

Giải pháp số hóa nguồn tài liệu và xây dựng các bộ sưu tập số trong thư viện các trường đại học

Xây dựng thư viện số là xu thế phát triển tất yếu của ngành thư viện nói chung và của thư viện Đại học nói riêng. Một xu thế phát triển không phải vì mong muốn chủ quan của những người làm thư viện mà vì đòi hỏi khách quan của sự phát triển Khoa học kỹ thuật cũng như đòi hỏi của xã hội phát triển.

Xây dựng thư viện số là xu thế phát triển tất yếu của ngành thư viện nói chung và của thư viện Đại học nói riêng. Một xu thế phát triển không phải vì mong muốn chủ quan của những người làm thư viện mà vì đòi hỏi khách quan của sự phát triển Khoa học kỹ thuật cũng như đòi hỏi của xã hội phát triển.

1. KHÁI NIỆM

Thư viện điện tử và Thư viện số là những khái niệm đang còn rất mới ở Việt nam và cũng tồn tại nhiều cách hiểu và định nghĩa khác nhau:

Có thể hiểu theo nghĩa tổng quát là một loại hình thư viện đã tin học hóa toàn bộ hoặc một số dịch vụ thư viện. Là nơi người sử dụng có thể tới để tra cứu, sử dụng các dịch vụ thường làm như với một thư viện truyền thống nhưng đã được tin học hóa. Nguồn lực của Thư viện điện tử bao gồm cả tài liệu in giấy và tài liệu đã được số hóa.

Thư viện số (TVS) là một Thư viện điện tử cao cấp trong đó toàn bộ các tài liệu chủ yếu của thư viện đã được số hóa và được quản lý bằng một phần mềm chuyên nghiệp có tổ chức giúp người dùng dễ dàng truy cập, tìm kiếm và xem được nội dung toàn văn của chúng từ xa thông qua hệ thống mạng thông tin và các phương tiện truyền thông.

Một Thư viện số hoàn chỉnh phải thực hiện được tất cả các dịch vụ cơ bản của thư viện truyền thống kết hợp với việc ứng dụng các lợi thế của công nghệ thông tin trong việc lưu trữ, tìm kiếm và phổ biến nội dung thông tin.

Khái niệm Thư viện số không chỉ tương đương với bộ sưu tập số, đó là một môi trường tập hợp các bộ sưu tập số theo chủ đề. Nguồn thông tin của thư viện số có thể nằm ngay trong thư viện và có thể cả bên ngoài thư viện (ví dụ: CSDL toàn văn mua quyền truy cập theo thời gian).

Khái niệm về bộ Sưu tập số: Bộ sưu tập số là một tập hợp có tổ chức nhiều tài liệu đã được số hoá dưới nhiều hình thức khác nhau (văn bản, hình ảnh, Audio, Video...) về một chủ đề. Mặc dù mỗi loại hình tài liệu có sự khác nhau về cách thể hiện, nhưng nó đều cung cấp một giao diện đồng nhất mà qua đó các tài liệu có thể truy cập, tìm kiếm và sử dụng dễ dàng.

Ví dụ: bộ sưu tập số về chủ tịch Hồ Chí Minh gồm các văn bản, các tác phẩm văn học, các văn kiện chính trị do Bác viết và do người khác viết về Bác; Những bài hát, bản nhạc viết về Hồ chí Minh; những đoạn phim, những băng video phản ánh cuộc đời và sự nghiệp hoạt động cách mạng của Người.

Như vậy, một Thư viện số có thể bao gồm nhiều bộ sưu tập số theo các chủ đề khác nhau, có thể do tập thể hoặc cá nhân tự xây dựng hoặc trao đổi, mua bán. Có thể nằm trong lưu trữ của thư viện nhưng cũng có thể nằm ngoài thư viện thông qua một kênh cung cấp từ phía đối tác.

Trong điều kiện về kinh phí, nhân lực và cơ sở hạ tầng nói chung của các thư viện đại học hiện nay thì việc đặt ra mục tiêu trước mắt để xây dựng một Thư viện số là chưa có tính khả thi. Nhưng với mục tiêu xây dựng các bộ sưu tập số thì các thư viện hoàn toàn có thể thực hiện được.

