

Nguyễn Huy Chương

THƯ VIỆN SỐ VÀ CÔNG NGHỆ NỘI DUNG

**(Giáo trình sau đại học
chuyên ngành Khoa học Thông tin - Thư viện)**

Hà Nội, 2017

Mục lục

LỜI NÓI ĐẦU.....	5
CHƯƠNG 1: TỔNG QUAN VỀ THƯ VIỆN SỐ.....	6
1.1. MỘT SỐ KHÁI NIỆM	6
1.1.1. Thư viện điện tử	6
1.1.2. Thư viện số.....	8
1.1.2. Thư viện ảo	11
1.2. VAI TRÒ , ĐẶC TÍNH VÀ LOẠI HÌNH THƯ VIỆN SỐ.....	12
1.2.1. Vai trò của thư viện số.....	12
1.2.2. Đặc tính của thư viện số	18
1.2.3. Loại hình thư viện số	18
1.3 CÁC CHỨC NĂNG CƠ BẢN VÀ DỊCH VỤ CỦA THƯ VIỆN SỐ .	20
1.3.1. Các chức năng cơ bản	20
1.3.2. Các dịch vụ	21
1.4 CÁC NGUYÊN TẮC CƠ BẢN, TIÊU CHÍ ĐÁNH GIÁ VÀ MỘT SỐ VẤN ĐỀ KHI XÂY DỰNG THƯ VIỆN SỐ	21
1.4.1. Các nguyên tắc cơ bản	21
1.4.2. Tiêu chí đánh giá	22
1.4.3. Một số vấn đề khi xây dựng thư viện số	25
1.5. THƯ VIỆN SỐ VÀ HỌC LIỆU MỞ, TRUY CẬP MỞ	27
1.5.1. Khái quát về học liệu mở và tài nguyên giáo dục mở	27
1.5.2. Truy cập mở (Open Access - OA).....	31
1.6. TÌNH HÌNH PHÁT TRIỂN THƯ VIỆN SỐ TRÊN THẾ GIỚI.....	35
1.6.1. Ý tưởng và tiền đề xây dựng thư viện số.....	35
1.6.2. Quá trình phát triển thư viện số	36
1.7 TÌNH HÌNH PHÁT TRIỂN THƯ VIỆN SỐ Ở VIỆT NAM	37
1.7.1. Chính sách phát triển thư viện số ở Việt Nam.....	37
1.7.2. Xây dựng thư viện số trong các trường đại học	38

CÂU HỎI ÔN TẬP CHƯƠNG 1	46
CHƯƠNG 2: CÁC THÀNH PHẦN CỦA THƯ VIỆN SỐ	47
2.1 VỐN TÀI LIỆU SỐ	48
2.1.1. Các đối tượng số.....	48
2.1.2. Siêu dữ liệu	49
2.2 HẠ TẦNG KỸ THUẬT	63
2.2.1. Phần cứng.....	63
2.2.2. Phần mềm ứng dụng.....	91
2.3 CÁN BỘ THƯ VIỆN SỐ	111
2.3.1. Nội dung công việc chủ yếu.....	112
2.3.2. Cách thức phục vụ	113
2.4. DỊCH VỤ THƯ VIỆN SỐ VÀ NGƯỜI DÙNG THƯ VIỆN SỐ	114
2.4.1. Dịch vụ thư viện số	114
2.4.2. Người dùng thư viện số.....	117
CÂU HỎI ÔN TẬP CHƯƠNG 2	121
CHƯƠNG 3: SỰ TẦM VÀ KHAI THÁC THÔNG TIN SỐ	122
3.1 THÔNG TIN TRÊN INTERNET	122
3.1.1 Công cụ tìm kiếm (Search Engine).....	124
3.1.2. Tài nguyên điện tử.....	129
3.1.3. Chiến lược tìm kiếm thông tin	130
3.2 CSDL TRỰC TUYẾN THƯƠNG MẠI	133
3.3 SỐ HÓA NGUỒN TIN NỘI SINH.....	136
3.3.1. Khái niệm và vai trò của số hóa.....	136
3.3.2. Chính sách, kế hoạch và quy trình số hóa.....	138
3.3.3. Thiết bị số hóa	142
3.3.4. Nhận dạng ký tự quang học: OCR	146
3.4. HỆ THỐNG PHÁT HIỆN VÀ TÌM KIẾM TẬP TRUNG	147

<i>Phần mềm tìm kiếm tập trung Vufind</i>	149
CÂU HỎI ÔN TẬP CHƯƠNG 3	151
CHƯƠNG 4: BỘ SƯU TẬP VÀ BIÊN MỤC TÀI LIỆU SỐ	152
4.1 CÁC KHÁI NIỆM.....	152
4.2 Ý NGHĨA VÀ NHU CẦU.....	154
4.3 MỘT SỐ BỘ SƯU TẬP MẪU.....	156
4.3.1. <i>Các bộ sưu tập theo loại hình xuất bản</i>	156
4.3.2. <i>Các bộ sưu tập theo dạng lưu trữ tài liệu</i>	160
4.4 SỬ DỤNG CÁC BỘ SƯU TẬP	166
4.5 CÔNG CỤ XÂY DỰNG BỘ SƯU TẬP.....	166
4.6 BIÊN MỤC TÀI LIỆU SỐ	174
4.6.1. <i>Biên mục Analog</i>	174
4.6.2. <i>Biên mục Digital</i>	174
4.6.3. <i>Chuyển đổi MARC sang Dublin Core</i>	175
CÂU HỎI ÔN TẬP CHƯƠNG 4	177
CHƯƠNG 5: CÁC VẤN ĐỀ VỀ SỞ HỮU TRÍ TUỆ VÀ BẢN QUYỀN	178
5.1 MỘT SỐ KHÁI NIỆM VỀ SỞ HỮU TRÍ TUỆ	178
5.2 CÁC VẤN ĐỀ VỀ BẢN QUYỀN	185
CÂU HỎI ÔN TẬP CHƯƠNG 5	188
CHƯƠNG 6: XÁC ĐỊNH CHÍNH SÁCH VÀ XÂY DỰNG KẾ HOẠCH PHÁT TRIỂN THƯ VIỆN SỐ	189
6.1 XÁC ĐỊNH CHÍNH SÁCH PHÁT TRIỂN THƯ VIỆN SỐ	189
6.1.1 <i>Cấu trúc của thư viện số</i>	189
6.1.2. <i>Hạ tầng cơ sở kỹ thuật</i>	190
6.1.3. <i>Kho tư liệu số hóa</i>	191

6.1.4. Các vấn đề bảo quản, khai thác và bản quyền.....	191
6.2 XÂY DỰNG KẾ HOẠCH PHÁT TRIỂN NGUỒN TIN SỐ.....	191
6.3 XÂY DỰNG KẾ HOẠCH TÀI CHÍNH.....	194
CÂU HỎI ÔN TẬP CHƯƠNG 6.....	196
CHƯƠNG 7: CÔNG NGHỆ NỘI DUNG VÀ CÔNG NGHIỆP NỘI DUNG	197
7.1 KHÁI NIỆM	197
7.1. Khái niệm nội dung.....	197
7.1.2. Khái niệm công nghệ nội dung.....	198
7.1.3. Khái niệm công nghiệp nội dung.....	199
7.2 CÔNG NGHỆ NỘI DUNG	200
7.2.1. Số hóa tài liệu	200
7.2.2. Xuất bản điện tử	201
7.2.3. Hệ quản trị nội dung.....	202
7.2.4. Công nghệ thư viện số	206
7.3 CÔNG NGHIỆP NỘI DUNG	210
7.3.1. Đặc điểm chung.....	210
7.3.2. Công nghiệp nội dung (số) thế giới.....	213
7.3.3. Công nghiệp nội dung Việt Nam	217
7.3.4. Công nghiệp nội dung trong hoạt động thông tin - thư viện.....	224
CÂU HỎI ÔN TẬP CHƯƠNG 7	227
TÀI LIỆU THAM KHẢO	228

LỜI NÓI ĐẦU

Môn học “Thư viện số và công nghệ nội dung” nằm trong chương trình đào tạo trình độ thạc sĩ chuyên ngành khoa học thông tin - thư viện của Khoa Thông tin - Thư viện, Trường Đại học Khoa học Xã hội và Nhân văn, Đại học Quốc gia Hà Nội. Môn học có dung lượng 3 tín chỉ nhằm trang bị những kiến thức cơ bản và quan trọng nhất về thư viện số và công nghệ nội dung cho học viên cao học.

Sau phần giới thiệu tổng quan về thư viện số như khái niệm, vai trò, chức năng, nguyên tắc, truy cập mở và tình hình phát triển thư viện số cả trong và ngoài nước, Giáo trình tập trung phân tích các đặc điểm và nội dung chính của thư viện số. Đó là: thành phần cấu tạo thư viện số; sưu tầm và khai thác thông tin số; bộ sưu tập và biên mục tài liệu số; vấn đề bản quyền và sở hữu trí tuệ; xây dựng kế hoạch và chính sách phát triển thư viện số.

Công nghệ và công nghiệp nội dung được đề cập trên các khía cạnh liên quan đến thư viện số và các dịch vụ công nghiệp nội dung phổ biến tại Việt Nam và một số quốc gia trên thế giới.

Thư viện số và công nghệ nội dung được đổi mới và phát triển rất nhanh chóng. Cập nhật kiến thức về hai chủ đề này cần được tiến hành thường xuyên. Tác giả rất mong nhận được sự góp ý của các nhà khoa học và đồng nghiệp để tiếp tục hoàn thiện Giáo trình này.

Xin chân thành cảm ơn!

Tác giả

CHƯƠNG 1: TỔNG QUAN VỀ THƯ VIỆN SỐ

Mục tiêu đào tạo: Trang bị kiến thức tổng quan về thư viện số bao gồm khái niệm, vai trò, chức năng, dịch vụ, nguyên tắc, truy cập mở và một số vấn đề cần lưu ý khi xây dựng thư viện số. Đồng thời giới thiệu bức tranh chung về tình hình phát triển thư viện số trên thế giới và tại Việt Nam.

1.1. MỘT SỐ KHÁI NIỆM

1.1.1. Thư viện điện tử

Thư viện điện tử là một khái niệm chưa được định nghĩa thống nhất và còn nhiều tranh luận, đôi khi dùng lẫn lộn và đồng nghĩa với các khái niệm "Thư viện không biên giới", "Thư viện được nối mạng", "Thư viện số", "Thư viện ảo", "Thư viện tin học hoá", "Thư viện đa phương tiện", "Thư viện lôgic", "Thư viện văn phòng",....

Thuật ngữ "thư viện điện tử" (electronic library) có thể dùng theo nghĩa tổng quát nhất cho mọi loại hình thư viện đã tin học hoá toàn bộ hoặc một số dịch vụ.

Thư viện điện tử có thể được coi như là nơi người sử dụng có thể tới để thực hiện những công việc mà họ vẫn thường làm với thư viện truyền thống, nhưng đã được điện tử hoá.

Năm 1998, tiến sĩ Ching-chih Chen - người đã có sáng kiến tổ chức một loạt hội nghị quốc tế về công nghệ thông tin mới (NIT) cho rằng hiện không có một tiêu chuẩn cố định, chính thức nào cho thư viện điện tử. Người ta sử dụng khái niệm này khá tự do, tùy tiện.

Theo quan điểm của Collier thì thư viện điện tử được định nghĩa như là một môi trường gồm các tài liệu dưới dạng điện tử, được cấu trúc nhằm cung cấp một lượng thông tin lớn thông qua các máy tính hoặc các mạng viễn thông quốc tế.

Phillip Barker cho rằng: Trong thư viện điện tử có sử dụng rộng rãi

máy tính và các phương tiện hỗ trợ khác (bảng tra trực tiếp, tìm văn bản đầy đủ, lưu các biểu ghi tự động hoá, ra các quyết định bằng máy tính,...). Tác giả nhấn mạnh đặc trưng của thư viện điện tử là sử dụng phổ biến các phương tiện điện tử trong lưu trữ, tìm kiếm và cung cấp thư viện điện tử. Theo ông trong thư viện điện tử, ngoài ấn phẩm điện tử vẫn còn tồn tại cả sách truyền thống.

