

QUẢN LÝ TÀI LIỆU ĐIỆN TỬ BẰNG DSPACE TẠI TRUNG TÂM THÔNG TIN-THƯ VIỆN, ĐẠI HỌC KINH TẾ-TÀI CHÍNH THÀNH PHỐ HỒ CHÍ MINH

PGS TSKH Bùi Loan Thùy, CN Trương Thị Ngọc Mai
Trường Đại học KHXH&NV Tp. Hồ Chí Minh

Trình bày các luận cứ khi chọn phần mềm mã nguồn mở Dspace để quản lý tài liệu điện tử của Trung tâm Thông tin-Thư viện, Trường Đại học Kinh tế-Tài chính Tp Hồ Chí Minh. Nêu rõ sáu bài học kinh nghiệm từ thực tế sử dụng Dspace vào công việc nói trên. Liệt kê những điểm yếu hiện chưa khắc phục được khi sử dụng Dspace.

1. Cơ sở lựa chọn phần mềm mã nguồn mở Dspace

Phần mềm mã nguồn mở (PMMNM) được cung cấp dưới dạng mã và nguồn, miễn phí về bản quyền. Người dùng được phép tự sửa đổi, cải tiến, phát triển, nâng cấp theo một số nguyên tắc chung quy định trong giấy phép PMMNM. Người dùng chỉ cần tải xuống phần mềm từ Internet cùng với mã nguồn, chỉnh sửa cho phù hợp với yêu cầu của mình và sử dụng. Bởi vậy, PMMNM đã phát triển rất nhanh, ngày càng phong phú, đa dạng và được nhiều cơ quan, tổ chức, cá nhân, trong đó có thư viện, đặc biệt quan tâm ứng dụng, phát triển. Ví dụ: phần mềm OpenOffice, MediaPortal, VLC media player, GIMP, Dspace, Greenstone,...

Trong số PMMNM, Greenstone và Dspace, hỗ trợ việc xây dựng và phân phối các bộ sưu tập số trên Internet, được khá nhiều thư viện ở Việt Nam sử dụng trong quá trình xây dựng thư viện điện tử.

Đầu năm 2008, Trung tâm thông tin-thư viện (TTTT-TV), Trường Đại học Kinh tế-Tài chính Tp HCM (UEF) đã chọn và ứng dụng Dspace để xây dựng, quản lý vốn tài liệu điện tử của mình dựa trên những luận cứ sau:

- Dspace do HP và THE MIT

Libraries phát triển vào năm 2002;

- Đây là PMMNM nên chỉ cần tải xuống, sửa đổi theo ý muốn rồi sử dụng. Nhờ được miễn phí bản quyền phần mềm, trong điều kiện nhà trường chưa cho phép trang bị phần mềm thương mại, nên đã giảm được chi phí phát triển phần mềm theo yêu cầu nghiệp vụ;

- Tận dụng triệt để khả năng của đội ngũ chuyên viên công nghệ thông tin thuộc Trung tâm Phát triển phần mềm của nhà trường để nghiên cứu phát triển, mở rộng, chỉnh sửa Dspace (sử dụng module có sẵn) cho phù hợp với yêu cầu nghiệp vụ và phục vụ của thư viện;

- Kéo dài thời gian sử dụng/tái sử dụng các phần cứng, thiết bị trong khi vẫn đảm bảo hiệu năng toàn hệ thống;

- Giảm sự phụ thuộc vào nhà cung cấp phần mềm thương mại;

- Có nhiều tiện ích cho hoạt động thư viện, cho phép xây dựng và quản lý các bộ sưu tập số, đáp ứng tốt hơn nhu cầu của giảng viên và sinh viên;

- PMMNM này tổ chức và quản lý file theo hình cây thư mục, thuận tiện và thân thiện với người sử dụng với các tính năng upload, download, tìm kiếm, xem toàn văn, thống kê tài liệu. Giao diện khá đơn

Nghiên cứu - Trao đổi

giản, người sử dụng dễ dàng tiếp cận với các bộ sưu tập, tài liệu một cách khoa học;

- Phần mềm có khả năng phân quyền vào từng tài liệu chứa trong từng bộ sưu tập. Tốc độ truy cập nhanh, đáp ứng tốt nhu cầu đọc tài liệu của người sử dụng;

- Dspace sử dụng metadata gồm 15 trường Dublin Core nên biên mục đơn giản, chỉ đưa những thông tin cơ bản, giúp giảm thời gian biên mục.