2. Ý NGHĨA CỦA BỘ SƯU TẬP SỐ

Một trong những nhiệm vụ ưu tiên hàng đầu trong kế hoạch chiến lược phát triển của một trường đại học, đó là việc tăng cường mở rộng ứng dụng công nghệ thông tin phục vụ thiết thực và hiệu quả cho công tác quản lý, công tác đào tạo, nghiên cứu hướng tới mục tiêu trở thành trường đại học tiên tiến đạt chuẩn trong khu vực và trên thế giới.

Chất lượng đào tạo và nghiên cứu khoa học của nhà Trường là tiêu chí quan trọng để đánh giá mức độ hoàn thành các mục tiêu mà nhà Trường đề ra trong Kế hoạch chiến lược phát triển. Để nâng cao chất lượng đào tạo và nghiên cứu khoa học; Trước hết phải tập trung nâng cao năng lực nghiên cứu và giảng dạy của đội ngũ giảng viên, tạo cho giảng viên có nhiều thời gian hơn dành cho nghiên cứu khoa học, nắm bắt những vấn đề mới, sát với thực tiễn khoa học công nghệ cũng như kinh tế, xã hội đất nước; Tăng cường và khuyến khích giảng viên trẻ nâng cao trình độ; Hệ thống giáo trình, bài giảng phải từng bước được cập nhật và biên soạn mới để đảm bảo nội dung chất lượng và phải được thiết kế trên cơ sở

áp dụng được những công nghệ hiện đại trong đào tạo: Hệ thống công nghệ thông tin và dịch vụ thông tin – thư viện nhà trường.

Việt Nam gia nhập WTO vào ngày 07/11/2006 – Chúng ta bắt đầu hội nhập toàn diện với quốc tế, thực hiện hàng loạt các cam kết cắt giảm hàng rào thuế quan và phi thuế quan, giảm bảo trợ trong nước, mở cửa thị trường.... Trong Giáo dục và đào tạo đến năm 2009 Nhà nước cho các cơ sở đào tạo nước ngoài đầu tư 100% vốn vào Việt Nam sẽ làm tăng tính cạnh tranh đối với các cơ sở giáo dục đại học của Việt Nam. Do đó yêu cầu các cơ sở đào tạo cần phải đổi mới phương thức đào tạo, nâng cao chất lượng, từng bước hội nhập với quốc tế, tạo sự cạnh tranh với các đối tác nước ngoài thì mới tồn tại và phát triển.

Xuất phát từ những yêu cầu nêu trên, công tác Thông tin – Thư viện ở các trường đại học phải có sự đổi mới mạnh mẽ và phải đi trước một bước mới đáp ứng được yêu cầu phục vụ thông tin – tri thức cho nâng cao chất lượng đào tạo của nhà trường. Giải pháp xây dựng các Bộ sưu tập số trong các thư viện đại học là một bước đi cần thiết để góp phần giải quyết các vấn đề về đổi mới và nâng cao năng lực, chất lượng đào tạo của nhà trường. Bởi lẽ bộ sưu tập số có những đặc tính nổi trội mà dịch vụ thư viện truyền thống chưa có như:

- Bộ sưu tập số tạo ra một môi trường và cơ hội bình đẳng rộng mở cho tất cả mọi người đều có cơ hội sử dụng nguồn tài liệu học tập bởi nó không bị giới hạn về không gian và thời gian. Loại bỏ khoảng cách tri thức giữa người giàu và người nghèo, giữa thành thị và nông thôn, giữa các quốc gia.

- Tính linh hoạt và khả năng đáp ứng của tài liệu số trong đào tạo thể hiện ở chỗ một bản tài liệu số có thể cùng lúc phục vụ cho nhiều đối tượng khác nhau, không phụ thuộc vào số lượng người dùng, thời gian và vị trí địa lý của người học.

- Tính hiệu quả của bộ sưu tập số là tiết kiệm thời gian và kinh phí: thư viện đỡ tốn kinh phí xây dựng kho tàng, kinh phí bổ sung tài liệu, bảo quản và kinh phí trả lương cho người phục vụ. Hơn hết là giúp cho người dùng tin được dễ dàng thuận tiện, tiết kiệm được thời gian, tiền bạc trong việc tìm thông tin.

- Bộ sưu tập số kết hợp với phương thức thư viện truyền thống sẽ phục vụ có hiệu quả hơn cho việc đổi mới và nâng cao chất lượng đào tạo, đặc biệt là đào tạo theo tín chỉ, đào tạo trực tuyến của nhà trường. Giúp cho người học chủ động trong việc sắp xếp thời gian học tập, họ không phải đến thư viện cũng có thể lấy được tài liệu qua hệ thống mạng thông tin ở mọi lúc, mọi nơi.

