

1. **Thư viện truyền thống** là thư viện lưu trữ và phục vụ tài liệu in ấn, có thể sử dụng máy tính và công nghệ thông tin (CNTT) để tự động hóa tất cả các hoạt động của thư viện.

2. **Thư viện điện tử** là thư viện truyền thống có tổ chức phục vụ thông tin điện tử. Thông tin điện tử là thông tin được chuyển tải và đọc trên máy vi tính hay mạng máy tính. Tài liệu điện tử bao gồm CD-ROM và cơ sở dữ liệu (CSDL) CD-ROM, CSDL trực tuyến thương mại, sách và tạp chí điện tử, tạp tin máy tính, vv... Những tài nguyên điện tử này thường là do các thư viện mua quyền sử dụng.

3. **Thư viện số** là thư viện điện

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 mà chúng ta cần phân biệt. 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 hัก những đặc điểm quan trọng của việc sưu

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.

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. Một bộ sưu tập có thể chứa nhiều loại tài liệu khác nhau. Ví dụ một bộ sưu tập về đề tài "Phố cổ Hội An" sẽ bao gồm những tài liệu dạng văn bản về lịch sử, văn hoá, phong tục, vv...; tài liệu dạng hình ảnh về những di tích, trang phục, các bản thiết kế, vv...; tài

Sử dụng phần mềm nguồn mở GREENSTONE để xây dựng thư viện số - cơ hội cho tất cả các thư viện Việt Nam

ThS. NGUYỄN MINH HIỆP

Giám đốc Thư viện Trường Đại học Khoa học Tự nhiên
Đại học Quốc gia TP.HCM

Hiện nay ở nước ta, tồn tại ba loại hình thư viện: Đa số là thư viện truyền thống, một số đang phát triển thư viện điện tử, tất cả đang đứng trước ngưỡng cửa của thư viện số - thế nhưng cơ hội phát triển loại hình thư viện này là dành cho tất cả mọi người.

tử có tạo lập nên tài nguyên điện tử bằng cách:

- Số hoá tài liệu in ấn;
- Tạo lập những bộ sưu tập thông tin điện tử.

TAI SAO XÂY DỰNG THƯ VIỆN SỐ?

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

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. 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. Đó là lý do chúng ta cần phải xây dựng Thư viện số.

XÂY DỰNG THƯ VIỆN SỐ NHƯ THẾ NÀO?

Theo Ian H. Witten, chuyên gia Thư viện số Đại học Waikato, New Zealand, Thư viện số là tập hợp những bộ sưu tập thông tin

liệu dạng âm thanh về những bài hát, dân ca, vv...; tài liệu dạng phim về những lễ hội, sinh hoạt cộng đồng, vv... Một bộ sưu tập chứa nhiều tài liệu với dạng thức khác nhau, tuy nhiên cung cấp một giao diện đồng nhất qua đó tất cả các tài liệu có thể được truy cập, mặc dù cách mà tài liệu đó hiển thị sẽ tuỳ thuộc vào phương tiện và dạng thức của tài liệu đó. Một bộ sưu tập như thế trước khi trình bày phải qua một quá trình hình thành để tạo nên những cấu trúc hỗ trợ cho việc truy tìm và lướt tìm được dùng cho việc truy cập sưu tập, một khi xây dựng xong, bộ sưu tập có thể được xuất bản trên Internet hoặc xuất ra CD-ROM một cách hoàn toàn tự động. Một khi sưu tầm thêm tài liệu mới, ta có thể dễ dàng bổ sung thêm vào bộ sưu tập bằng cách tái xây dựng. Một thư

► viện nói chung bao gồm nhiều bộ sưu tập khác nhau, mỗi sưu tập tổ chức mỗi khác, tuy nhiên hoàn toàn giống nhau về phương cách hiển thị. Những bộ sưu tập như thế có thể được tạo nên bằng một phần mềm nguồn mở Thư viện số Greenstone.

PHẦN MỀM NGUỒN MỞ THƯ VIỆN SỐ GREENSTONE

Phần mềm nguồn mở khác với Phần mềm thương mại là phần mềm đóng được ví như hòn đá có thể mòn đi nhưng hòn đá vẫn là hòn đá; trong khi Phần mềm nguồn mở được ví như một thực thể sinh học, có thể phát triển để lớn lên. Phần mềm nguồn mở là miễn phí, người sử dụng tự do sử dụng theo yêu cầu của mình, có quyền chỉnh sửa phần mềm cho phù hợp với những ứng dụng của mình cũng như làm lợi cho cộng đồng, và có quyền phân phối lại bản chỉnh sửa của mình có hay không có thu phí. Hiện nay TP. Hồ Chí Minh đang cổ suý việc sử dụng Phần mềm nguồn mở trong tất cả những lĩnh vực ứng dụng CNTT.

Greenstone là một bộ phần mềm dùng để xây dựng và phân phối các bộ sưu tập thư viện số. Nó cung cấp một phương pháp mới để tổ chức và xuất bản thông tin trên Internet hoặc trên CD-ROM; ngoài ra còn cung cấp phương tiện dễ dàng cho người sử dụng truy tìm toàn văn và lướt tìm dựa vào metadata. Greenstone là sản phẩm của dự án New Zealand Digital Library của đại học Waikato, được phát triển và phân phối với sự tham gia của UNESCO và Human Info NGO. Đây là phần mềm mã nguồn mở được cung cấp trên theo thoả thuận của GNU General Public License.

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 các thư viện 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 20 ngôn ngữ bao gồm: Anh, Pháp, Đức, Nga, Tây Ban Nha, Bồ Đào Nha, Italia, Hà Lan (Dutch), A-rập, Thổ Nhĩ Kỳ, Hebrew, Bosnia, Armenia, Ukraina, Farsi, Croatia, Serbia, Hindi, Trung Hoa, Nhật Bản, Thái Lan, Indonesia, vv... Phiên bản Greenstone Tiếng Việt cũng đã được phát hành và phổ biến trong hệ thống Greenstone trên thế giới.