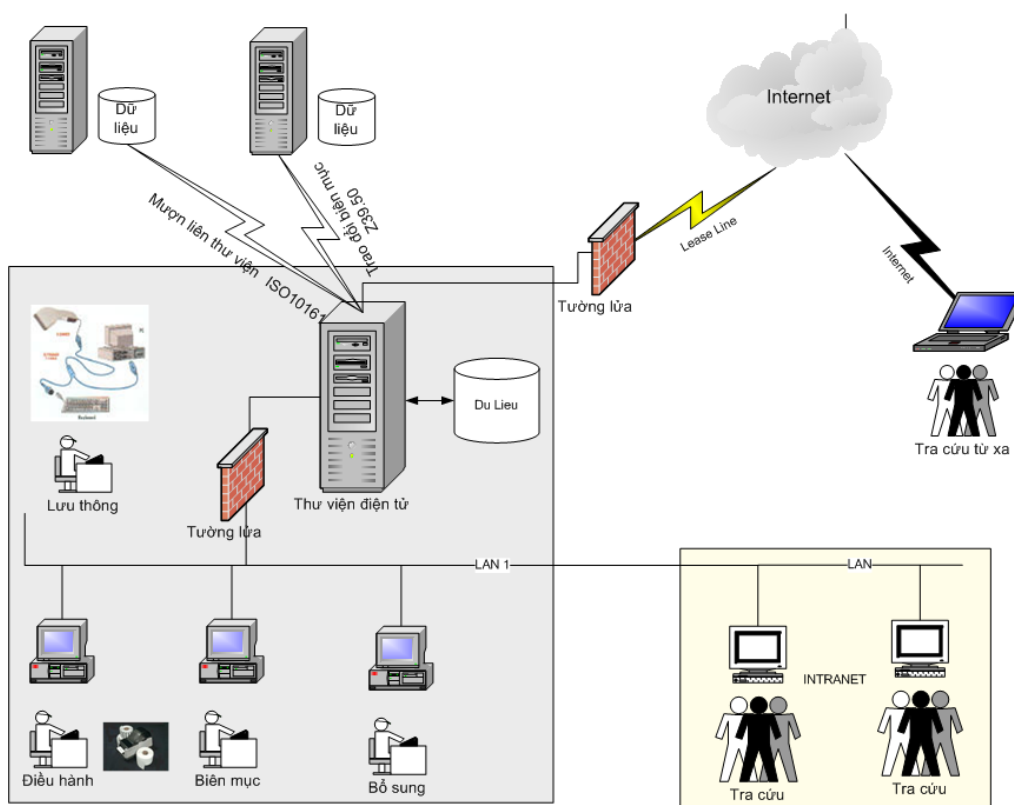
Sylvie Tellier thì đưa ra định nghĩa về thư viện điện tử như sau: Thư viện điện tử là thư viện có sử dụng hệ thống máy vi tính và các hệ thống phụ kiện của nó để lưu trữ, xử lý, cung cấp dịch vụ thông tin cho người sử dụng. Theo cách hiểu như vậy thì thư viện điện tử ở đây có sử dụng máy tính trong việc quản lý, lưu trữ và phục vụ tìm kiếm thông tin [47]

Tuy ý kiến chưa hoàn toàn thống nhất, nhưng tựu chung lại, ta có thể nhận dạng một số đặc điểm của thư viện điện tử như sau:

- Thư viện phải có vốn tư liệu điện tử (là những tư liệu được lưu trữ dưới dạng số sao cho có thể truy nhập được bằng các thiết bị xử lý dữ liệu).
- Phải được tin học hoá, phải có một hệ quản trị thư viện tích hợp (bổ sung, biên mục, quản trị xuất bản phẩm định kỳ, kiểm soát lưu thông tư liệu, tổ chức mục lục truy nhập công cộng trực tuyến,...); phải nối mạng (ít nhất là mạng cục bộ).
- Phải cung cấp và tạo điều kiện cho người dùng sử dụng các dịch vụ điện tử (yêu cầu và gia hạn mượn qua mạng, tìm tin trong các cơ sở dữ liệu, truy nhập và khai thác các nguồn tin tại chỗ và các nguồn tin ở nơi khác,...).

Thư viện điện tử ra đời là kết quả của sự hợp tác giữa các chuyên gia thư viện, xuất bản, các nhà khoa học và công nghệ hướng về mục tiêu tiếp cận tới đầy đủ thông tin, ở mọi nơi và mọi lúc.

Nói tóm lại, ***thư viện điện tử phải sử dụng các phương tiện điện tử trong thu thập, lưu trữ, xử lý, tìm kiếm và phổ biến thông tin.***



Hình 1.1 Mô hình thư viện điện tử

1.1.2 Thư viện số

Có ý kiến cho rằng thư viện số là một bước tiến xa hơn của thư viện điện tử hay có thể nói cách khác, **là thư viện điện tử cấp cao**, cho phép đọc được thông tin toàn văn sau khi đã số hoá hầu hết tư liệu, đặc biệt là các tư liệu dưới dạng đồ hoạ (như tranh ảnh, bản đồ,...) và đa phương tiện (multimedia) nói chung.

Tác giả Barnes [3] thì phân biệt thư viện điện tử và thư viện số theo một kiểu khác. Ông cho rằng thư viện điện tử lưu trữ và phục vụ **cả ấn phẩm (tài liệu in ấn truyền thống) lẫn tư liệu điện tử (tư liệu số hoá)**, trong khi đó thư viện số chỉ lưu trữ các tư liệu điện tử mà thôi.

Một thư viện điện tử có thiên hướng sử dụng linh hoạt và phổ biến các nguồn tin điện tử nhưng đồng thời cũng tham gia vào việc tạo ra các nguồn tin đó.

Các thư viện số cũng có nhiều định nghĩa khác nhau và nhiều công

trình nghiên cứu với quan điểm khác nhau.

Thư viện số theo quan điểm của Liên đoàn Thư viện số thế giới (DLF - Digital Library Federation): Thư viện số là tổ chức cung cấp các nguồn lực - tài nguyên, bao gồm cả các chuyên gia để lựa chọn, cấu trúc, cung cấp khả năng truy cập tới các nguồn tri thức, phân phối, bảo đảm tính vẹn toàn và tính lâu dài của các bộ sưu tập số để cho một cộng đồng hoặc một tập hợp cộng đồng người dùng tin xác định luôn có thể sử dụng một cách nhanh chóng, kịp thời và kinh tế [21].

Theo quan điểm của Liên hiệp Thư viện số của Mỹ (American Digital Feder - ADF) thì: Thư viện số là cơ quan, tổ chức có các nguồn lực, kể cả nguồn nhân lực chuyên môn hoá để lựa chọn cấu trúc, diễn giải, phổ biến, bảo quản sự toàn vẹn, đảm bảo sự ổn định trong thời gian dài của bộ sưu tập các công trình số hoá mà chúng ta có ở dạng sẵn sàng để sử dụng một cách kinh tế cho một hay một số cộng đồng nhất định [13].

Theo Michael Lesk [22]: Thư viện số là bộ sưu tập thông tin số hoá có tổ chức. Được xây dựng bằng cách cấu trúc và thu thập thông tin là các công việc mà các thư viện truyền thống vẫn luôn phải làm và các máy tính có nhiệm vụ trình bày các thông tin số đó,... Một thư viện số thực sự cũng tạo ra các nguyên tắc quản lý những yếu tố cấu thành thư viện và các phương thức tổ chức thư viện.

Theo Borgman [4]: Các thư viện số được xây dựng, lựa chọn và tổ chức cho một cộng đồng người dùng tin và chúng có khả năng thoả mãn các nhu cầu tin và cung cấp ích lợi cho cộng đồng ấy. Thư viện số là một bộ phận cấu thành lên các cộng đồng mà ở đó các cá nhân và các nhóm có thể tương tác với nhau, sử dụng dữ liệu, thông tin, các tài nguyên và hệ thống tri thức. Ở định nghĩa này, thư viện số chính là sự phát triển ở mức cao hơn và là sự tích hợp của các tổ chức thông tin ở dạng vật lý, nơi mà các tài nguyên thông tin được lựa chọn, thu thập, tổ chức, bảo

quản và được truy cập để phục vụ cho cộng đồng người dùng tin. Những tổ chức thông tin này gồm các thư viện, viện bảo tàng, cơ quan lưu trữ và các trường học. Song các thư viện số này lại phát triển và vươn tới phục vụ các cộng đồng khác bao gồm cả các lớp học, công sở, văn phòng, phòng thí nghiệm, gia đình, các khu vực công cộng...

Theo Ian Written [45]: Thư viện số là tập hợp những bộ sưu tập thông tin của các đối tượng số hoặc đã được số hoá 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 chức 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).

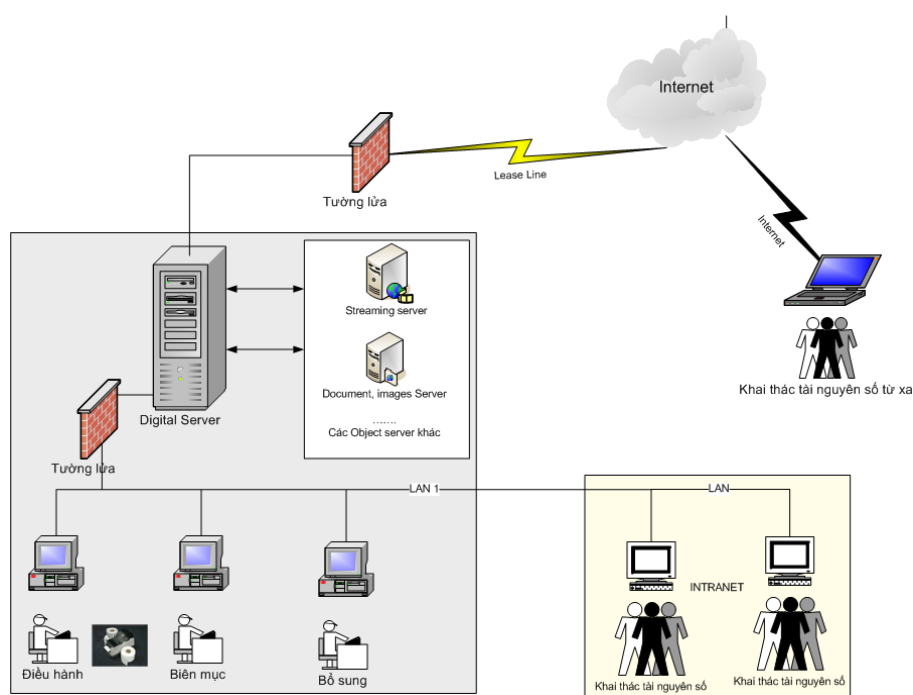
Theo Fox Thư viện số là tập hợp của các máy tính, các thiết bị máy móc lưu trữ và trao đổi thông tin cùng với bối cảnh số và phần mềm cần thiết để sản xuất và cung cấp các dịch vụ thông tin tương tự như các thư viện truyền thống vẫn làm đối với tài liệu giấy và các loại hình tài liệu truyền thống khác trong quá trình thu thập, biên mục, tìm kiếm và phổ biến thông tin,... Một thư viện số đúng nghĩa và hoàn chỉnh phải bao gồm tất cả các dịch vụ cơ bản của thư viện truyền thống, đồng thời tận dụng được các lợi thế của việc lưu trữ, tìm kiếm và cung cấp thông tin số hoá.

Theo Gladney: Một thư viện số phải là một tập hợp các thiết bị máy tính, lưu trữ, truyền thông cùng với các nội dung số và phần mềm để tái tạo, thúc đẩy và mở rộng các dịch vụ thông tin của các thư viện truyền thống chứa các tài liệu trên giấy và các vật mang tin khác vẫn làm như thu thập, biên mục, tìm kiếm, phân phát thông tin. Một dịch vụ của thư viện số đầy đủ trọn vẹn phải bao gồm đảm bảo có cả các dịch vụ chính yếu

của các thư viện truyền thống và khai thác tối đa các ích lợi của công nghệ lưu trữ số, tìm kiếm thông tin số và truyền thông số [14].

Năm 2004, trong cuốn Từ điển khoa học thông tin thư viện của Nhà xuất bản Libraries Unlimited, tác giả Joan M.Reitz lại khẳng định **thư viện điện tử và thư viện số là một** [20]

Theo tác giả, **thư viện số là bộ sưu tập thông tin có tổ chức, là tập hợp các đối tượng dữ liệu số và tài liệu in ấn truyền thống bao gồm cả văn bản, hình ảnh, âm thanh... được quản trị, truy cập, khai thác thông qua hệ thống hạ tầng công nghệ thông tin và truyền thông.**



Hình 1.2 Mô hình thư viện số

1.1.2. Thư viện ảo

Là "thư viện không có bức tường", trong đó bộ sưu tập không tồn tại trên giấy, trên dạng vi thể hoặc các hình thức hữu hình khác ở vị trí thực nhưng có thể truy cập dưới dạng điện tử thông qua mạng máy tính. Một số thư viện và hệ thống thư viện tự gọi mình là "ảo" vì chúng cung cấp các dịch vụ trực tuyến (ví dụ: Thư viện Ảo Colorado).

Thuật ngữ "thư viện ảo" (virtual library) dùng theo nghĩa trừu tượng,

là một dạng của thực tế ảo (virtual reality), được xây dựng trên cơ sở công nghệ ảo (đôi khi phối hợp với kỹ thuật âm thanh nổi và hình ảnh nổi để tạo ảo giác như thực), nhấn mạnh đến tính chất "phi không gian" của loại hình thư viện này về phương diện vốn tư liệu và dịch vụ. Bất cứ thư viện nào tạo điều kiện cho người đọc tiếp cận được những tư liệu nằm tại bất cứ nơi nào khác trên thế giới đều có thể được coi là "thư viện ảo". Nói cách khác, thư viện ảo không phụ thuộc vào một địa điểm cố định và cho phép truy nhập thông tin từ xa thông qua mạng. Còn thư viện điện tử có một địa điểm cụ thể, hữu hình, nơi bạn đọc hay người sử dụng có thể tới để nhận những sản phẩm và dịch vụ được cung cấp dưới dạng điện tử. Tuy nhiên, cũng có ý kiến cho rằng thư viện ảo nằm trong phạm trù thư viện điện tử, trong thư viện điện tử có thư viện ảo.