- Trong điều kiện còn thiếu nguồn tài liệu tham khảo học tập in giấy, thì việc có thêm giải pháp tài liệu số hóa sẽ giúp cho người học có thêm nhiều lựa chọn để phục vụ cho kế hoạch học tập của cá nhân.

- Bộ sưu tập số góp phần giải phóng kiến thức, mở rộng đối tượng phục vụ: Phạm vi phục vụ các tài liệu của thư viện không bị bó hẹp trong khuôn viên của nhà trường mà nó vươn tới các vị trí địa lý khác.

- Bộ sưu tập số là lựa chọn tối ưu để bảo tồn được lâu dài các tài liệu quý hiếm, ngăn chặn những rủi ro hủy hoại do thời gian, thiên tai, khí hậu và tàn suất sử dụng.

3. GIẢI PHÁP XÂY DỰNG BỘ SƯU TẬP TÀI LIỆU SỐ.

Quy trình để xây dựng một bộ sưu tập số bao gồm: 2 bước

1. Số hoá nguồn tài liệu:

- Lựa chọn tài liệu đầu vào;
- Số hóa nguồn tài liệu;
- Tạo siêu dữ liệu liên kết;

2. Lựa chọn công nghệ thực hiện

- Lựa chọn phần mềm quản lý và xây dựng cơ sở dữ liệu tài liệu số
 - Biên mục tài liệu
 - Vận hành, bảo quản và cung cấp dữ liệu;

- Xuất, nhập dữ liệu từ bên ngoài để trao đổi.

3.1.1 Lựa chọn tài liệu đầu vào:

Chúng ta không thể số hoá tất cả các tài liệu có trong thư viện, do đó trước khi thực hiện số hóa nguồn tài liệu, chúng ta phải đưa ra các tiêu chí để làm căn cứ lựa chọn những tài liệu nào cần thiết đưa vào bộ sưu tập. Có những tiêu chí mà chúng ta phải quan tâm như sau:

- Tiêu chí tình trạng bản quyền của tài liệu: Vấn đề bản quyền thực sự đang là rào cản làm nhụt chí những người có tâm huyết với công việc số hoá tài liệu thư viện. Sở dĩ như vậy là vì chúng ta chưa thực sự nắm vững những quy định cụ thể của Luật bản quyền (Luật sở hữu trí tuệ). Các cơ quan chức năng nhà nước cũng chưa có những văn bản dưới luật để hướng dẫn thi hành.

Theo Luật sở hữu trí tuệ của Việt nam ban hành năm 2005 thì tại điều 25 khoản (a) và (đ) có quy định: những trường hợp sử dụng tác phẩm đã công bố không phải xin phép, không phải trả tiền nhuận bút, thù lao:cụ thể như sau:

a) TỰ sao chép một bản nhằm mục đích nghiên cứu khoa học, giảng dạy cá nhân;

đ) Sao chép tác phẩm để lưu trữ trong thư viện với mục đích nghiên cứu;

Một khái niệm về phạm vi không gian khuôn viên thư viện hiện nay cũng cần phải xem xét lại: đối với một thư viện truyền thống thì phạm vi khuôn viên thư viện là một không gian nằm trong hàng rào của trụ sở thư viện. Nhưng với một thư viện số thì khuôn viên thư viện không phải như vậy mà là phạm vi những bạn đọc được thư viện cho phép truy cập sử dụng tài liệu dưới sự giám sát bằng một công cụ tin học cho dù người đó sử dụng từ bên ngoài thư viện. Với khái niệm này thì việc vận dụng Luật sở hữu trí tuệ vào số hoá tài liệu thư viện sẽ dễ dàng hơn.

- Theo tiêu chí nhóm người dùng mà thư viện xác định mức độ ưu tiên phục vụ:

- + Cán bộ lãnh đạo;

- + Giảng viên, cán bộ nghiên cứu;

- + Cán bộ kỹ thuật;

- + Sinh viên;

- + Các đối tượng khác.

