩ Kỳ, Trung Hoa (phổ thông), Trung Hoa (truyền thống), Ukrainian, Việt Nam.

- Các ấn bản của các ngôn ngữ đang tiến hành phiên dịch bao gồm: Bislama, Blackfoot/Siksika, Bosnian, Esperanto (Quốc tế ngữ), Hungarian, Korean, Macedonian, Marathi, Nepali, Romanian, Shona, Tamil.



Hình 2: Giao diện đa ngôn ngữ của sưu tập Greenstone

Phân cảnh nang sử dụng bao gồm: “*Từ giấy đến bộ sưu tập*”, “*Hướng dẫn sử dụng*”, “*Hướng dẫn cài đặt*” được trình bày dưới các ngôn ngữ: Anh, Pháp, Tây Ban Nha, Nga, Kazakh, và Việt Nam. Các bản hướng dẫn Tiếng Việt có thể được download tại các website ở Anh, Brazil, Đài Loan, Đức, Hoa Kỳ, Ireland, Nga, Nhật, Pháp, Thụy Sĩ, và Úc thông qua một địa chỉ trung gian được tìm thấy tại web site của Greenstone thể giới là <http://prdownloads.sourceforge.net/greenstone/Paper-2.39-vi.pdf>.

Xây dựng bộ sưu tập với Greenstone

Phương thức chung.

Có 3 cách để xây dựng bộ sưu tập với Greenstone:

- Xây dựng thủ công bằng công cụ ORGANIZER;
- Xây dựng tự động bằng công cụ GREENSTONE COLLECTOR;
- Xây dựng bán tự động bằng công cụ LIBRARIAN INTERFACE.

Thường thì ta dùng công cụ LIBRARIAN INTERFACE để tổ chức một tài liệu, công việc này trong nghiệp vụ thông tin - thư viện gọi là biên mục (cataloging) và chỉ mục (indexing), còn trong công nghệ kỹ thuật số thì gọi là xác định metadata. Một tài liệu sau khi được tổ chức biên mục theo chuẩn Dublin Core (thủ công) và xác định metadata (tự động) sẽ trở thành một thư mục chứa năm thư mục con: *Import* (chứa tài liệu thô); *Archives* (chứa tập tin tự tạo dạng GML), *Index* (chứa các tập tin cuối cùng phục vụ người dùng kèm theo metadata), *Building* (thư mục trung gian trong quá trình xây dựng bộ sưu tập), -etc (thư mục hỗ trợ chứa các tập tin điều khiển quá trình kiến tạo bộ sưu tập).

Giao diện LIBRARIAN INTERFACE trình bày 15 yếu tố của Dublin Core cho ta biên mục tài liệu. Quá trình này khiến ta chọn những dẫn mục (entry) hay điểm truy cập (access point) của mỗi tài liệu để phục vụ việc truy tìm và lướt tìm sau này. Cũng bằng giao diện này, Greenstone sẽ cho ta xác định những dẫn mục và hình thức truy tìm hay lướt tìm được trình bày trên giao diện của bộ sưu tập. Chẳng hạn như: Nhan đề (Title), Tác giả (Author), Từ khoá (Keyword), Đề mục (Subject), vv... hoặc Tìm kiếm (Search) thì tìm kiếm trên những điểm truy cập nào. Trong quá trình tổ chức tài liệu, việc thêm, bớt hay thay thế những yếu tố của Dublin Core là rất dễ dàng. Công việc sưu tầm và tổ chức tài liệu cứ tiếp diễn liên tục, tài liệu được lưu vào máy tính cá nhân. Greenstone hoàn toàn xử lý một cách tự động và nhanh chóng một khi chúng ta muốn xuất bản tài liệu như một bộ sưu tập lên Internet hay CD-ROM. Dĩ nhiên chúng ta vẫn có thể cập nhật tài liệu vào bộ sưu tập mỗi khi cần thiết; khi đó thì ta phải xuất bản lại bộ sưu tập.