Theo tác giả, thư viện ảo thực chất là thư viện điện tử/thư viện số nhưng được đặt tên như vậy nhằm nhấn mạnh tính chất phi không gian mà thôi.

1.2. VAI TRÒ , ĐẶC TÍNH VÀ LOẠI HÌNH THƯ VIỆN SỐ

1.2.1. Vai trò của thư viện số

Với những đặc điểm ưu việt như:

- Mang thông tin đến với người dùng tin mà không phụ thuộc vào khoảng cách và thời gian: Thư viện số có thể sử dụng ở khắp mọi nơi, người dùng tin có thể truy cập và tìm kiếm thông tin theo một cơ chế quản trị quyền truy cập ở mọi nơi, mọi lúc;

- Tăng cường khả năng tìm kiếm thông tin;

- Tăng cường khả năng chia sẻ thông tin: Thư viện số thúc đẩy quá trình chia sẻ, công bố thông tin của các tổ chức với nhau; chia sẻ thông tin giữa các cá nhân, tổ chức với cộng đồng (thông qua việc hình thành các bộ sưu tập cá nhân của các nhà nghiên cứu; thông qua việc chia sẻ các tài liệu học tập của các thành viên trong đơn vị đào tạo ...);

- Tăng cường sự cộng tác: Thư viện số tăng cường sự cộng tác giữa

các bộ phận nghiệp vụ trong cùng một cơ quan thông tin - thư viện (giữa bộ phận bổ sung; bộ phận biên mục, phân loại; bộ phận lưu thông); tăng cường sự cộng tác giữa thủ thư với người dùng tin (trong thư viện số người dùng tin đồng thời đóng vai trò là người sáng tạo, tạo lập thông tin cho thư viện số); tăng cường sự cộng tác giữa các cơ quan thông tin - thư viện thông qua các hoạt động liên kết và chia sẻ nguồn tin ...;

- Giảm khoảng cách số: Sự phát triển của công nghệ thông tin và truyền thông phát triển, đặc biệt Internet và thư viện số, đang làm phẳng thế giới và làm giảm khoảng cách giữa mọi người trên thế giới, mọi người đều có cơ hội tiếp cận thông tin khắp nơi trên thế giới một cách bình đẳng, không phụ thuộc vào vị trí địa lý, thời gian;

Thư viện số đóng vai trò rất quan trọng trong tất cả các lĩnh vực hoạt động chính trị, kinh tế, khoa học, văn hóa, giáo dục ... của mọi quốc gia.

Thư viện số, với chức năng rất quan trọng là cung cấp hệ thống tri thức khoa học đầy đủ và luôn cập nhật những thông tin mới, là bộ phận không thể thiếu của một môi trường giáo dục & đào tạo. Đối với trường đại học, nơi mà khả năng tự học và tự nghiên cứu khoa học của sinh viên được đề cao, thì vai trò của thư viện số càng được khẳng định.

- Đối với hoạt động học tập: bên cạnh giáo trình chính thức bắt buộc phải có, giảng viên còn yêu cầu sinh viên phải tham khảo một số lượng lớn tài liệu khác có liên quan đến môn học để mở mang nhận thức, khi đó thư viện số sẽ là nơi cung cấp đầy đủ cho sinh viên nguồn kiến thức và cả hệ thống máy tính nối mạng để truy cập tìm những tài liệu cần thiết.

- Đối với hoạt động giảng dạy: Không chỉ sinh viên, mà ngay cả giảng viên trước khi giảng dạy họ cũng thường đến thư viện để tham khảo các tài liệu có liên quan, góp phần hoàn thiện hơn giáo trình giảng dạy. Từ đó cho thấy, hoạt động dạy học giữa giảng viên và sinh viên rất cần được sự tiếp sức của thư viện

- Đối với hoạt động nghiên cứu khoa học: Vai trò của thư viện số càng thể hiện rõ ràng hơn. Khi làm bất cứ công trình nghiên cứu nào, người ta đều phải hiểu rõ lịch sử nghiên cứu vấn đề, đối tượng và phạm vi nghiên cứu, phương pháp và phương tiện nghiên cứu... Hệ thống tài liệu tham khảo phong phú mà thư viện số cung cấp, bao gồm cả tài liệu sách in và tài liệu điện tử, sẽ phần nào giúp chúng ta giải quyết được những câu hỏi đó.

Thư viện số và đào tạo từ xa

Cùng với sự phát triển vượt bậc của công nghệ, việc tiến hành đào tạo từ xa (ĐTTX) đã và đang trở nên dễ dàng, thuận tiện hơn nhờ sự trợ giúp của thông tin viễn thông và các chương trình học trên mạng. Khuynh hướng cung cấp các chương trình đào tạo thông qua mạng diện rộng toàn cầu (World Wide Web) của các trường cao đẳng, đại học và cao học, của các trung tâm và viện nghiên cứu đang trở nên phổ biến hơn tại Hoa Kỳ cũng như nhiều quốc gia khác trên thế giới.

Đào tạo từ xa có thể được định nghĩa như một hình thức giáo dục trong đó giảng viên và học viên không “gần gũi” về mặt địa lý, do đó họ phải thông qua các hình thức giao tiếp từ xa để thực hiện quá trình dạy và học.

Đối tượng của ĐTTX là những người không thể tham gia vào các chương trình đào tạo truyền thống được tổ chức ở một địa điểm cố định trong một khoảng thời gian nhất định. Dạy và học dựa trên mạng diện rộng toàn cầu và máy tính, các lớp học ảo và sự hỗ trợ kỹ thuật số là những hình thức giảng dạy đã ra đời để đáp ứng đòi hỏi về một loại hình đào tạo không lệ thuộc vào không gian và thời gian dành cho những đối tượng kể trên.

Mục tiêu của ĐTTX là đưa giáo dục đến với mọi người, thay vì mỗi người tự tìm đến giáo dục. Trong môi trường ĐTTX, không có những giới hạn của không gian và thời gian; học viên và giảng viên có thể tiếp

cận các nguồn tài nguyên, cũng như tương tác và tác động đến các nguồn tài nguyên “mọi lúc, mọi nơi”. Nói một cách khác, thư viện số là trái tim của các tổ chức giáo dục và đào tạo cung cấp chương trình ĐTTX. Học viên không cần phải đến thư viện để truy cập và tìm kiếm tài liệu, mà chính thư viện số đem tài liệu đến người dùng ở bất cứ nơi đâu và trong mọi thời điểm. Có thể khẳng định rằng ĐTTX không thể thực hiện được nếu không có thư viện số.

Thư viện số được xây dựng theo mô hình của một “siêu thị bách hóa” (nơi người mua có thể tìm thấy tất cả mọi thứ cần mua), bao gồm các mạng lưới tri thức, “kho chứa” các tài nguyên điện tử và những công cụ dạy và học ảo cho cộng đồng ĐTTX. Kho chứa của “siêu thị bách hóa” này phải đáp ứng được những yêu cầu của chương trình đào tạo cũng như những nhu cầu của người dạy và học. Vì thế mục tiêu của thư viện số là giúp đỡ và hướng dẫn cộng đồng ĐTTX học tập và nghiên cứu để họ trở thành những người biết lập luận, phân tích những thông tin do chính mình thu thập được và sáng tạo nên tri thức. Người làm công tác thư viện trong thư viện số không còn là người thụ động chờ bạn đọc đến đặt ra những nhu cầu tìm kiếm thông tin. Người làm công tác thư viện trong thư viện số chủ động cộng tác chặt chẽ với giảng viên, trợ giảng và học viên để chủ động hướng dẫn sử dụng thư viện số cũng như cung cấp các dịch vụ tham khảo ảo trên mạng. Dịch vụ tham khảo kỹ thuật số trong môi trường ĐTTX là sự kết hợp của việc hướng dẫn sử dụng nguồn tài nguyên thư viện số và việc cung cấp câu trả lời. Người làm công tác thư viện tích cực phục vụ như những chuyên viên kỹ thuật giúp đỡ cộng đồng người dùng, hướng dẫn giảng viên thiết kế những tài liệu giảng dạy hoặc tư vấn những vấn đề luật bản quyền và quyền sở hữu trí tuệ đối với các tài liệu điện tử.

- ❖ *Thư viện số lưu trữ, “sản xuất” và phổ biến các tài nguyên điện tử*

Thư viện số chọn lọc và lưu trữ các tài nguyên điện tử. Tất cả các loại tài liệu của thư viện truyền thống như sách, bài báo, phim, ảnh chụp, bản nhạc, bản đồ và các loại tài liệu lưu trữ khác đều có thể được số hóa và lưu trữ trong kho chứa của thư viện số. Tài liệu điện tử có thể bao gồm nguồn lưu trữ toàn văn các bài báo, tạp chí do các cơ sở dữ liệu hoặc các báo điện tử cung cấp. Tài nguyên điện tử có thể được tạo thành từ chương trình nhận dạng ký tự khi “quét” văn bản hoặc bài viết để xuất ra các văn bản có thể tìm kiếm được theo từ khóa. Người làm công tác thư viện và các kỹ thuật viên thực hiện các dự án kỹ thuật số, chuyển dạng các ấn phẩm hoặc các loại tài liệu khác của thư viện sang dạng số hóa. Quá trình số hóa các loại tài liệu của thư viện truyền thống bao gồm phân loại, tạo metadata (dữ liệu về các dữ liệu), kiểm tra tình trạng của tài liệu và lập chỉ mục. Sau khi được chọn lọc, xử lý, lưu trữ và bảo quản, bộ sưu tập tài liệu số hóa sẽ được chuyển giao và phổ biến đến bạn đọc thông qua các trang web của thư viện.

Định dạng số hóa, metadata lưu trữ trong thư viện kỹ thuật số có thể được “tái sử dụng” để “sản xuất” ra những “sản phẩm dạy và học điện tử” (e-learning productions), ví dụ như những “gói tài liệu giáo khoa” (course-packages), tài liệu giảng dạy dựa trên ứng dụng web, v.v. Đội ngũ giảng viên ở đây không chỉ là những giảng viên, giáo viên, trợ giảng mà còn bao gồm cả người làm công tác thư viện, kỹ thuật viên tin học. Trong đội ngũ này, mọi người cộng tác chặt chẽ với nhau để số hóa các ấn phẩm, kết nối từng đoạn nhạc, từng chương sách, từng đoạn báo và các hình ảnh kỹ thuật số để tạo thành một văn bản điện tử (e-texts), hoặc các “gói tài liệu” dành cho những sinh viên được đào tạo tại chỗ hoặc học viên ĐTTX.

❖ *Thư viện số phục vụ như những công cụ thông minh*

Thư viện số cung cấp những công cụ và dịch vụ thông qua các trang web hoặc các giao thức của thư viện để giúp cộng đồng người sử dụng

tìm kiếm, “lướt” và đánh giá các nguồn tài nguyên, thao tác thành thạo với nhiều định dạng số hóa khác nhau. Trợ giảng, giảng viên và học viên tái sử dụng nguồn tài nguyên số hóa để tạo nên tài liệu dạy và học, thực hiện công tác nghiên cứu hoặc chỉ đơn giản là làm bài tập. Các công cụ và dịch vụ ảo sẽ giúp nâng cao kỹ năng sử dụng máy tính của người dùng, khuyến khích hoạt động nghiên cứu để sáng tạo nên những sản phẩm tri thức, những sản phẩm này cũng chính là nguồn tài nguyên thích ứng với nhu cầu người dùng. Thêm vào đó, công nghệ thông tin tạo cơ hội cho mọi người trong môi trường ĐTTX tham gia trực tiếp vào quá trình dạy và học cũng như nghiên cứu.

❖ *Thư viện số là thành viên của “cộng đồng học giả”*

Thư viện số của các cơ sở giáo dục và đào tạo đang tích cực cùng làm việc với nhau trong các dự án số hóa, các hiệp hội và các sáng kiến nhằm thu hoạch, thiết kế theo yêu cầu và mở rộng các cơ sở dữ liệu, dữ liệu về các cơ sở dữ liệu (metadata) nhưng còn gọi là siêu dữ liệu của họ. Có thể kể đến các dự án mang tính tiên phong như DSpace Federation (DSpace) của Học viện Kỹ thuật Massachusetts (MIT), Diễn đàn Kho Tư liệu Mở của các nước châu Âu (Open Archives Forum - OAForum) và Các Công trình Nghiên cứu Kinh tế (Research Papers in Economics - RePEc). Thư viện số tham gia các dự án số hóa và các hiệp hội với tư cách là nhà cung cấp dữ liệu hoặc nhà cung cấp dịch vụ để xây dựng và trao đổi tri thức nhằm thúc đẩy sự phát triển và nâng cao chất lượng của ĐTTX.