- Theo tiêu chí nội dung tài liệu: Trên cơ sở xác định nhóm người dùng mục tiêu mà thư viện lựa chọn các chủ đề tài liệu theo nội dung tài liệu phục vụ; tài liệu phục vụ phát triển giáo trình, bài giảng; tài liệu có tần suất sử dụng cao

- Theo tiêu chí điều kiện bảo quản hiện tại: Tùy tình hình cụ thể của từng thư viện trong tình trạng điều kiện bảo quản kết hợp với nội dung tài liệu mà quyết định lựa chọn tài liệu đưa vào. Ví dụ: ưu tiên cho các tài liệu in trên giấy tròn, dễ rách, có hóa chất bảo quản.

- Theo tiêu chí các loại tài liệu đặc biệt: Tài liệu độc bản, tài liệu quý hiếm, thời gian xuất bản (Luận án tiến sỹ, tài liệu cổ, tài liệu cẩm nang chuyên ngành)...

3.1.2. Số hoá nguồn tài liệu

Đây là công đoạn đòi hỏi đầu tư nhiều công sức, kinh phí nhưng lại là khâu dễ dàng thực hiện nhất. Bởi vì hiện nay công nghệ số hóa tài liệu đã tiến bộ rất nhiều. Nếu như trước đây, khi ta muốn số hóa một cuốn sách khoảng 2000 trang thì phải mất hàng mấy ngày để quét từng trang sách. Nhưng hiện nay cũng với cuốn sách đó chỉ mất vài giờ đồng hồ là cho ra một sản phẩm tài liệu số đảm bảo chất lượng tốt, sắc nét, hình ảnh đẹp, giống 100% bản gốc và đặc biệt còn cho phép tự động tạo các siêu dữ liệu mô tả và siêu dữ liệu cấu trúc của tài liệu ở định dạng XML. Hiện nay ở Việt nam đã có các thiết bị số hóa tài liệu của công nghệ KIRTAS APT 1200, công nghệ này cùng với thiết bị BookScan APT 1200 có thể giúp các thư viện có thể số hóa nguồn tài liệu với số lượng lớn, giá cả hợp lý và đảm bảo chất lượng, thiết bị nhận dạng quang học OCR. Đặc biệt là công nghệ KIRTAS APT 1200 có một phần mềm biên tập BookScan Editor cho phép tự động biên tập, tạo siêu dữ liệu theo yêu cầu; BookScan APT 1200 không làm hư hỏng tài liệu gốc do không phải tháo gáy tài

liệu đối với tài liệu có độ dày trang khi thực hiện Scan.

3.2 Lựa chọn công nghệ

Việc lựa chọn công nghệ để tiến hành thực hiện rất quan trọng bởi vì nó là công cụ đắc lực giúp ta thực hiện các công việc trong quy trình tạo lập và vận hành của bộ sưu tập số. Do đó công nghệ để thực hiện phải đáp ứng các yêu cầu:

- Là công cụ, môi trường để đảm bảo các tài liệu số hóa sau khi được tạo lập sẽ dễ dàng, thuận tiện cho người dùng tiếp cận;
- Có đủ độ tin cậy cho người quản trị và kỹ thuật viên trong quá trình tạo lập, bảo quản và cung cấp dữ liệu trong quá trình hoạt động của bộ sưu tập;
- Đáp ứng đầy đủ các yêu cầu về chuẩn nghiệp vụ thông tin – thư viện;
- Dễ dàng trao đổi dữ liệu với các chuẩn khác, có công cụ sao lưu an toàn dữ liệu.

Dựa theo các yêu cầu nêu trên, để bộ sưu tập số phát huy được hết tác dụng, thư viện khi thực hiện tạo lập bộ sưu tập số cần phải có cơ sở hạ tầng sau:

+ Phải có hệ thống mạng Intranet được kết nối Internet với đường truyền đủ đáp ứng cho số người dùng tối thiểu của thư viện (hiện nay hầu hết các trường đều đã xây dựng. Mạng LAN của thư viện là một nhánh của hệ thống Intranet của trường).

+ Hệ thống máy chủ đủ mạnh để đáp ứng việc lưu trữ, bảo quản, cung cấp dữ liệu và quản lý người dùng và các phần mềm hệ thống có bản quyền.

+ Trang Web đăng tải và là cổng truy cập của người dùng vào bộ sưu tập.