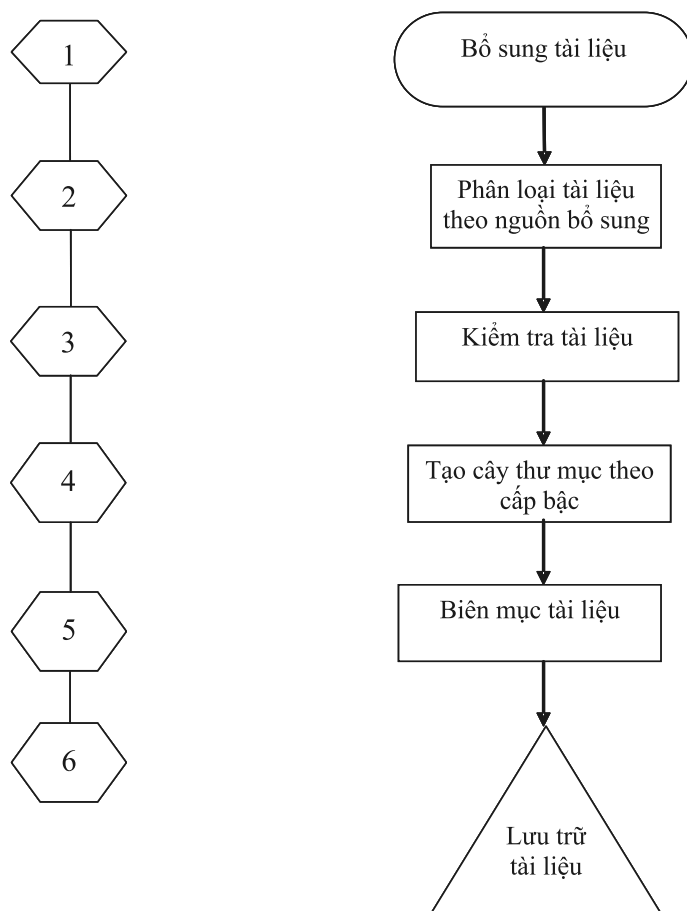
Dspace được Trung tâm Phát triển phần mềm của UEF Việt hóa và phát triển, được TTTT-TV sử dụng để xây dựng các bộ sưu tập số và quản lý việc khai thác, sử dụng tài liệu điện tử. Sau bốn năm nỗ lực làm việc (tính đến 02/2012), sáu bộ sưu tập số phục vụ học tập, nghiên cứu của sinh viên, giảng viên,

cán bộ quản lý của UEF đã được xây dựng hoàn chỉnh, cập nhật thường xuyên, bao gồm: “Kiến thức đại cương và chuyên ngành”, “Tài liệu học tiếng Anh”, “Tài liệu các lĩnh vực khoa học khác”, “Tài liệu đa phương tiện”, “Bài trích báo - tạp chí”, và “Luận văn-luận án”. 22.722 nhan đề tài liệu điện tử hiện đang được quản lý khai thác, sử dụng khá hiệu quả.

2. Kinh nghiệm sử dụng phần mềm mã nguồn mở Dspace

Tổng kết quá trình sử dụng phần mềm Dspace, TTTT-TV UEF đã rút ra sáu bài học kinh nghiệm như sau.

2.1. Xác định quy trình tạo lập bộ sưu tập số một cách rõ ràng, nhất quán và chi tiết



Hình 1: Sơ đồ quy trình tạo lập bộ sưu tập số trên phần mềm Dspace tại trường UEF

Nghiên cứu - Trao đổi

Quy trình này gồm sáu bước như sau (Hình 1):

- Bước 1. Bổ sung tài liệu điện tử: tài liệu điện tử được chọn lọc, bổ sung từ nhiều nguồn khác nhau như: số hóa, mua, trao đổi, nhận tặng, tải từ Internet xuống, v.v..

- Bước 2. Phân loại tài liệu điện tử theo nguồn bổ sung: để thuận tiện cho việc quản lý, thống kê, tài liệu được phân loại theo nguồn như: nguồn nội sinh và nguồn bên ngoài.

- Bước 3. Kiểm tra tài liệu điện tử:

+ Kiểm tra kiểu file: phần mềm quản lý và viewer chỉ hỗ trợ một số kiểu file thông dụng như .pdf, .prc... Nếu tài liệu thuộc kiểu file khác thì phải chuyển về các kiểu file thông dụng.

+ Kiểm tra file có mật khẩu: cần dùng phần mềm mở mật khẩu để đưa file về trạng thái bình thường.

+ Kiểm tra dung lượng file: nếu file có dung lượng quá lớn, không thể đính kèm vào phần mềm thì dùng ACDsee để giảm dung lượng, chuyển sang chế độ trắng đen, giảm độ phân giải hay tách file dung lượng lớn thành nhiều file nhỏ hơn.