Thư viện số đóng vai trò quan trọng trong ĐTTX. Thư viện số bảo quản các tài liệu số hóa, cung cấp tài liệu, công cụ và những dịch vụ để tạo nên hình thức học tập dựa trên các nguồn tài nguyên dành cho cộng đồng ĐTTX. Trong môi trường ĐTTX, người làm công tác thư viện, học viên và người tham gia giảng dạy tương tác qua lại trên mạng diện rộng toàn cầu để chia sẻ tài liệu số hóa và xuất bản những sản phẩm tri thức

nhằm mở rộng vốn kiến thức của nhân loại. Thư viện số phục vụ như những công cụ thông minh để cung cấp phương thức xây dựng kiến thức, hỗ trợ quá trình học tập, nghiên cứu và chuyển giao những sản phẩm tri thức vượt qua sự giới hạn của không gian và thời gian không chỉ cho học viên ĐTTX mà còn cho cả cộng đồng người dùng thư viện.

1.2.2. Đặc tính của thư viện số

Theo quan điểm của Liên hiệp Thư viện số Mỹ (ADF) về khái niệm thư viện số - được xác nhận như một định nghĩa khoa học và xác đáng, thì thư viện số có đặc tính sau:

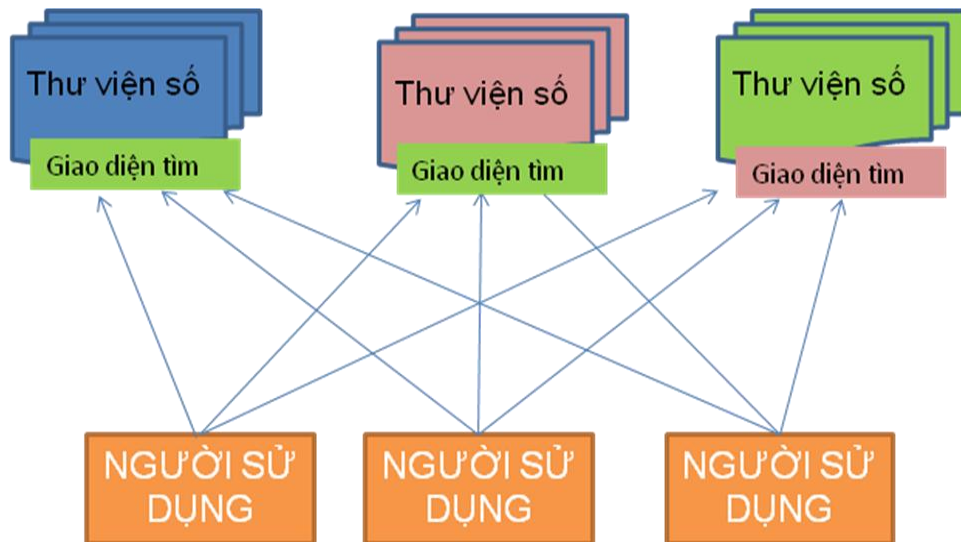
- Thư viện số là cơ quan, tổ chức chứ không phải là “ảo” trên mạng.
- Thư viện số là tập hợp, lưu trữ các sưu tập tài liệu số.
- Nó là phương diện số của thư viện truyền thống mà ở đó lưu trữ các tập tài liệu số và tài liệu truyền thống.
- Thư viện số bao gồm các quá trình và nhiệm vụ nòng cốt và trung tâm của hệ thống thư viện.
- Thư viện số cung cấp những dịch vụ truy cập đến tài liệu số không phụ thuộc vào loại hình khổ mẫu của chúng.
- Do tài liệu dễ bị thay đổi, bị mất cho nên việc bảo quản tài liệu số trong điều kiện môi trường điện tử phân tán và thay đổi nhanh chóng của công nghệ thông tin là công việc khó khăn và phức tạp.
- Các hoạt động của thư viện số phụ thuộc hoàn toàn máy tính và hệ thống mạng.
- Cán bộ thư viện số cần có kỹ năng cao về tin học. [13]

1.2.3. Loại hình thư viện số

Có 3 loại thư viện số là thư viện số độc lập, thư viện số liên kết và thư viện số thu thập.

Thư viện số độc lập (Stand-alone Digital Library)

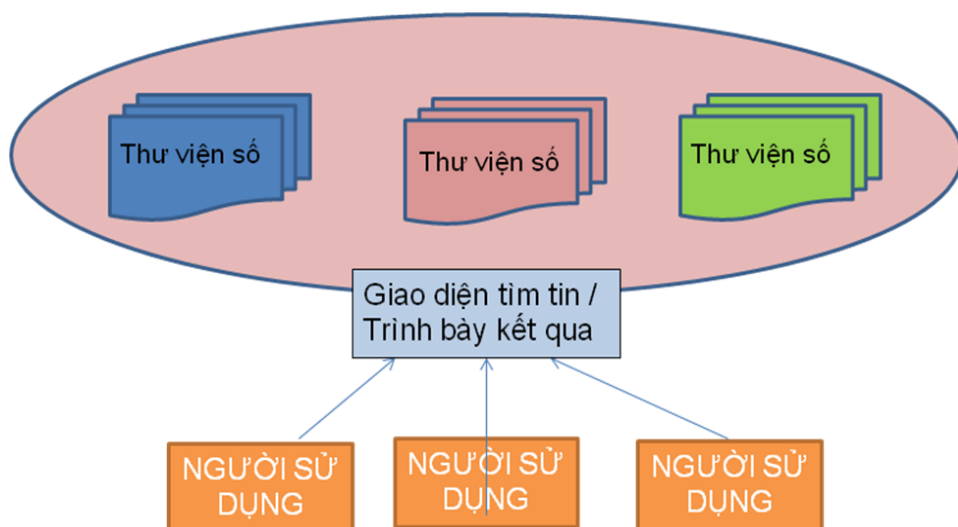
- Là thư viện truyền thống được vận hành hoàn toàn tin học hoá.
- Suu tập thư viện được số hoá
- Tài liệu thư viện được tập trung tại chỗ



Hình 1.3 Thư viện số độc lập

Thư viện số liên kết (Federated Digital Library)

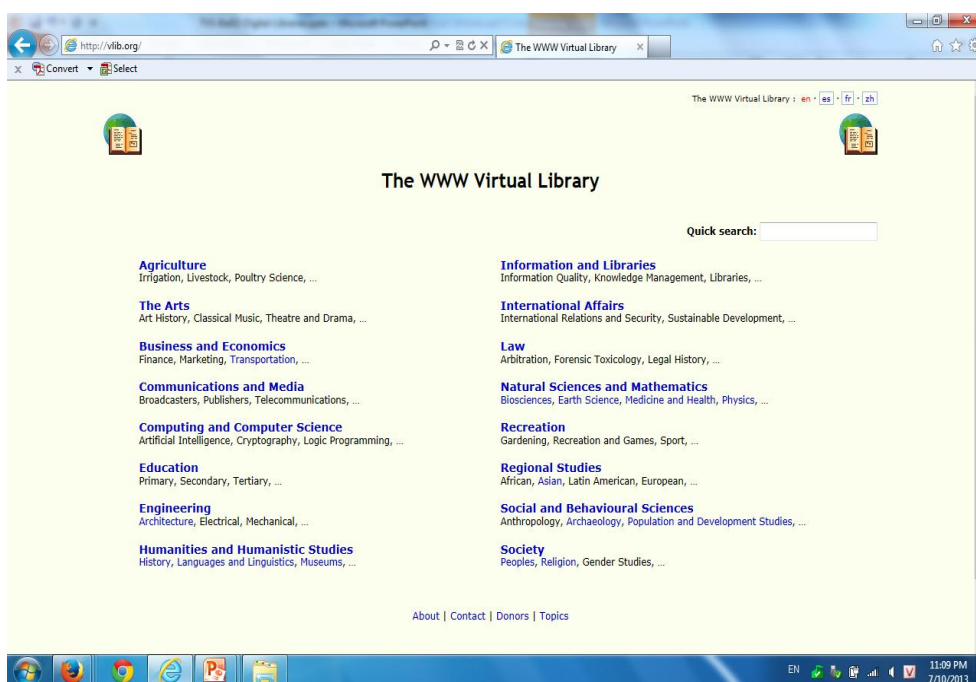
- Là liên hợp của một số thư viện số độc lập trên mạng có cùng chủ đề
- Có giao diện tìm kiếm và trình bày chung
- Người sử dụng chỉ cần vào một nơi để tìm tin



Hình 1.4 Thư viện số Liên kết

Thư viện số thu thập (Harvested Digital Library)

- Tạo khả năng truy cập đến thông tin thư mục của các tài nguyên phân tán trên mạng
- Thông tin được thu thập do cán bộ thông tin chịu trách nhiệm tìm kiếm và thu thập về



Hình 1.5: Thư viện số thu thập

1.3 CÁC CHỨC NĂNG CƠ BẢN VÀ DỊCH VỤ CỦA THƯ VIỆN SỐ

1.3.1. Các chức năng cơ bản

Thư viện số và tất cả các loại hình thư viện hiện đại vẫn luôn đảm bảo các chức năng cơ bản của thư viện truyền thống. Đó là:

- Chức năng văn hóa;
- Chức năng giáo dục;
- Chức năng thông tin;
- Chức năng giải trí.

Ngoài ra, với việc ứng dụng mạnh mẽ thành tựu của công nghệ thông tin và truyền thông, thư viện số còn có những điểm mạnh nổi bật như:

- Giúp cho người dùng có thể với tới được các dạng nguồn tri thức, thông tin
- Tạo cơ chế phát hiện nguồn, cho phép người dùng nhận dạng, xác định được các nguồn tin cần tìm và vị trí lưu giữ các nguồn tin ấy
- Cung cấp cơ chế chuyển giao các nguồn tin riêng biệt tới người dùng, kể cả quá trình nhận các nguồn tin ở nơi khác và chuyển cho người dùng. (Thư viện như là người môi giới thông tin)

1.3.2. Các dịch vụ

- Dịch vụ tàng trữ, nơi các đối tượng số được ký gửi và lưu giữ
- Dịch vụ định danh (đặt tên) đảm bảo cho mỗi đối tượng số có một tên duy nhất và có ít nhất một vị trí lưu trữ
- Dịch vụ chỉ mục: mô tả tập hợp các đối tượng số, chuyển đổi các câu hỏi thành tập hợp kết quả tìm có chứa các tên nguồn duy nhất.
- Dịch vụ thu thập: lựa chọn theo các tiêu chí xác định, dựa vào các mục lục chuyên môn hoá hoặc các phương tiện trợ giúp phát hiện nguồn khác

1.4 CÁC NGUYÊN TẮC CƠ BẢN TIÊU CHÍ ĐÁNH GIÁ VÀ MỘT SỐ VẤN ĐỀ KHI XÂY DỰNG THƯ VIỆN SỐ

1.4.1. Các nguyên tắc cơ bản

Trong khi xây dựng thư viện số, điều quan trọng là phải xem xét các nguyên tắc quan trọng đảm bảo sử dụng dễ dàng thư viện đó và giá trị lưu trữ lâu dài:

- Cần phải có những dạng tư liệu tiêu biểu. Các thành phần tư liệu phải được trình bày dưới hình thức tự nhiên, cụ thể là những đối tượng có thể được vận hành bởi người dùng quen thuộc với chúng.
- Kết hợp cả ba lĩnh vực: Xã hội (kỹ năng và kiến thức thông tin của người sử dụng, ảnh hưởng xã hội đối với dây chuyền chuyển giao thông tin, luật pháp và chính sách), thông tin (tổ chức, phát hiện nguồn, vai trò của siêu dữ liệu,...), và hệ thống (tương tác người - máy, phần

mềm và cấu trúc, qui mô và tương tác)

- Các đường liên kết phải được ghi lại, giữ gìn, tổ chức và tổng quát hoá.

- Cần phải có sự phân tách giữa thư viện số và giao diện người dùng cho thư viện đó. Đối tượng của thư viện số được sử dụng khác với đối tượng được lưu trữ. Người dùng tin cần nội dung trí tuệ của tư liệu chứ không phải là đối tượng số.

- Sử dụng những phương pháp tìm kiếm tiên tiến.

- Phải phát triển các hệ thống mở, bao gồm người dùng và địa điểm nơi mà một số chức năng của cán bộ thư viện sẽ do máy tính thực hiện.

- Phải hỗ trợ việc truy nhập theo nhiệm vụ tới các nguồn lưu trữ điện tử

- Phải có quan điểm phát triển lấy người dùng làm trung tâm. Người dùng phải làm việc với những đối tượng ở mức tổng quát hoá thích hợp.

1.4.2. Tiêu chí đánh giá

Như các loại hình thư viện khác, thư viện số được đánh giá theo các tiêu chí chung trong TCVN 11774:2016, ISO 11620:2014 Thông tin và tư liệu - Bộ chỉ số đánh giá hoạt động thư viện, được Tổng cục Tiêu chuẩn Đo lường Chất lượng thuộc Bộ Khoa học và Công nghệ Việt Nam ban hành năm 2016. Theo đó, thư viện được đánh giá theo 4 nhóm chỉ số: Các nguồn lực, sự truy cập và cơ sở hạ tầng; Sử dụng; Hiệu suất; Tiềm năng và sự phát triển. Tuy nhiên, do đặc thù của thư viện số là sử dụng phần mềm trong toàn bộ hoạt động nên ngoài các tiêu chí chung, căn cứ quan trọng đánh giá thư viện số được dựa trên các tiêu chí đánh giá phần mềm. Có 3 loại phần mềm chính được ứng dụng trong thư viện là phần mềm quản trị thư viện tích hợp; phần mềm tạo lập, quản trị và khai thác tài liệu số và phần mềm tìm kiếm tập trung (theo công nghệ điện toán đám mây).