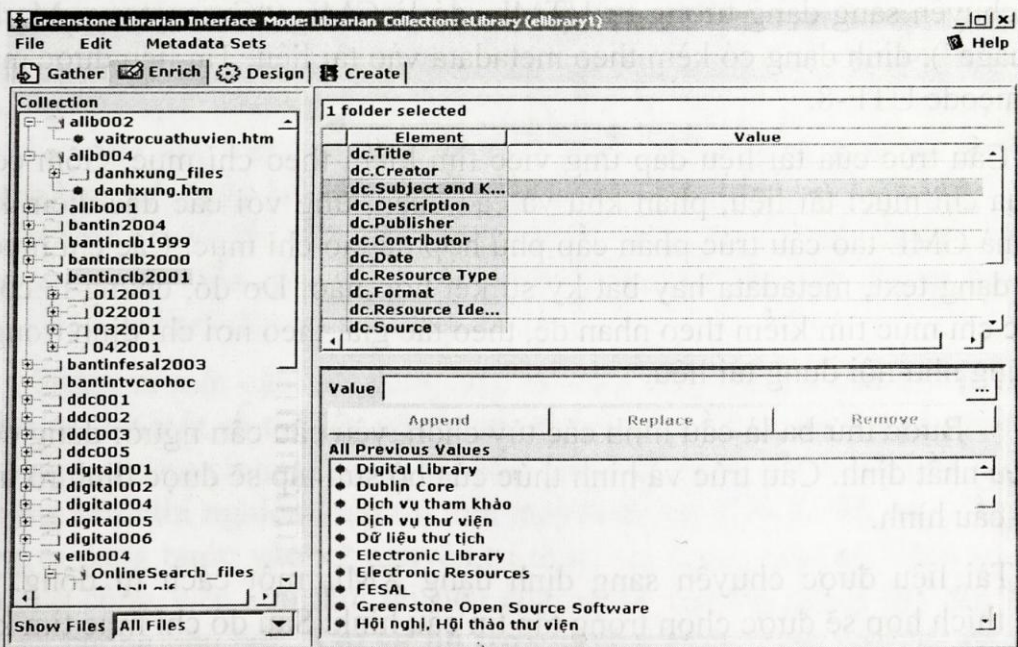
Mỗi bộ sưu tập được xuất bản lên Internet hay CD-ROM đều có một giao diện Greenstone kèm theo. Nếu trên CD-ROM thì giao diện Greenstone có chứa sẵn một phần trình duyệt (web browser) Netscape để tải xuống (download) cho những máy cá nhân nào không sử dụng web. Như thế cho chúng ta thấy rằng mỗi khi một bộ sưu tập được xuất ra CD-ROM thì người sử dụng có thể dùng bất kỳ một máy tính với bất kỳ một hệ điều hành nào đều có thể đọc, truy tìm, lướt tìm, in ra những thông tin trên bộ sưu tập với giao diện thân thiện của Greenstone. Nếu chúng ta sử dụng và truy cập vào các bộ

su tập của các quốc gia khác nhau khắp nơi trên thế giới thì có thể sử dụng giao diện của nhiều ngôn ngữ, trong đó có cả giao diện Tiếng Việt.

Giao diện (INTERFACE)

Giao diện INTERFACE cung cấp bốn giao diện tương tác phản ánh các bước thực hiện như sau:

- GATHER: Dùng để thu gom tài liệu tập trung vào bộ sưu tập;
- ENRICH: Trình bày 15 yếu tố của Dublin Core để biên mục tài liệu. Động tác biên mục được làm thủ công. Người biên mục có thể chọn trên danh sách những yếu tố có sẵn hoặc “append” một yếu tố mới vào; cũng có thể “replace” một yếu tố hay hơn và “remove” một yếu tố khác đi;
- DESIGN: Dùng để thiết kế giao diện bộ sưu tập kèm theo những chỉ thị tìm kiếm theo những dẫn mục cho ta chọn, chẳng hạn như *tác giả, nhan đề, đề mục, năm, nguồn, vv...*;
- CREATE: Dùng để xuất bộ sưu tập lên Internet hay ra đĩa CD. Đây là một thao tác tự động.



Hình 3: Giao diện LIBRARY INTERFACE của Greenstone với 15 yếu tố Dublin Core

Các bước thực hiện.

Để tạo được một bộ sưu tập phải thực hiện nhiều bước, cụ thể là: khai báo thông tin bộ sưu tập, xác định dữ liệu nguồn, cấu hình và xây dựng bộ sưu tập. Trong đó, bước cấu hình đóng vai trò quan trọng nhất vì nó quyết định bộ sưu tập có hoạt động được hay không.

- Bước đầu tiên là khai báo tên của bộ sưu tập và các thông tin liên quan như địa chỉ email, thông tin mô tả sưu tập. Bước này sẽ giúp chúng ta quản lý dễ dàng khi đã xây dựng được nhiều bộ sưu tập.

- Bước thứ hai là xác định nguồn dữ liệu để đưa vào bộ sưu tập.