- Bước 4. Tạo cây thư mục theo cấp bậc:

+ Phân loại tài liệu: sử dụng bảng phân loại DDC để phân loại tài liệu.

+ Sắp xếp tài liệu theo môn loại khoa học, trong môn loại khoa học sắp xếp theo thứ tự chữ cái nhan đề tài liệu.

+ Đối với luận văn, luận án: sắp xếp theo trình độ tiến sĩ, thạc sĩ, sau đó tiếp tục phân theo môn loại khoa học và theo chữ cái tên luận văn, luận án.

+ Đối với bài trích báo, tạp chí: sắp xếp theo tên báo, tạp chí, năm, số, bài trong số.

- Bước 5. Biên mục tài liệu điện tử trên phần mềm Dspace:

+ Dspace sử dụng siêu dữ liệu (metadata) gồm 15 trường Dublin Core, trong các trường còn có các trường con để biên mục chi tiết cho tài liệu. Tùy vào từng loại hình tài liệu mà sử dụng các trường chính và các trường con cho phù hợp.

+ Sử dụng bộ từ khóa, từ điển từ chuẩn để xác định chủ đề, từ khóa cho tài liệu.

+ Sử dụng quy tắc biên mục AACR2 để biên mục tài liệu.

+ Kiểm tra các biểu ghi biên mục trước khi đưa ra phục vụ bạn đọc.

+ Xây dựng bookmarks cho tài liệu.

- Bước 6. Lưu trữ tài liệu:

+ Lưu trữ tài liệu sang các ổ cứng rời nhằm đề phòng rủi ro.

+ Tài liệu điện tử được lưu trữ trên một server khác.

+ Backup metadata theo ngày.

2.2. Xây dựng chính sách khai thác và sử dụng tài liệu điện tử

Thư viện của UEF quản lý việc khai thác và sử dụng tài liệu điện tử trên Dspace bằng cách tạo cho người sử dụng một tài khoản để đăng nhập. Tài khoản này cũng là tài khoản đăng nhập trên webmail trường của người sử dụng. Người dùng tin (NDT) có thể tìm kiếm tài liệu bằng các hình thức tìm kiếm cơ bản và nâng cao. Trong mục tìm kiếm nâng cao NDT có thể kết hợp các toán tử AND, OR, NOT và có thể tìm kiếm với các điểm truy cập như nhan đề, tác giả,

từ khóa, chủ đề,... NDT có thể chọn tìm kiếm trên tất cả các bộ sưu tập số hay từng bộ sưu tập số. Nhằm tạo thuận lợi cho NDT, các bộ sưu tập số trên phần mềm Dspace được liệt kê theo CSDL, thời gian xuất bản, tác giả, nhan đề, từ khóa.

Chính sách khai thác và sử dụng tài liệu điện tử đã quy định: muốn xem toàn văn tài liệu, người sử dụng phải đăng nhập vào tài khoản của mình, xem được biểu ghi biên mục chi tiết tài liệu và đọc được toàn văn của tài liệu.

Tài liệu điện tử trên Dspace được quản lý theo hai chế độ: đọc nhưng không tải xuống được; đọc và tải xuống được.

Thư viện cho phép NDT tải xuống nhằm phục vụ học tập, nghiên cứu đối với những tài liệu được thư viện tải xuống miễn phí từ Internet. Để tải xuống các tài liệu bổ sung từ nguồn mua, nguồn trao đổi, ... NDT phải liên hệ với thủ thư và tuân thủ Luật Quyền tác giả.

Thực tế phục vụ cho thấy, NDT tra cứu, xem toàn văn tài liệu điện tử trên Dspace với tốc độ khá nhanh, giao diện thân thiện, dễ dàng sử dụng. NDT có thể di chuyển giữa các trang với nhau với tốc độ nhanh. Ngoài ra, việc xây dựng bookmarks cho tài liệu đã tạo điều kiện thuận lợi cho NDT nhanh chóng tiếp cận những thông tin của tài liệu, tiết kiệm thời gian, công sức cho họ.