Năm 2005, nhóm tác giả Tạ Bá Hưng, Nguyễn Điển và Nguyễn Thắng tại Cục Thông tin KH-CN Quốc gia đề xuất “Các tiêu chí đánh giá và lựa chọn phần mềm cho thư viện điện tử ở Việt Nam” gồm 3 nhóm tiêu chí: nhóm tiêu chí về CNTT, nhóm tiêu chí về tiêu chuẩn nghiệp vụ TT-TV và nhóm tiêu chí đối với các phân hệ chức năng. Trong đó nhóm tiêu chí về CNTT là:

- Nguyên tắc thiết kế mở;
- Xây dựng theo mô hình khách/chủ (client/server);
- Làm việc trên mạng TCP/IP;
- Làm việc trong môi trường Web, bằng tiếng Việt và tiếng Anh hoặc một số ngôn ngữ phổ biến khác;
- Xây dựng theo kiến trúc nhiều lớp, hệ thống bao gồm các phân hệ chức năng và ñược tích hợp thành một hệ thống thống nhất;
- Sử dụng hệ quản trị CSDL mô hình quan hệ;
- Máy chủ sử dụng hệ ñiều hành Windows 2000 Advanced Server, Windows NT Server và có thể chuyển sang Linux một cách dễ dàng;
- Máy trạm có thể sử dụng bất cứ hệ ñiều hành nào để hỗ trợ Web;
- Quản trị và giám sát: Cho phép theo ñõi và giám sát ñược mọi hoạt ñộng trên hệ thống;
- An ninh hệ thống: Phần mềm phải hỗ trợ nhiều mức và cơ chế ñảm bảo an ninh hệ thống khác nhau;
- Ngôn ngữ giao diện: Ngôn ngữ sử dụng trên giao diện các phân hệ của chương trình là tiếng Anh và Việt
- Bảng mã lưu trữ dữ liệu trong hệ thống: là Unicode tiêu chuẩn Việt Nam TCVN 6909 với bảng mã ñựng sẵn (pre-

- compound), có bộ chuyển đổi (convertor) để chuyển đổi sang bảng mã tổ hợp (compound) khi cần thiết;
- Bảng mã hiển thị dữ liệu trên giao diện phải hỗ trợ đồng thời các bảng mã TCVN 6909 (dựa trên Unicode); VNI; TCVN 5712;
 - Sắp xếp tiếng Việt: Phần mềm phải có khả năng sắp xếp dữ liệu tiếng Việt theo đúng trật tự từ điển;
 - Vận hành hiệu quả trên CSDL liệu lớn;
 - Khả năng sao lưu/khôi phục dữ liệu: đảm bảo toàn vẹn dữ liệu cho hệ thống;
 - Khả năng mở rộng: Khả năng bổ sung thêm các phân hệ, tính năng, máy trạm và máy chủ với số lượng người dùng không hạn chế;
 - Khả năng tự quản cao: Cài đặt dễ dàng, người dùng có khả năng tự đặt cấu hình cho hệ thống với trợ giúp tối thiểu của nhà cung cấp;
 - Hỗ trợ mã vạch;
 - Khả năng tùy biến cao trong việc tạo khuôn dạng báo cáo dữ liệu;
 - Liên kết với các phần mềm khác như E-mail trong một mạng.

Năm 2017, tác giả Đoàn Phan Tân đề xuất 2 loại tiêu chí quan trọng khi đánh giá phần mềm ứng dụng trong hoạt động thông tin – thư viện là yêu cầu về tính năng và yêu cầu về kỹ thuật – công nghệ.

Yêu cầu về tính năng gồm :

- Phần mềm phải đáp ứng đầy đủ các yêu cầu ứng dụng mà hệ thống thông tin đặt ra;
- Có khả năng xử lý dữ liệu, cập nhật dữ liệu dễ dàng;
- Có khả năng tiếp nhận dữ liệu với nhiều định dạng khác nhau;

- Có công cụ tìm kiếm thông tin mạnh;
- Có khả năng kết xuất thông tin, đáp ứng đầy đủ yêu cầu thông tin của người sử dụng;
- Có khả năng đảm bảo làm việc ổn định với dữ liệu lớn;
- Có khả năng tương tác với các hệ thống khác;
- Giao diện người dùng thân thiện, thẩm mỹ và dễ sử dụng và có khả năng tùy biến cao.

Yêu cầu về kỹ thuật – công nghệ gồm:

- Phần mềm phải được thiết kế trên nền tảng công nghệ thông tin và truyền thông hiện đại;
- Đáp ứng đầy đủ các chuẩn các chuẩn nghiệp vụ của lĩnh vực được ứng dụng;
- Có thể cài đặt và chạy được trên các hệ điều hành thông dụng Windows XP, Windows 7, Linux, Unix;
- Vận hành với giao diện Web trên môi trường Internet;
- Cơ sở dữ liệu được thiết kế theo mô hình dữ liệu tiên tiến và được quản lý bởi các hệ quản trị cơ sở dữ liệu có uy tín;
- Có khả năng tương tác với phần mềm khác;
- Có khả năng phát triển mở rộng tính năng chương trình theo yêu cầu.

1.4.3. Một số vấn đề khi xây dựng thư viện số

Nỗ lực của các nhà tin học (thông qua các phần mềm tương tác và hệ chuyên gia có khả năng mô phỏng các dịch vụ do con người thực hiện) và chất lượng xử lý thông tin cao đã giúp người sử dụng bớt được một số khó khăn thường gặp khi sử dụng các loại thư viện số như:

- Khó tìm do thiếu công cụ hoặc do bộ máy tìm kiếm tổ chức kém.
- Thiếu các tham chiếu qua lại và mối liên kết với các tư liệu

khác.

- Quá nhiều đường liên kết vu vơ hoặc dẫn tới các thông tin vô dụng.
- Thường xuyên cải tổ khiến người dùng nhiều khi phải phán đoán mò mẫm nơi có các thông tin đã được định vị trước đây.
- Thiếu sự nhất quán trong khi trình bày những thông tin tương tự.
- Thông tin lạc hậu, không cập nhật, sai ngữ pháp và chính tả.
- Khổ mẫu không tương hợp nên gặp rắc rối khi tham chiếu trực tuyến và in ra.

Tuy nhiên, quá trình xây dựng thư viện số còn gặp nhiều thách thức trên các khía cạnh khác nhau. Đó là:

Về phương diện kinh tế

Bước đầu tiên và khó khăn nhất trong việc phát triển thư viện số là chuyển đổi các xuất bản phẩm truyền thống trên giấy sang dạng số. Vấn đề giá thành / hiệu quả của quá trình số hoá khó thuyết phục do chi phí công nghệ và sự cần thiết phải đầu tư liên tục vào công nghệ mới: Số hoá một trang tài liệu giá thành khoảng 0,1 - 0,5 đôla (cách đây 10 năm là 2 đôla). Một thư viện lớn có hàng triệu đơn vị tài liệu khó có thể số hoá toàn bộ kho tài liệu. Thiết bị cũng rất tốn kém vì các dự án đòi hỏi những siêu máy tính có trang bị các bộ vi xử lý cực mạnh, bộ nhớ và khối lượng lưu trữ trong ổ cứng cực lớn. Kể cả trường hợp chuyển đổi số hoá toàn bộ ấn phẩm vào thư viện (số) truy nhập toàn cầu, thì chắc chắn phần lớn các tài liệu này sẽ lại in ra từ máy để đọc trên giấy.

Về phương diện kỹ thuật

Mặc dầu vấn đề chất lượng sao lại chính xác dữ liệu trong quá trình số hoá đã từng đặt ra và đến nay hầu như đã được giải quyết được bằng các máy quét hiện đại chất lượng cao, nhưng vấn đề vật liệu lưu trữ vẫn không thể bền được bằng các giấy không có axit.

Về phương diện con người

Một thực tế không ai phủ nhận được là khả năng áp dụng các tiếp cận công nghệ thông tin của cán bộ là công tác thông tin còn rất hạn chế. Không những thế những trang thiết bị hiện đại các phương thức hoạt động mới trong hoạt động thư viện không được cập nhật. Chính vì vậy đây là một rào cản lớn khi thư viện Việt Nam tiến hành xây dựng thư viện số, thư viện điện tử.

Vấn đề bản quyền

Bản quyền là một trở ngại đối với việc phát triển thư viện số, bởi vì thư viện số bị ràng buộc bởi những điều khoản của luật bản quyền có liên quan đến việc xuất bản lại các tư liệu dưới hình thức mới, không có phép. Một thư viện phải dung hoà giữa quyền sở hữu trí tuệ và quyền lợi của công chúng. Một mặt tính chất công cộng của các thư viện xuất phát từ nguyên tắc phục vụ không lấy tiền và không vụ lợi và bình đẳng trong truy nhập thông tin và tri thức. Khi chuyển sang thư viện số thì chức năng này cũng không thể thay đổi. Mặt khác, truy nhập toàn cầu tới thư viện số là điều kiện tiên quyết để phát triển thị trường số, như vậy việc truy nhập công cộng miễn phí khó có thể dung hoà với cơ chế thị trường của việc phổ biến thông tin trong một nền kinh tế mà thông tin được coi là hàng hoá.

1.5. THƯ VIỆN SỐ VÀ HỌC LIỆU MỞ, TRUY CẬP MỞ

1.5.1. Khái quát về học liệu mở và tài nguyên giáo dục mở

1.5.1.1. Khái niệm

Thuật ngữ Học liệu mở (OpenCourseWare - OCW) được Viện Công nghệ Massachusetts - MIT (Mỹ) khai sinh vào năm 2002 khi MIT quyết định đưa toàn bộ nội dung giảng dạy của mình lên web và cho phép người dùng Internet ở mọi nơi trên thế giới truy nhập hoàn toàn miễn phí. Hiện nay trang web về học liệu mở của MIT có trên 2000 môn học (course) bao gồm bài giảng, lịch học, danh mục tài liệu tham khảo, bài

tập về nhà, bài thi, bài thí nghiệm để người dùng có thể tham khảo cho việc giảng dạy, học tập và nghiên cứu của mình.

Với tiêu chí “Tri thức là của chung của nhân loại và tri thức cần phải được chia sẻ”, rất nhiều trường đại học và viện nghiên cứu trên thế giới đã tham gia phong trào học liệu mở và lập lên Hiệp hội Học liệu mở (OpenCourseWare Consortium) để chia sẻ nội dung, công cụ cũng như phương thức triển khai học liệu mở sao cho đạt được hiệu quả cao nhất. Giảng viên, sinh viên và người tự học ở mọi nơi trên thế giới, đặc biệt là từ các nước đang phát triển như Việt Nam, đều có cơ hội như nhau trong việc tiếp cận các tri thức mới.

Tài nguyên giáo dục mở (Open Educational Resources - OER) là các tài liệu học tập được sử dụng linh hoạt, trao đổi một cách tự do. Thuật ngữ Tài nguyên giáo dục mở lần đầu tiên được thông qua tại Diễn đàn UNESCO năm 2002 về Tác động của học liệu mở đối với giáo dục đại học ở các nước đang phát triển. Tài nguyên giáo dục mở bao gồm:

- Nội dung học: Các khóa học đầy đủ, giáo trình, bài giảng, các tài liệu học tập, các module nội dung, các đối tượng học tập, tuyển tập, tạp chí...
- Công cụ: Phần mềm hỗ trợ việc tạo lập, phân phối, sử dụng và cải thiện nội dung học tập mở bao gồm cả việc tìm kiếm và tổ chức nội dung, các hệ thống quản lý nội dung và quá trình học tập, các công cụ phát triển nội dung, và các cộng đồng học tập trực tuyến.
- Các tài nguyên bổ sung khác: Các giấy phép sở hữu trí tuệ để thúc đẩy xuất bản các tài liệu mở, các nguyên tắc thiết kế, và việc bản địa hóa nội dung.

1.5.1.2. Chương trình học liệu mở Việt Nam

Chương trình Học liệu mở Việt Nam ra đời vào tháng 11/2005 với sự hợp tác giữa Bộ Giáo dục và Đào tạo, Công ty Phần mềm và Truyền thông VASC, và Quỹ Giáo dục Việt Nam. Mục tiêu của chương trình là xây dựng các phương thức để xoá bỏ các rào cản đối với người dùng Việt Nam để có thể tận dụng một cách tối đa các nguồn học liệu mở sẵn có. Ngày 12/12/2007, trang tin chính thức của chương trình, website www.vocw.edu.vn đã được bấm nút khai trương tại Hà Nội, Đà Nẵng và Thành phố Hồ Chí Minh

Hiện nay, website VOCW đã có hơn 200 course và hơn 1.000 modules. Chủ yếu là do cán bộ thuộc các trường chủ động đưa lên, phần còn lại có được thông qua các hoạt động tài trợ và chuyển đổi các kho giáo trình đã có, cụ thể như sau:

- 24 khóa học mẫu do VEF tài trợ

Trong giai đoạn thử nghiệm (2006-2008), Quỹ Giáo dục Việt Nam đã tài trợ kinh phí để xây dựng nội dung cho chương trình học liệu mở Việt Nam. Hơn ba mươi chuyên gia Việt Nam là giảng viên từ các trường đại học và viện nghiên cứu trong cả nước đã được mời tham gia xây dựng nội dung cho 24 môn học.