+ Phần mềm quản lý tài liệu số:

Hiện nay cũng có nhiều phần mềm, trong đó có phần mềm nguồn mở Greenstone (Hòn đá xanh); Dspace và một số phần mềm do các công ty và cá nhân xây dựng. Công ty Tinh Vân đã phát triển thêm một phần hệ quản lý tài liệu số trong hệ quản trị thư viện điện tử LIBOL phiên bản 6.0. Công ty CMC cũng phát triển một phần hệ Dilib trong hệ quản trị thư viện điện tử Ilib...

Phần mềm quản lý tài liệu số phải đáp ứng các yêu cầu như:

+ Tạo siêu dữ liệu: có 3 dạng siêu dữ liệu:

- Siêu dữ liệu mô tả: Mô tả các thông tin về tài liệu;
- Siêu dữ liệu cấu trúc: Mô tả các liên kết giữa các đối tượng thông tin liên quan của tài liệu như mục lục, chương, phần, trang sách, hình ảnh minh họa, phụ lục... giúp người dùng dễ dàng di chuyển đến các thành phần của tài liệu.
- Siêu dữ liệu quản trị: gồm tạo kích cỡ tập tin; Định dạng tài liệu (PDF); Đặc tính sử dụng và tình trạng của tài liệu.

+ Cho phép upload tài liệu từ máy trạm người dùng theo lô hoặc theo từng tài liệu riêng lẻ.

+ Mô tả dữ liệu: (theo một trong các chuẩn siêu dữ liệu: MARC; Dublin core; MODS; METS, ISO 2709) trong đó chuẩn Dublin Core là dùng tương đối phổ biến vì có khả năng tùy biến cho các tiêu chuẩn khác với 16 trường biên mục.

+ Vận hành liên kết là tạo ra một giao diện tra cứu tích hợp cả nhất quán cho người dùng trên nhiều bộ sưu tập cùng một lúc dựa trên các điểm truy cập nhất quán như: tác giả; Nhan đề tài liệu, từ khóa; chủ đề; chỉ mục quốc gia...

+ Quản lý các nguồn dữ liệu truy cập được cho phép (phần này chưa có trong các phần mềm nguồn mở), theo đó phải có công cụ quản lý để chỉ có các thành viên đã được đăng ký mới được quyền truy cập vào tài liệu.

+ Xuất - nhập dữ liệu để trao đổi với các hệ thống khác theo các chuẩn chung (ISO 2709 hoặc theo MARC). Có công cụ Backup cơ sở dữ liệu định kỳ để bảo quản dữ liệu.

Kết luận: Sự phát triển mạnh mẽ của khoa học và công nghệ cùng với xu thế hội nhập là những đòi hỏi và thách thức cho ngành Thông tin – Thư viện nói chung và cho thư viện đại học nói riêng cần phải có những đổi mới hoạt động, bắt kịp những tiến bộ của thời đại phục vụ đắc lực cho sự nghiệp CNH, HĐH đất nước. Chúng tôi thiết nghĩ giải pháp xây dựng các Bộ sưu tập tài liệu số tại các thư viện đại học là một việc làm cần thiết, một động thái tích cực để đổi mới phương pháp phục vụ nhằm góp phần nâng cao chất lượng đào tạo và nghiên cứu trong trường đại học.

Trung tâm TT-TV Đại học Nông nghiệp Hà nội trong 5 năm qua đã nỗ lực học hỏi, tìm tòi và ứng dụng. Đã số hoá và đưa vào quản lý phục vụ được hơn 10.000 trang tài liệu nội sinh, 04 Bộ sưu tập với gần 5.000 tài liệu, bao gồm các Luận án tiến sỹ, luận văn thạc sỹ, Giáo trình, các Kết quả nghiên cứu KHKT cùng một số tài liệu tham khảo. Bạn đọc có thể truy cập để sử dụng tài liệu của chúng tôi theo địa chỉ : <http://infolib.hua.edu.vn>

Để làm được điều này, ngoài những nỗ lực của cán bộ thư viện tại các thư viện đại học, cũng cần phải có sự quan tâm chỉ đạo và có định hướng từ các cấp lãnh đạo của nhà trường, của Bộ GD & ĐT, Bộ Thông tin truyền thông và các ngành liên quan.

Một tương lai với nhiều cơ hội đang chờ đón chúng ta.

PGS.TS. Hoàng Đức Liên, TVVC. Nguyễn Hữu Ty

Trung tâm Thông tin – Thư viện ĐH Nông nghiệp HN