2.3. Xây dựng cơ chế phân quyền và bảo mật

Cho đến nay, việc phân quyền trên Dspace của TTTT-TV của UEF còn bị hạn chế, người dùng và người quản trị đều có quyền như nhau. Điều này dẫn

đến khó khăn trong quản lý và ảnh hưởng đến độ an toàn dữ liệu. Bài học rút ra là: cần phát triển tính năng phân quyền của phần mềm Dspace, cần phân biệt rõ ràng quyền của người quản trị và người sử dụng. Trong từng bộ sưu tập, phải có cơ chế phân quyền riêng. Cần xây dựng chế độ backup dữ liệu để đảm bảo tính an toàn cho dữ liệu. Bên cạnh đó, nên phân chia cấp độ mật (từ thấp đến cao) cho từng tài liệu khác nhau trong bộ sưu tập. Việc quản lý mật khẩu sử dụng phần mềm phải được thực hiện một cách cẩn trọng, không nên đặt mật khẩu quá đơn giản, không được tiết lộ mật khẩu với người khác nhằm đảm bảo tính bảo mật, độ an toàn cho dữ liệu.

2.4. Việc tích hợp giữa phần mềm DSPACE và các phần mềm quản lý thư viện khác không đơn giản, có thể phải chấp nhận việc thiếu khả năng tích hợp

Khi ứng dụng Dspace với mục đích quản lý bộ sưu tập số, TTTT-TV của UEF đã muốn tích hợp Dspace với phần mềm quản lý thư viện LIBOL để quản lý các biểu ghi thư mục của các vật mang tin khác nhau. Tuy nhiên, thực tế cho thấy việc làm này thật sự không đơn giản, không thể quản lý tập trung được. Trong quá trình biên mục, xảy ra khả năng phải biên mục hai lần, lần đầu trên Dspace theo chuẩn Dublin Core và lần sau trên phần mềm quản lý thư viện theo chuẩn MARC21. Điều này liên quan đến việc cấp quyền, phân quyền cho các đối tượng khác nhau. Phải cấp quyền trên cả hai phần mềm do CSDL người dùng khác nhau. Ngoài ra, cán bộ thư viện sẽ gặp khó khăn trong việc phân nhóm, phân quyền các bạn đọc khác nhau tiếp cận với các nguồn tư liệu cấp độ mật

khác nhau, hoặc nguồn tư liệu không/có thu phí. Khó khăn nhất là việc kiểm tra tài khoản hay việc thanh toán của NDT vì CSDL bạn đọc bị rời rạc.

2.5. Thiết lập mối quan hệ thân hữu với bộ phận công nghệ thông tin (CNTT) của cơ quan, tổ chức chủ quản nếu thư viện không có bộ phận CNTT riêng

Hoạt động thư viện luôn cần các phần mềm đáp ứng yêu cầu cao về mặt nghiệp vụ. Việc cài đặt phần mềm Dspace khá phức tạp, khó tùy biến, trong khi số chuyên viên thư viện am hiểu về PMMNM lại ít. Phần mềm Dspace sử dụng hệ quản trị CSDL PostgreSQL và ngôn ngữ lập trình Java ít người biết. Thực tế cho thấy, để các chuyên viên học và sử dụng thành thạo phần mềm cần phải có thời gian và thường xuyên cần sự hỗ trợ của CNTT, đặc biệt trong thời gian đầu.

Tại UEF, để đảm bảo sự vận hành liên tục, ổn định của hệ thống phục vụ, để có đội ngũ CNTT đủ năng lực phát triển PMMNM nói chung và Dspace nói riêng theo yêu cầu của các đơn vị chức năng, lãnh đạo nhà trường đã trả lương cao cho đội ngũ CNTT này. Tuy nhiên, do công việc quá nhiều, họ buộc phải xếp cho chúng thứ tự ưu tiên. Vì vậy, thư viện cần thiết lập quan hệ chặt chẽ, thân thiết với bộ phận CNTT để công việc của mình được CNTT ưu tiên giải quyết. Bên cạnh đó, cán bộ quản trị Dspace của thư viện phải thường xuyên kiểm tra các phân hệ trong phần mềm, phát hiện nhanh chóng vấn đề để phối hợp với bộ phận CNTT giải quyết kịp thời.

Về lâu dài, do sử dụng phần mềm mở

nên việc hỗ trợ, phát triển sản phẩm, khắc phục lỗi trong quá trình sử dụng tốt nhất do CNTT của bản thân thư viện đảm nhận.

2.6. Tích cực huấn luyện người sử dụng

Thư viện phải tổ chức huấn luyện người sử dụng các kỹ năng tìm kiếm (từ cơ bản đến nâng cao) và khai thác tài liệu điện tử; hỗ trợ người sử dụng cài đặt một số phần mềm hỗ trợ đọc các kiểu file như .prc, .pdf,... và phải luôn sẵn sàng hỗ trợ người sử dụng khi họ yêu cầu.