Một số môn học được xuất bản và cập nhật hàng ngày trên VOCW tại địa chỉ: <http://www.vocw.edu.vn/lenses/vef/vef>. Những nội dung này hiện đã sẵn sàng để các giảng viên và sinh viên tham khảo trong giảng dạy, học tập và nghiên cứu.

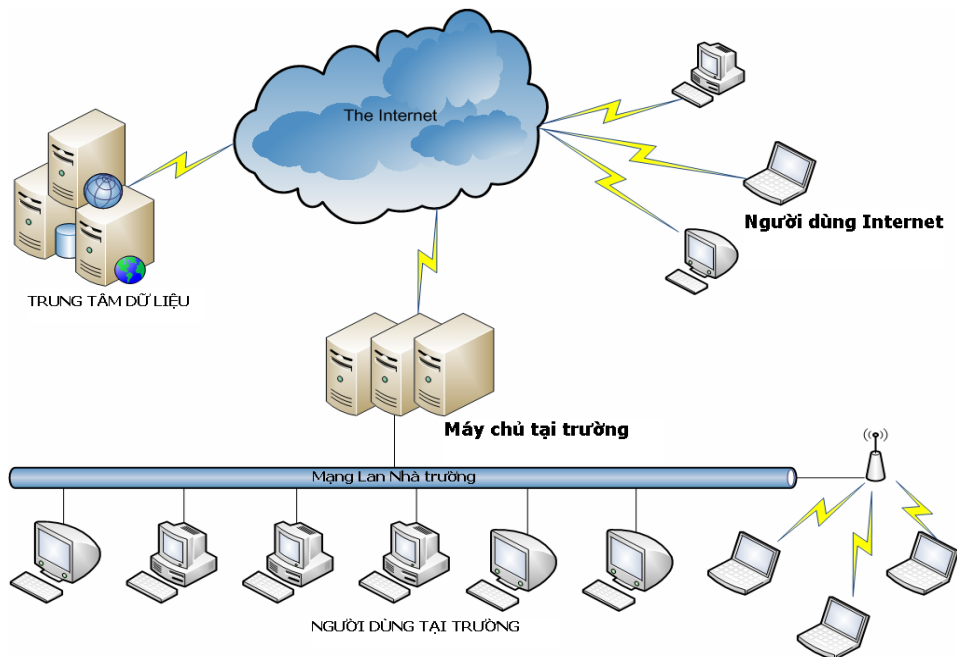
- Hơn 100 giáo trình chuyển đổi từ kho giáo trình điện tử của Bộ GD&ĐT

Với mục đích đóng góp cho trang tin điện tử Học liệu mở Việt Nam các giáo trình tiếng Việt có nội dung phong phú và có thể được truy cập, chia sẻ bởi các giảng viên, nghiên cứu viên, sinh viên và người tự học, Bộ giáo dục và Đào tạo đã đồng ý bước đầu chuyển đổi hơn 100 giáo trình

trong kho giáo trình sẵn có của mình hiện ở các định dạng MS Word, PDF,... sang định dạng của hệ thống VOCW. Số lượng, hình thức và nội dung các giáo trình sẽ liên tục được cập nhật tại địa chỉ <http://www.vocw.edu.vn/lenses/moet/> và sẽ là địa chỉ tin cậy để các Thầy cô, các bạn Sinh viên có thể tham khảo và tái sử dụng.

- Kho học liệu cập nhật từ MIT OpenCourseware

Học viện công nghệ Masachuset MIT vừa công bố con số giáo trình đăng tải hiện có trên website www.mit.edu là 1.800. Theo các hoạt động hợp tác thường niên giữa VOCW và MIT, một đĩa cứng chứa toàn bộ số bài giảng này sẽ được chuyển tới Việt Nam để host trên Server Việt Nam. Số giáo trình này sẽ được đăng tải trên máy chủ VOCW nhằm giúp người dùng tại Việt Nam nhanh chóng truy cập và sử dụng.



Hình 1.5 Mô hình máy chủ địa phương VOCW đặt tại mạng LAN của các cơ quan, tổ chức

Như vậy, với nguồn thông tin và mô hình kết nối như trên, việc tham gia vào Hệ thống học liệu mở Việt Nam, các thư viện sẽ giúp độc giả của mình tiếp cận được một nguồn lực thông tin vô cùng to lớn và tuân thủ

chặt chẽ các nguyên tắc học thuật nhằm hỗ trợ độc giả tiếp cận nhanh, dễ dàng và hiệu quả cao

1.5.2. Truy cập mở (Open Access - OA)

1.5.2.1. Khái niệm

Có nhiều định nghĩa khác nhau về tài liệu truy cập mở. Peter Suber cho rằng “Tài liệu truy cập mở là tài liệu ở dạng số, trực tuyến và không bị ràng buộc bởi hầu hết những hạn chế về bản quyền và cấp phép sử dụng”. [38]

Thư viện Khoa học Công cộng (The Public Library of Science) định nghĩa truy cập mở là “truy cập không bị hạn chế và sử dụng lại không hạn chế”, “sử dụng miễn phí và không hạn chế”. [19]

Năm 2002, Open Society Institute họp tại Budapest và đưa ra *Sáng kiến Budapest về Truy cập mở*. Tiếp đó Năm 2003 có 2 tuyên bố là *Tuyên bố Bethesda về Xuất bản truy cập mở* (Bethesda Statement on Open Access Publishing) và *Tuyên bố Berlin về Truy cập mở đến tri thức khoa học và nhân văn* (Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities) với nội dung như sau: “Có nhiều cấp độ và loại truy cập mở tài liệu rộng hơn và dễ dàng hơn. Khi nói ‘Truy cập mở’ tài liệu này, nghĩa là nói tài liệu đó được cung cấp miễn phí trên Internet công cộng, cho phép tất cả người dùng đọc, tải, sao chép, phân phát, in, tìm kiếm hoặc kết nối đến bài toàn văn của các bài viết này, quét các tài liệu này để tạo chỉ mục, chuyển chúng thành dữ liệu cho phần mềm hoặc sử dụng chúng cho các mục đích hợp pháp khác mà không bị các rào cản về tài chính, luật pháp hoặc kỹ thuật. Vấn đề duy nhất về tái bản và phân phối và vai trò đối với bản quyền trong lĩnh vực này nên để tác giả kiểm soát sự toàn vẹn tác phẩm của họ và quyền hợp pháp được biết và được trích dẫn”.

Nói một cách ngắn gọn, truy cập mở là truy cập miễn phí tới các kết quả nghiên cứu và các dữ liệu trên Internet mà không có bất cứ một yêu

cầu nào về tài chính, pháp lý và các rào cản kỹ thuật hay công nghệ. Việc cho phép người dùng tự do truy cập giúp cho tác giả phổ biến một cách rộng rãi, nhanh chóng và kiểm soát được việc đạo văn, đồng thời tăng số trích dẫn công trình khoa học của mình. Các tài liệu truy cập mở thường tập trung vào các bài báo khoa học, báo cáo hội thảo, kết quả nghiên cứu khoa học, luận văn, luận án ...

1.5.2.2. Mô hình truy cập mở

Hiện tại, có 2 loại mô hình truy cập mở chủ yếu là Tạp chí truy cập mở (OA journals) và Truy cập mở thông qua tự lưu trữ (Open Access via self-archiving).

- a) Tạp chí truy cập mở gồm 3 phương thức truy cập:
 - Truy cập mở hoàn toàn (Full Open Access). Phương thức này cho phép truy cập đến toàn bộ nội dung của tạp chí ngay sau khi được xuất bản trực tuyến mà không mất bất cứ chi phí nào. Tuy nhiên, việc truy cập phải tuân thủ những quy định được ghi trên giấy phép do tác giả gán cho tài liệu.
 - Truy cập mở lai (Hybrid Open Access): Cho phép truy cập miễn phí ngay lập tức đến một phần của nội dung tạp chí. Phương thức này còn được gọi là truy cập mở lựa chọn.
 - Truy cập mở trì hoãn (Delayed Open Access). Cho phép truy cập miễn phí đến nội dung tạp chí, nhưng không phải là ngay lập tức sau khi xuất bản mà sau một thời gian trì hoãn nhất định. Mục đích chính của việc áp dụng phương thức này nhằm đảm bảo quyền lợi cho người mua tạp chí.
- b) Kho lưu trữ truy cập mở (OA repositories) là những sưu tập trực tuyến hoặc CSDL các bài báo khoa học do tác giả tự lưu trữ những bài viết hay tài liệu của mình vào kho và cho phép truy cập mở tới các tài liệu chứa trong đó. Kho lưu trữ mở cho

phép chứa đựng các bản thảo cả trước và sau khi được bình duyệt (“Preprint và postprint”), nhưng chúng luôn luôn sẵn có, miễn phí và vẫn có thể được phản biện trực tiếp bởi cộng đồng người sử dụng.

1.5.2.3. Xuất bản truy cập mở

Một xuất bản phẩm truy cập mở là một xuất bản phẩm đáp ứng hai điều kiện sau:

- Tác giả và người giữ bản quyền cho phép tất cả người dùng quyền truy cập vĩnh viễn, trên toàn thế giới, không thể huỷ bỏ và miễn phí, cho phép sao chép, sử dụng, phân phát, chuyển và giới thiệu tác phẩm đó tới công chúng, tạo và phân phát các tác phẩm phát sinh, trong bất kỳ môi trường số cho bất kỳ mục đích có trách nhiệm nào, quyền tác giả, cũng như quyền sao chụp với khối lượng nhỏ cho mục đích sử dụng cá nhân.
- Một phiên bản tác phẩm hoàn chỉnh và toàn bộ tài liệu bổ sung gồm bản sao giấy phép nêu trên, dưới dạng điện tử theo tiêu chuẩn phù hợp được gửi ngay khi xuất bản lần đầu tới ít nhất một kho trực tuyến. Việc này được sự hỗ trợ của một học viện, tổ chức khoa học, cơ quan chính phủ hoặc một tổ chức khác có uy tín đang tìm kiếm khả năng truy cập mở, phân phát không hạn chế, hoạt động xuyên suốt và lưu trữ lâu dài.

1.5.2.4. Cấp phép truy cập mở

Giấy phép truy cập mở (giấy phép mở) do tổ chức Tài sản sáng tạo công cộng (Creative Commons - CC) là một tổ chức phi lợi nhuận quốc tế đề xuất với mục đích thúc đẩy, phát triển số lượng các tác phẩm sáng tạo được đưa ra cho công chúng sử dụng tự do, sáng tạo lại hoặc chia sẻ. Đồng thời nhằm bảo vệ quyền tác giả trong môi trường mở khi tài liệu được số hóa có thể được sao chép và chia sẻ dễ dàng mà không cần

xin phép. Giấy phép tạo điều kiện cho sự sao chép và chia sẻ được thực hiện trong khuôn khổ pháp lý linh hoạt hơn, với hình thức được lưu giữ theo bản quyền tự động. Những giấy phép này giúp cho người sáng tạo sớm trao quyền cho những người khác để sử dụng tác phẩm theo những điều kiện nhất định.

Các tác phẩm được cấp phép theo các điều khoản và điều kiện của giấy phép mở, được bảo hộ theo luật bản quyền tác giả và/hoặc các luật áp dụng khác. Bất kỳ việc sử dụng tác phẩm nào ngoài phạm vi được phép theo giấy phép mở hoặc theo luật bản quyền tác giả đều bị nghiêm cấm.

Việc thực hiện bất kỳ quyền nào quy định tại giấy phép mở đối với tác phẩm, đã được chấp thuận bởi các bên liên quan đều bị ràng buộc bởi các điều khoản của giấy phép. Bên cấp phép trao cho bên liên quan các quyền được nêu trong giấy phép này trên cơ sở sự chấp thuận của đối tác đối với các điều khoản và điều kiện đã được quy định.

Việc cấp phép mở cũng yêu cầu những người sử dụng tác phẩm ghi nhận công của người sáng tạo. Điều này có nghĩa là mọi người phải luôn luôn nhận thức được người sáng tạo tác phẩm mà mình đang sử dụng, cũng như đưa ra bất kỳ thông tin phù hợp nào về bản quyền. Nội dung cấp phép phải thể hiện được các yếu tố ghi công, phi thương mại, chia sẻ chứa đựng đầy đủ những nội dung như: tác phẩm phái sinh, tuyển chọn, phân phối, các yếu tố cấp phép, bên cấp phép, tác giả, chủ sở hữu quyền liên quan, tác phẩm, truyền đạt đến công chúng, sao chép, chuyển nhượng ảo, chấm dứt ...