- Tên thư mục trong cùng hệ thống sẽ bắt đầu bởi "file://"
- Địa chỉ bắt đầu bằng "http://" các tập tin sẽ được lấy về từ Web.
- Địa chỉ bắt đầu bằng "ftp://" các tập tin sẽ được lấy về qua FTP.

Bộ sưu tập sẽ gồm tất cả các tập tin, các thư mục con trong thư mục được chỉ định. Đối với "http://" thì bộ sưu tập sẽ ánh xạ đến web site được chỉ định.

Khi tạo mới một bộ sưu tập hay thêm tài liệu vào sưu tập đã sẵn có, đầu tiên tài liệu nguồn được đưa vào hệ thống – *importing*. Đồng thời tài liệu được chuyển sang dạng tương tự HTML, đó là GML (“Greenstone Markup Language”), định dạng có kèm theo metadata vào tài liệu. Tài liệu được mang mã Unicode UTF-8.

Cấu trúc của tài liệu đáp ứng việc tìm kiếm theo chỉ mục. Gồm có ba cấp của chỉ mục: tài liệu, phân khu và các đoạn, ứng với các đặc điểm khác nhau mà GML tạo cấu trúc phân cấp phù hợp để tạo chỉ mục. Các chỉ mục có thể là dạng text, metadata hay bất kỳ sự kết hợp nào. Do đó, chúng ta có thể tạo các chỉ mục tìm kiếm theo nhan đề, theo tác giả, theo nơi chỉ định trong tài liệu cũng như nội dung tài liệu.

- Bước thứ ba là cấu hình các tùy chọn, yêu cầu cần người dùng có sự am hiểu nhất định. Cấu trúc và hình thức của bộ sưu tập sẽ được định rõ trong tập tin cấu hình.

Tài liệu được chuyển sang định dạng XML một cách tự động. Các plugin thích hợp sẽ được chọn trong tập tin cấu hình. Sau đó chỉ mục tìm kiếm và cấu trúc trình duyệt toàn tài liệu được qui định trong tập tin cấu hình đã tạo. Cuối cùng, kết quả của tiến trình xây dựng tạo ra bộ sưu tập hoạt động được.

Giai đoạn xây dựng có thể mất nhiều thời gian. Những bộ sưu tập nhỏ mất khoảng một vài phút nhưng với những sưu tập lớn thì có thể mất hàng giờ hay nhiều hơn.

Kho Tài nguyên học tập

Kho tài nguyên học tập là hình thức tiêu biểu nhất của Thư viện số trong một trường đại học. Đây là kho tài nguyên điện tử, tùy theo qui mô và chức năng của Trường đại học mà một kho tài nguyên học tập có thể trở nên rất đa dạng. Thường bao gồm:

1. Tài liệu đa phương tiện: Bản đồ, hình ảnh, mẫu vật, đoạn băng hình thí nghiệm, băng hình bài giảng, hội nghị khoa học, khối giáo trình (course block), vv... phục vụ như là công cụ trợ giảng, đào tạo từ xa.
2. Bộ sưu tập chuyên ngành phục vụ nghiên cứu, giảng dạy, học tập.

Trong những đại học đa ngành có quy mô lớn, chẳng hạn như ĐHQG TP. Hồ Chí Minh, có thể tập trung kho tài nguyên học tập từ các trường thành viên để thành lập một **Trung tâm tài nguyên học tập** (Learning Resource Center). Sử dụng công nghệ đa phương tiện để tích hợp tài nguyên học tập và phục vụ trực tuyến việc học tập, giảng dạy, hội nghị, vv...

Trong phạm vi bài này chúng tôi chỉ đề cập đến việc tạo lập những bộ sưu tập chuyên ngành bằng Phần mềm nguồn mở Thư viện số Greenstone để đóng góp cho Kho Tài nguyên học tập. Đây là công việc do đội ngũ giảng viên, nghiên cứu, nói chung là những người sử dụng cùng với đội ngũ nhân viên thư viện xây dựng, theo đúng tinh thần **"Thư viện số là sự tương tác giữa người sử dụng với thư viện để phục vụ chính người sử dụng"**. Công việc này sẽ trở thành một sinh hoạt bình thường trong một trường đại học trong giai đoạn hiện nay; tuy nhiên bước khởi đầu cần phải có sự đầu tư và động viên của lãnh đạo nhà trường.