Quản lý tài liệu điện tử tại TTTT-TV của UEF bằng phần mềm Dspace còn tồn tại một số điểm yếu hiện chưa khắc phục được như:

- Chưa phát triển được tính năng phân quyền giữa người dùng và người quản trị nhằm tăng cường bảo mật cho các CSDL được khai thác phục vụ trên phần mềm;

- Cơ chế phân chia cấp độ mật chưa hoàn thiện: phần mềm Dspace không quản lý và không gán được cấp độ mật khác nhau (từ thấp đến cao) cho từng tài liệu trong bộ sưu tập số;

- Dspace không hỗ trợ tất cả các kiểu file điện tử ở chế độ view only, gây khó khăn trong quản lý việc tải xuống, in ấn của người sử dụng. Chưa phát triển được tính năng cấp quyền cho in ấn và tải xuống đối với các tài liệu chỉ xem (view only);

- Chưa phát triển được tính năng giới hạn in ấn, tải xuống ở bất kỳ nơi nào có thể kết nối tới máy in (người sử dụng chỉ được phép in 5 trang/tài liệu);

- Chưa giới hạn được số lượng tài liệu người sử dụng được phép tải (tối đa 3 tài

Nghiên cứu - Trao đổi

liệu/ngày/người);

- Chưa thống kê được số lượng người sử dụng có đăng ký, số lượng trang được in, tài liệu được in, số lượng truy cập cho từng bộ sưu tập;

- Chưa phát triển được tính năng phát hiện trùng khi nhập vào một tài liệu có nhan đề giống tài liệu đã được nhập trước đó;

- Chưa tạo được công cụ nhập/xuất dữ liệu để cán bộ thư viện chủ động trong việc này;

- Chưa tạo được công cụ sao lưu (backup) dữ liệu;

- Chưa tạo được mục *người biên mục* trong phần biên mục và hiển thị ngoài các tiêu chí tra cứu để lọc lại tất cả các biểu ghi của từng người biên mục, nhằm chỉnh sửa và kiểm tra lại dữ liệu của từng người. (Hiện thư viện đang dùng tạm mục Người tài trợ).

- Chưa hỗ trợ tìm toàn văn và các chuẩn Unicode (có trường hợp khi nhập liệu chọn Unicode dựng sẵn thì khi tìm bắt buộc cũng phải theo Unicode dựng sẵn, không thuận tiện cho người sử dụng);

- Chưa phát triển tính năng tạo *người dùng* khi phục vụ các đối tượng bên ngoài cơ quan;

- Chưa có khả năng tạo các nhóm bạn đọc với những chính sách và quyền truy cập khác nhau vào từng loại hình tài liệu

với những cấp độ mật khác nhau;

- Không liên thông được với các CSDL của các thư viện khác;

- Không hỗ trợ quản lý rõ ràng báo, tạp chí, bài viết, sự liên thông giữa các CSDL;

- Việc biên mục theo chuẩn Dublin Core tuy đơn giản nhưng còn nhiều hạn chế như:

- + Chưa lột tả được chính xác các thông tin thư mục giữa các biểu ghi;

- + 15 trường Dublin Core chỉ phù hợp cho mô tả ấn phẩm đơn, hiển thị thông tin trong một ấn phẩm, không mô tả được sự liên đới giữa biểu ghi con và biểu ghi chủ. Trong quá trình sử dụng Dublin Core có thể có sự không thống nhất về tác giả biên soạn, tác giả hiệu đính, dấu phân cách, v.v...

3. Kết luận

Mặc dù còn một số tồn tại kể trên, việc ứng dụng Dspace để xây dựng, quản lý tài liệu điện tử tại TTTT-TV của UEF đã mang lại hiệu quả cao, thiết thực. TTTT-TV được giảng viên, sinh viên UEF và lãnh đạo nhà trường, các đoàn kiểm tra của Bộ Giáo dục và Đào tạo, Quốc hội đánh giá cao. Trong điều kiện kinh phí còn hạn hẹp, chưa thể trang bị phần mềm thương mại, hy vọng những bài học kinh nghiệm này sẽ bổ ích đối với các thư viện muốn ứng dụng PMMN Dspace vào hoạt động của mình.

Tài liệu tham khảo

1. Vũ Thị Minh Hương (2007), "Công nghệ thông tin góp phần thực hiện nhiệm vụ "bảo vệ và phát huy giá trị của tài liệu lưu trữ", Văn thư lưu trữ, (1), tr. 8- 11.

2. Các Báo cáo tổng kết hoạt động hàng năm của TTTT-TV UEF.

3. <http://www.uef.edu.vn>; <http://www.mediafire.com>; <http://www.thietkewebvietnam.net>; <http://www.ctu.edu.vn>