Các hình thức phổ biến trong giấy phép mở của Tổ chức Tài sản sáng tạo công cộng là:

- Giấy phép ghi công **CC BY**



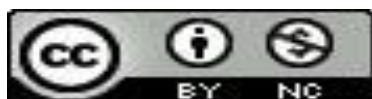
- Giấy phép ghi công - Chia sẻ tương tự **CC BY - SA**



- Giấy phép ghi công - Không có tác phẩm phái sinh **CC BY-ND**;



- Giấy phép ghi công - Phi thương mại **CC BY-NC**;



- Giấy phép ghi công - Phi thương mại - Chia sẻ tương tự **CC BY-NC-SA**



- Giấy phép ghi công - Phi thương mại - Không có tác phẩm phái sinh **CC BY-NC-ND**



1.6. TÌNH HÌNH PHÁT TRIỂN THƯ VIỆN SỐ TRÊN THẾ GIỚI

1.6.1. Ý tưởng và tiền đề xây dựng thư viện số

Thư viện số có một lịch sử khá lâu đời trên thế giới đặc biệt ở các nước có nền khoa học, mà cụ thể là khoa học thông tin thư viện, đã có rất nhiều dự án nhằm phát triển thư viện số trên thế giới.

Năm 1954, trong dự án MEMEX, Vannevar Bush đã có ý tưởng xây dựng hệ thống lưu trữ hồ sơ, quản lý tư liệu tự động, kích thước lưu trữ nhỏ. Có 5 tính năng sau:

- Sử dụng công nghệ hiện đại lưu trữ tìm kiếm tạo ra dấu vết thông tin;
- Tạo ra dấu vết thông tin

- Tạo chú giải điện tử;
- Có kích thước để bàn nhưng có khả năng lưu trữ bằng cả thư viện đại học;
- Có khả năng chia sẻ và trao đổi thông tin.

Vào những năm 50 của thế kỷ XX, dự án INTREX nhằm lưu trữ tài liệu trên vi phim có kèm hệ thống biên mục và định chỉ số trực tuyến, làm thử nghiệm và thành công.

Từ năm 1970-1980, dự án STAIRS (Storage and information Retrieval Systems) của hãng IBM, có khả năng lưu trữ thông tin điện tử và tra cứu thư mục trực tuyến. Hệ thống này phát triển mạnh nhưng chỉ lưu trữ văn bản không lưu trữ hình ảnh.

1.6.2. Quá trình phát triển thư viện số

Năm 1990, với sự phát triển mạnh mẽ của công nghệ thông tin và công nghệ truyền thông, các công nghệ mới xuất hiện tạo khả năng lưu trữ lớn, đa dạng thông tin và tạo cơ hội phát triển thư viện số. Năm 1994, Quốc hội Mỹ thông qua dự án National Digital Library. Quỹ khoa học quốc gia Hoa Kỳ đưa ra sáng kiến Thư viện số pha 1 (Digital library Innitiative) với kinh phí 24 triệu \$ cho 4 năm để triển khai 6 dự án thư viện số ở 6 trường đại học, Pha 2 từ 8-10 triệu \$ trong 4 năm kế tiếp

Gần đây một dự án xây dựng thư viện số thế giới đã được tiến hành, Google hợp tác với Thư viện Quốc hội Mỹ (LOC) để tiến hành dự án xây dựng thư viện số thế giới (World Digital Library). Dự án có tham vọng số hóa và đưa lên web toàn bộ tài nguyên của các thư viện nước Mỹ cùng các viện nghiên cứu khác trên toàn cầu. James H. Billington, người quản lý thư viện cho biết: World Digital Library được xây dựng dựa trên bộ sưu tập American Memory. Bắt đầu từ năm 1994 cho đến nay, American Memory đã số hóa và đưa lên www.loc.gov/memory/ hơn 10 triệu danh mục, bao gồm cả bản viết tay của các vĩ nhân như Abraham

Lincoln, Thomas Jefferson, những bức ảnh về cuộc chiến tranh Nam Bắc Mỹ, cho đến những tư liệu giản dị về cuộc sống đời thường, tạo nên một bộ sưu tập khổng lồ về nền văn hóa Mỹ.

World Digital Library đã nhận được sự đồng thuận đầu tiên từ Thư viện Quốc gia Ai Cập, cho phép số hóa các tài liệu khoa học Hồi giáo từ thế kỉ 10 đến thế kỉ 16. James H. Billington hi vọng Trung Quốc, Đông Á, Ấn Độ, Nam Á và các quốc gia Hồi giáo khác cũng sẽ ủng hộ dự án này.

Ở những trung tâm khoa học lớn trên thế giới vấn đề thư viện số là một trong những chiến lược hàng đầu trong lĩnh vực thông tin thư viện, tiêu biểu như ở Trung Quốc, Mỹ, Anh và rất nhiều quốc gia khác có nền khoa học kỹ thuật phát triển

1.7 TÌNH HÌNH PHÁT TRIỂN THƯ VIỆN SỐ Ở VIỆT NAM

1.7.1. Chính sách phát triển thư viện số ở Việt Nam

Ngành Thư viện Việt Nam đã có một lịch sử phát triển khá lâu đời song ý tưởng số hoá tài liệu, hay xây dựng thư viện điện tử/ thư viện số thì vẫn còn là một trở ngại trước mắt, một mục tiêu phấn đấu của ngành thư viện Việt Nam. Trong khi đó trên thế giới các dự án, các kế hoạch về phát triển thư viện số đã được triển khai trên một quy mô rộng với sự đầu tư khá kỹ càng về ngân sách. Nếu đứng trên bình diện về các kỹ năng và sản phẩm thư viện, Việt Nam cũng đã có sự thua thiệt khá xa với các nước bạn, và đặc biệt về thư viện số thì đây thực sự là một khoảng cách quá lớn đòi hỏi nhà nước ta phải có sự quan tâm thích đáng và đầu tư có hiệu quả, tập trung triển khai thí điểm tại các thư viện lớn đầu ngành trong cả nước. Ý thức rõ được điều này trong chiến lược phát triển thư viện đến năm 2020 của Bộ Văn hoá Thông tin (nay là Bộ Văn hoá, Thể thao và Du lịch) đã chỉ rõ trong quyết định phê duyệt số 10/2007/QĐ-BVHTT ngày 4/5/2007 của Bộ trưởng Bộ Văn hoá Thông tin Lê Doãn Hợp: “Ứng dụng khoa học công nghệ cao nhằm tự động

hoá, hiện đại hoá trong khâu hoạt động của thư viện điện tử và thư viện kỹ thuật số”.

Quy hoạch cũng liệt kê một số dự án đầu tư phát triển sự nghiệp thư viện số:

- Dự án điện tử hóa thư viện công cộng
- Dự án số hóa tài liệu Việt Nam (tiếng Việt và tiếng dân tộc có chữ viết)

Qua các chính sách, quyết định trên chúng ta có thể dễ dàng nhận thấy sự quan tâm của Đảng và Nhà nước trong việc xây dựng hệ thống thông tin thư viện, đặc biệt là thư viện số ở Việt Nam. Điều này mở ra một cơ hội mới cho các thư viện Việt Nam trong công tác số hóa tài liệu, đưa Việt Nam vào một kỷ nguyên mới: Kỷ nguyên số.

1.7.2. Xây dựng thư viện số trong các trường đại học

Trước nhu cầu ngày càng cao của bạn đọc đối với các sản phẩm và dịch vụ thư viện trong Kỷ nguyên số, đồng thời, sự phát triển mạnh mẽ của công nghệ thông tin và truyền thông đã tạo điều kiện thuận lợi cho các thư viện Việt Nam xây dựng và phát triển thư viện số. Ngoài các cơ quan thông tin thư viện quy mô quốc gia như Thư viện Quốc gia Việt Nam, Cục Thông tin Khoa học và Công nghệ Quốc gia, Trung tâm Thông tin – Tư liệu thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam, Viện Thông tin Khoa học Xã hội thuộc Viện Hàn lâm Khoa học Xã hội Việt Nam, hệ thống thư viện đại học là hệ thống thư viện triển khai nhanh chóng và chất lượng tốt nhất mô hình thư viện số.

Nghị định của Chính phủ số 11/2014/NĐ-CP ban hành ngày 18/2/2014 về hoạt động thông tin KH&CN đã xác định các tổ chức thông tin thư viện đại học là một bộ phận cấu thành hệ thống thông tin KH&CN Quốc gia. Nghị định cũng xác định chức năng chính của các tổ chức trong Mạng thông tin nghiên cứu và đào tạo quốc gia, trong đó có các thư

viện đại học. Có thể sơ bộ xác định những chức năng, nhiệm vụ này như sau:

- Là cơ quan thực hiện việc tổ chức và phát triển thư viện trong trường đại học. Tổ chức và phát triển thư viện KH&CN hướng đến việc hình thành trên phạm vi quốc gia mạng thông tin thư viện đại học¹;
- Thực hiện chức năng đầu mối của trường đại học trong việc tổ chức, xây dựng và quản lí các nguồn tin KH&CN (thành phần của Cơ sở dữ liệu quốc gia về khoa học và công nghệ) nhằm đáp ứng nhu cầu thông tin được hình thành trong các hoạt động đào tạo, nghiên cứu và quản lí của trường đại học²;
- Thực hiện công tác tuyên truyền phổ biến thông tin phục vụ các nhiệm vụ và yêu cầu đặt ra trong quá trình phát triển của trường đại học;
- Đáp ứng các nhu cầu trao đổi thông tin dưới mọi hình thức (hội nghị, hội thảo, trình diễn, triển lãm, các dịch vụ trên mạng...);
- Xây dựng và thực hiện các kế hoạch dài hạn và ngắn hạn, các chương trình, đề tài, đề án, dự án phát triển tổ chức và hoạt động thông tin trong trường đại học;
- Quản lí và phát triển ứng dụng CNTT trong hoạt động thông tin;

¹ Từ năm 1998 ở Trung Quốc đã hình thành Hệ thống thông tin thư viện đại học (China Academic Library and Information System-CALIS), và cùng với nó là mạng Nghiên cứu giáo dục (CERNET) trong khuôn khổ Dự án 211 – Dự án phát triển giáo dục đại học quốc gia cho thập kỉ đầu tiên (1) của thế kỉ 21 (21).

² Trong nội dung này, cần đặc biệt chú ý đến vấn đề nguồn tin nội sinh. Tại đây, cơ quan thông tin thư viện đại học cần chú trọng thực hiện chức năng là cơ quan thu nhận, lưu trữ, bảo quản và tổ chức phục vụ theo quy định chung nguồn tin nội sinh (luận án, luận văn khoa học, tư liệu điền dã, kỷ yếu hội nghị, hội thảo khoa học, kết quả nghiên cứu và phát triển, các chương trình điều tra và điều tra cơ bản được thực hiện bằng nguồn ngân sách do Nhà nước cấp hoặc do các tổ chức bên ngoài tài trợ,...) của trường đại học.

- Điều hoà, phối hợp hoạt động thông tin thư viện giữa các cơ quan hữu quan ở trong và ngoài trường đại học, đặc biệt đối với các đại học đa ngành, lớn.

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 sưu tập số 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 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

Với những ưu điểm nổi bật như đã phân tích, thư viện điện tử, thư viện số và tài nguyên số không còn là xu thế mà đã trở thành nhu cầu cấp bách đối với toàn hệ thống TVĐH nhằm cung cấp cho bạn đọc khả năng tiếp cận nhiều sản phẩm và dịch vụ hiện đại, truy cập các CSDL đa dạng, phong phú trong nước và quốc tế. Các thư viện đại học, bằng nhiều nguồn kinh phí khác nhau, từ các dự án cấp quốc gia, từ nguồn tài trợ bên ngoài hay từ vốn ngân sách được cấp hàng năm... đều chú trọng xây dựng và

phát triển TNS, đều coi đây là một trong những nhiệm vụ quan trọng cần tiến hành thường xuyên và có định hướng lâu dài. Tuy nhiên, do những khó khăn về chủ trương, chính sách chung và đặc thù của từng trường đại học; khả năng về tài chính, cơ sở vật chất; trình độ, năng lực của cán bộ thư viện... công tác phát triển TNS trong hệ thống TVĐH Việt Nam còn nhiều bất cập.