Phàm là một cán bộ giảng dạy đại học hay nghiên cứu thì ai cũng tự mình sưu tầm một số tài liệu về chuyên ngành của mình, đó là sách tham khảo, photocopy những chương sách, bài báo, mẫu vật, hình ảnh minh họa, bản thiết kế, băng hình thí nghiệm, tập tin trên máy tính hay CD-ROM, vv... Ngày nay chúng ta đang bước vào kỷ nguyên tri thức với công nghệ kỹ thuật số có thể giúp cho công việc sưu tầm tài liệu của người cán bộ giảng dạy đại học dễ dàng và phong phú hơn. **Với sự hỗ trợ của thư viện, từng cán bộ giảng dạy và nghiên cứu sử dụng phần mềm nguồn mở Greenstone để xây dựng bộ**

suu tập cho chính mình rồi đóng góp vào thư viện để làm phong phú Kho tài nguyên học tập.

Xây dựng Kho Tài nguyên học tập bằng phương thức trên có nghĩa là ta suu tâm thông tin và làm cho thông tin đó trở nên phổ biến hơn đối với người khác. Đây là điều liên quan đến vấn đề xã hội, chúng ta cần lưu ý và am hiểu quyền sở hữu trí tuệ để hành động đúng luật pháp.

KẾT LUẬN

Thư viện luôn luôn đóng vai trò quan trọng trong việc hỗ trợ công tác học tập và giảng dạy trong một trường đại học. Nhu cầu này ngày càng cao do đó giá trị nghiệp vụ thư viện cũng thay đổi từ chỗ thư viện là nơi tập trung thông tin chờ người đến sử dụng, người cán bộ thư viện đóng vai trò của người giữ sách thụ động; ngày nay thư viện là nơi đáp ứng tức thì nhu cầu của người sử dụng, người cán bộ thư viện có vai trò chủ động của người cung cấp thông tin và hợp tác với người sử dụng để làm phong phú nguồn tri thức vì lợi ích chính người sử dụng.

Thư viện ĐH Khoa học Tự nhiên đã có một quá trình nghiên cứu và tham gia việc chuyển đổi Phần mềm nguồn mở Thư viện số Greenstone sang Tiếng Việt nên có kinh nghiệm trong việc triển khai và sử dụng. Chúng tôi sẵn sàng tham gia công việc tập huấn và chuyển giao công nghệ Greenstone Tiếng Việt cho đội ngũ giảng viên và nhân viên thư viện trong ĐHQG TP. Hồ Chí Minh để xây dựng Kho tài nguyên học tập với mục tiêu cuối cùng là nâng cao chất lượng đào tạo nguồn nhân lực.

Nâng cao chất lượng đào tạo nguồn nhân lực để phục vụ cho cộng đồng là công việc đầu tư cho tương lai. **"Thư viện có truyền thống là người giữ gìn quá khứ; nhưng ngày nay thư viện ngày mỗi ngày trở thành đường dẫn đến tương lai"**.

(*) thuật ngữ **tài nguyên** – **resource** ngày nay được dùng phổ biến trong ngành Thông tin - Thư viện để chỉ tất cả các loại hình tài liệu, in ấn, điện tử, và đa phương tiện. (Thuật ngữ **tài nguyên** dần dần thay thế cho **tài liệu**)

Trong chuẩn biên mục Dublin Core và Bộ Quy tắc biên mục RDA (Sẽ phát hành vào năm 2009 thay cho AACR2), người ta gọi một đối tượng để biên mục: một cuốn sách, một bài báo, một tấm hình, một tài liệu điện tử, một video clip, một bài hát, vv... là **một resource** – **một tài nguyên**. Chẳng

hạn như thuật ngữ *serials* - **ấn phẩm liên tục** ngày nay được thay thế bằng *continuing resources* – **tài nguyên liên tục** (Chương 12 của AACR2-2002)

Chúng ta không nên nhầm lẫn giữa *information resource* là **tài nguyên thông tin** với *information source* là **nguồn thông tin**. Do đó *open source* là **nguồn mở** còn *open resource* là **tài nguyên mở** tức là tài nguyên miễn phí trên mạng.

TÀI LIỆU THAM KHẢO

1. PANDIAN, B. Maruthu, SONKER, Sharad Kumar và MOORTHY R. ***Creating Digital Libraries: An Experiment with Greenstone Digital Library Open Source Software***
2. TILLET, B. Barbara. ***AACR3: Resource Description and Access*** .- CPSO, Library of Congress, 2004.
3. WITTEN, Ian H., BAIBRIDGE, David và BODDIE, Stefan J.. ***Greenstone Open-Source Digital Library Software***, D-Lib Magazine, October 2001, Volume 7, Number 10 (www.dlib.org)
4. WITTEN, Ian H. và BAIBRIDGE, David. ***How to Build a Digital Library***. – New York : Morgan Kaufmann, 2003.