Greenstone có những đặc điểm sau:

- Truy cập rộng rãi. Bộ sưu tập được truy cập qua một trình duyệt web chuẩn.
- Đa hệ. Bộ sưu tập có thể chạy trên Window và Unix.
- Hướng metadata. Chỉ mục lướt tìm được tạo nên từ metadata. Metadata có thể kết hợp với mỗi tài liệu hoặc một phần của tài liệu.
- Đa ngôn ngữ. Unicode được dùng để hỗ trợ việc chuyển đổi ngôn ngữ. Chỉ mục riêng biệt có

thể tạo ra cho những ngôn ngữ khác nhau.

- Thang độ (scale) lớn. Những bộ sưu tập chứa hàng triệu tài liệu, nhiều gigabytes có thể được tạo ra. Truy tìm toàn văn nhanh chóng. Có thể nén để giảm kích thước văn bản và chỉ mục.

- Tương thích Z39.50. Giao thức Z39.50 hỗ trợ việc truy cập máy chủ bên ngoài cũng như giới thiệu bộ sưu tập Greenstone cho người sử dụng bên ngoài.

- Truy tìm linh hoạt. Người sử dụng có thể truy tìm toàn văn tài liệu với những dẫn mục thích hợp.

- Lướt tìm linh hoạt. Người sử dụng có thể lướt tìm danh mục tác giả, danh mục nhan đề, danh mục đề mục, danh mục từ khoá, những cấu trúc phân cấp khác, vv...

- Đa phương tiện. Bộ sưu tập có thể chứa hình ảnh, âm nhạc, đoạn băng ghi âm và hình.

- Xuất ra CD-ROM. Bộ sưu tập có thể xuất ra một CD-ROM tự khởi động.

- Greenstone là phần mềm nguồn mở cho nên dễ dàng chỉnh sửa.

XÂY DỰNG BỘ SƯU TẬP VỚI GREENSTONE

Để xây dựng một bộ sưu tập mới thường ta phải đổi mới với cả hai loại tài liệu: tài liệu đã ở dạng điện tử rồi và tài liệu in ấn cần phải số hóa. Nếu tài liệu ở dạng điện tử thì công việc hết sức dễ dàng, ngay cả việc sưu tầm, tổ chức tập tin và chuyển đổi dạng thức. Còn đối với tài liệu in ấn thì ta phải tiến hành số hóa bằng cách quét tài liệu và chuyển qua dạng PDF; khi sử dụng Greenstone để lưu trữ tài liệu đó trong CSDL thì tự động mang hai dạng HTML và PDF.

Chúng ta dùng công cụ LIBRARIAN INTERFACE của Phần mềm Greenstone 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 ít nhất là 2 biểu ghi: một biểu ghi nội dung dạng HTML và biểu ghi metadata. Nếu nội dung mang nhiều dạng thức khác nhau thì thư mục tài liệu có nhiều biểu ghi hơn.

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

(Xem tiếp trang 22)

Sử dụng phần mềm... (Tiếp theo trang 8)

nhanh. 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 sưu 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.

Ứng dụng của Bộ sưu tập

1. Dịch vụ tham khảo: Đáp ứng yêu cầu thông tin của người sử dụng bằng cách sưu tầm thông tin theo chủ đề hay đề tài cụ thể rồi xuất ra CD-ROM cho người sử dụng;

2. Xây dựng Kho tài nguyên học tập: Với sự hỗ trợ của thư viện, từng cán bộ giảng dạy và nghiên cứu trong Trường đại học sử dụng phần mềm nguồn mở Greenstone để xây dựng bộ sưu tập cho chính mình rồi đóng góp vào thư viện để làm phong phú Kho tài nguyên học tập.

3. Công tác địa chí: Tất cả tài liệu địa chí từ văn bản, di chỉ, mẫu vật, hình ảnh di tích lịch sử, đến phim ảnh đều có thể số hóa như những tài liệu và tổ chức thành Bộ sưu tập lưu giữ trên Internet hay CD-ROM.

Một điều cần lưu ý là xây dựng những Bộ sưu tập bằng phương

thức trên có nghĩa là ta sưu tầm 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 quan tâm đến và am hiểu quyền sở hữu trí tuệ để hành động đúng luật pháp.

KẾT LUẬN

Xây dựng Thư viện số bằng cách sử dụng Phần mềm nguồn mở để tạo nên những Bộ sưu tập là xu thế hiện nay trên thế giới. Phần mềm nguồn mở Thư viện số Greenstone được UNESCO khuyến khích sử dụng đặc biệt trong các quốc gia đang phát triển là cơ hội để hệ thống thư viện Việt Nam nắm bắt công nghệ mới để phát triển. Greenstone dễ sử dụng và rất tiện ích. Những thư viện đã có một phần mềm quản lý thư viện rồi thì tích hợp Greenstone để xây dựng những bộ sưu tập; những thư viện chưa có phần mềm quản lý thư viện vẫn có thể dùng Greenstone để tạo lập sưu tập, đồng thời có thể phát triển từ Greenstone những ứng dụng khác đáp ứng yêu cầu của thư viện mình nhằm đi đến một hệ thống hoàn chỉnh.

Thư viện Đại học Khoa học Tự nhiên TP. Hồ Chí Minh 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 hỗ trợ đồng nghiệp nắm bắt công nghệ mới này để cùng nhau xây dựng Thư viện số.

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. Thư viện số ĐH Khoa học Tự nhiên
<http://www.glib.hcmuns.edu.vn/greenstone>

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. □