Về phần mềm tạo lập, quản trị, khai thác tài nguyên số

Trừ Trung tâm Thông tin - Thư viện, Đại học Quốc gia Hà Nội, được đầu tư nhiều dự án lớn, đang sử dụng phần mềm thư viện số Content Pro (là PM chuyên dụng có nhiều tính năng ưu việt cho việc tạo lập, quản trị và khai thác TNS của Tập đoàn Innovative Interfaces, Hoa Kỳ), các thư viện khác đều đang dùng phần mềm (PM) thương mại trong nước hoặc PM mã nguồn mở. Trung tâm Học liệu ĐH Thái Nguyên và TV Nguyễn Thúc Hào, ĐH Vinh tự phát triển PM dùng riêng. Nhiều đơn vị đang sử dụng PM Dspace (được đánh giá tốt nhất trong các PM mã nguồn mở hiện nay) như Trung tâm Học liệu ĐH Đà Nẵng, Thư viện Trung tâm ĐHQG TpHCM, TVĐH Kiến Trúc, THĐH Bách Khoa HN, TVĐH Mở HN... Ngay Trung tâm TT - TV, ĐHQGHN, mặc dù đang sử dụng PM Content Pro, vẫn kết hợp sử dụng Dspace để tổ chức và khai thác các tài liệu học thuật dạng mở. Riêng TV Trường ĐHKHTN, ĐHQG TpHCM dùng PM mã nguồn mở Green Stone. Phần mềm KIPOS tích hợp tự động hóa thư viện và quản lý thư viện số đã được triển khai tại ĐH Y khoa thuộc ĐH Vinh, ĐH Nha Trang, Viện ĐH Mở HN... Vẫn còn khá nhiều thư viện sử dụng phân hệ tài liệu số trong PM quản trị thư viện tích hợp của các công ty tin học trong nước (Libol, Dlib, Ilib, Lạc Việt...), mặc dù tính năng còn nhiều hạn chế. Có lẽ chủ yếu do khó khăn về kinh phí nên các thư viện này chưa có điều kiện đổi sang dùng PM khác, hiệu quả hơn.

Hầu hết các TVĐH đều sử dụng chính phần mềm đang tạo lập, quản trị TNS để tìm kiếm, khai thác thông tin. Chỉ có 3 đơn vị dùng PM tìm

kiếm, khai thác thông tin chuyên dụng là Trung tâm Thông tin – Thư viện ĐHQGHN và Học viện Chính trị Quốc gia HCM (sử dụng Content Pro) trong khi ĐHQG TpHCM (sử dụng SFX, Primo).

Có thể nhận xét, đánh giá về các PM được hệ thống TVĐH Việt Nam đang sử dụng để hỗ trợ việc tạo lập, quản trị và khai thác TNS như sau:

- Hiện nay chưa có phần mềm nào được sử dụng ở Việt Nam có khả năng liên thông giữa biểu ghi thư mục MARC 21 và biểu ghi Dublincore; KIPOS đã liên thông được biểu ghi MARC 21 và METS; còn lại các phần mềm đều phải biên mục trên phần mềm thư viện điện tử tích hợp sau đó export ra Dublincore và import vào phần mềm thư viện số.
- Ngoại trừ Content Pro, các phần mềm thư viện số hiện đang sử dụng ở Việt Nam đều chưa có tính năng phân quyền linh hoạt (theo đối tượng bạn đọc, theo items, theo bộ sưu tập ...). Dspace cho phép xem hoặc tải về (tùy theo mức độ phân quyền), còn Greenstone thì không có tính năng phân quyền. Các phần mềm thương mại của Việt Nam hiện mới chỉ dừng lại ở mức độ nếu bạn đọc có tài khoản thì được phép tải về; bạn đọc không có tài khoản thì chỉ được xem thông tin thư mục (tương tự như Dspace)
- Nhược điểm của các phần mềm thư viện số do Việt Nam phát triển là không có khả năng sắp xếp kết quả tìm kiếm theo mức độ thích hợp (relevant); không hỗ trợ tìm kiếm toàn văn (KIPOS có chức năng tìm kiếm toàn văn nhưng sử dụng dựa trên nền tảng mã nguồn mở Lucene, không tích hợp được việc tìm kiếm toàn văn và tìm kiếm thuộc tính của tài liệu). Các sản phẩm phần mềm của nước ngoài và phần mềm mã nguồn mở đã giải quyết tốt các yêu cầu này (relevant kết quả tìm kiếm, tìm kiếm toàn văn, tìm kiếm toàn văn tích hợp với tìm kiếm thuộc tính).

- Phần lớn các phần mềm Việt Nam đều không có chức năng streaming documents, streaming video (bạn đọc muốn xem tài liệu thì phải tải về máy tính), điều này sẽ làm giảm tốc độ khai thác (mặc dù người dùng tin chỉ quan tâm đến 1 nội dung nào đó của tài liệu cũng phải tải toàn bộ tài liệu về).
- Tính tuân thủ chuẩn Dublincore của các phần mềm mã nguồn mở và phần mềm thương mại của nước ngoài thường tốt hơn phần mềm trong nước (phần mềm trong nước thường gặp lỗi khi kết xuất để chia sẻ dữ liệu).
- Đặc biệt Dspace hỗ trợ nhận diện với Google Scholar rất tốt, các sản phẩm trong nước và một số sản phẩm của nước ngoài chưa làm được việc này.

Về thiết bị, công cụ tạo lập, quản trị tài nguyên số

Để tạo lập TNS, do định dạng phổ biến hiện nay thường là file PDF nên các TVĐH thường áp dụng phương pháp chuyển dạng từ các tệp văn bản file text (thường là file MS word) sang dạng file PDF bằng tính năng save as PDF của MS Office hoặc các công cụ chuyển đổi định dạng khác. Đối với các tài liệu in được xuất bản trước đó đã lâu, không có hoặc không còn bản mềm, các TVĐH phải dùng thiết bị quét (scanner) để tạo file ảnh (JPEG, PDF), tiếp theo là nhận dạng ký tự quang học (OCR), chuyển sang PDF 2 lớp,... Sau đó lưu giữ vào các thiết bị chuyên dụng hoặc biên mục, đính file vào các phần mềm quản trị tài liệu số. Một vài thư viện còn tạo thêm bookmark cho tài liệu để bạn đọc thuận tiện trong sử dụng.

Do không có dự án đầu tư lớn, hầu hết các TVĐH dùng máy quét bằng tay (scanner) của các hãng HP, Canon, FUJITSU... tạo file ảnh để xây dựng kho tài liệu số.

Tuy nhiên, qua khảo sát thực tế, có thể thấy giải pháp quản lý chất lượng ảnh của các thiết bị quét tự động cũng rất khác nhau. Trong khi scanrobot của nhiều hãng chỉ có máy quét + phần mềm xử lý ảnh (về mặt

nào đó tương tự như PHOTOSHOP), dẫn tới việc kiểm soát chất lượng, tối ưu hóa nguồn nhân lực, ... rất khó khăn và chấp vá, thì duy nhất TREVENTUS có giải pháp tổng thể để kiểm soát toàn bộ quá trình từ đầu vào đến đầu ra nhằm bảo đảm chất lượng cho người dùng. Bộ giải pháp tổng thể gồm:

- Thiết bị quét
- Phần mềm xử lý ảnh hoàn hảo
- Phần mềm quản lý hoạt động (luồng công việc, tối ưu hóa, kiểm soát chất lượng, nhân lực...)
- Phần mềm “ thư viện số”

Ngoài ra, TREVENTUS còn đáp ứng mọi nhu cầu từ lưu trữ, xuất bản sách in, E-book, tái bản ...)

Phương thức và kết quả xây dựng kho tài liệu số của các thư viện

Kho TNS của các TVĐH được hình thành từ 3 nguồn gồm tự số hóa, mua và trao đổi. Nguồn tự số hóa được đa số các thư viện tiến hành và đối tượng thường là các tài liệu nội sinh. Loại hình tài liệu được các thư viện quan tâm số hóa nhiều nhất là luận án, luận văn, tiếp đó là sách giáo khoa, giáo trình, rồi đến kết quả NCKH và tạp chí. Cuối cùng là các tài liệu đặc biệt như các bộ sưu tập chuyên đề, tiêu chuẩn kỹ thuật, phát minh sáng chế...

Các thư viện đều xây dựng nguyên tắc lựa chọn tài liệu số hóa. Tiêu chí lựa chọn tài liệu nội sinh để số hóa có sự khác nhau giữa các thư viện. Song, nhìn chung, các tiêu chí thường được xác định theo các cấp độ ưu tiên giảm dần như sau:

- Tiêu chí 1 – Truy cập: Tài liệu được sử dụng nhiều hoặc tài liệu có giá trị khoa học, tư tưởng, nghệ thuật... cao. Tài liệu phục vụ các chương trình đào tạo, nghiên cứu chiến lược, trọng điểm của Nhà nước và của trường. Ưu tiên các tài liệu nội sinh có hàm lượng khoa học cao. Những tài liệu đơn bản hoặc hiếm cũng được xếp trong tiêu chí này.

- Tiêu chí 2 – Bảo quản: Tài liệu dễ hỏng, dễ phân hủy, khó bảo quản.
- Tiêu chí 3 – Cộng đồng: Tài liệu phục vụ cho các sự kiện, các triển lãm và xây dựng thương hiệu của nhà trường.
- Tiêu chí 4 – Tiềm năng phát triển: Là tài liệu có kinh phí từ dự án, tương lai được đầu tư, bổ sung hoặc số hóa theo yêu cầu của lãnh đạo hay nhà tài trợ.

Do điều kiện, khả năng về tài chính, thiết bị, nhân lực... rất khác nhau nên số lượng và chất lượng TNS có sự khác biệt khá lớn giữa các thư viện. Trong khi Trung tâm Thông tin – Thư viện, ĐHQGHN đã số hóa (có bản quyền) 3.500 sách, 20.000 luận án, luận văn; 7000 bài đăng tạp chí và kỷ yếu hội nghị khoa học (HNKH); 10.000 thư tịch Hán Nôm; 1000 đề mục trong bộ CSDL Địa chất và tài nguyên địa chất VN..., được chuyển PDF 2 lớp và làm bookmark thì nhiều TVĐH khác chỉ có vài trăm giáo trình, tổng số cả luận án, luận văn, bài đăng tạp chí, kỷ yếu HNKH, đề tài ... chưa đến 10.000 đơn vị.

Một số TVĐH lớn, có nguồn ngân sách dồi dào, có quyền tự chủ cao hoặc được nhận dự án đầu tư đã trực tiếp đặt mua CSDL sách, tạp chí... trực tuyến của các nhà cung cấp như Nam Hoàng, iGroup, Xunhasaba... Nhiều thư viện tham gia Liên hợp bổ sung tài liệu điện tử do Cục Thông tin KH&CN Quốc gia làm đầu mối để được quyền truy cập CSDL Proquest Central và một số CSDL của Cục. Qua quan hệ hợp tác quốc tế, một số thư viện được nhận tặng biếu, trao đổi CSDL với các trường đại học, các tổ chức khoa học nước ngoài. Tuy nhiên, tài liệu từ kênh này không ổn định và chất lượng khoa học không thật cao.

CÂU HỎI ÔN TẬP CHƯƠNG 1

1. Thư viện điện tử, thư viện số và thư viện ảo có phải là một không?
2. Hãy phân tích vai trò, chức năng và các nguyên tắc xây dựng thư viện số!
3. Phân tích nội dung chính của học liệu mở và truy cập mở
4. Trình bày tình hình phát triển thư viện số trên thế giới và tại Việt Nam!

CHƯƠNG 2: CÁC THÀNH PHẦN CỦA THƯ VIỆN SỐ

Mục tiêu đào tạo: Giúp người học hiểu rõ 4 thành phần cấu tạo của thư viện số. Đó là vốn tài liệu số, hạ tầng kỹ thuật, cán bộ thư viện và các dịch vụ cho người dùng thư viện số

Theo cách tiếp cận truyền thống, chúng ta có thể xem xét các thành phần cấu tạo nên Thư viện số được xác định bởi 4 yếu tố như sau:

1) Vốn tài liệu số (sưu tập số - Digital Collection): Bao gồm các đối tượng số (Digital Objects) và các siêu dữ liệu để hỗ trợ tra cứu và định vị tài nguyên số.

2) Hạ tầng kỹ thuật: Các thiết bị tin học; mạng; các công cụ hỗ trợ tìm kiếm; phần mềm; kho dữ liệu số...

3) Cán bộ thư viện số: Đội ngũ cán bộ có các khả năng và kỹ năng quản trị thư viện số; tổ chức thông tin và tri thức số; phổ biến thông tin số; phục vụ các dịch vụ tra cứu thông tin số, cung cấp tri thức từ các nguồn thông tin hiếm có; số hoá, xử lý, lưu trữ và bảo quản thông tin số; tìm và phục vụ thông tin số...

1) Dịch vụ thư viện số và người dùng thư viện số: Các dịch vụ của thư viện số thay đổi một cách đáng kể dựa trên các thành tựu của khoa học kỹ thuật, đặc biệt là công nghệ thông tin. Trong môi trường thông tin điện tử, người sử dụng không bị giới hạn bởi không gian và thời gian, có thể truy cập tới nguồn tin thông qua một máy tính có nối mạng (Tương tác trên máy tính - giao diện Interface). Người dùng thư viện số đòi hỏi phải có các kỹ năng máy tính, kiến thức chuyên môn, kiến thức thông tin...

2.1 VỐN TÀI LIỆU SỐ

2.1.1. Các đối tượng số

Trong lịch sử phát triển, vốn tài liệu đã từng xuất hiện dưới dạng sách trên đất sét, giấy papirut, giấy da,.. đến thế kỷ XV xuất hiện sách dưới dạng in. Từ thế kỷ XV, ngoài sách còn có vật mang tin khác như băng từ, đĩa từ. Ngày nay vốn tài liệu còn bao gồm cả các dữ liệu điện tử trên máy tính và mạng các máy tính.

Vốn tài liệu số là một trong bốn yếu tố cấu thành thư viện số, có ý nghĩa vô cùng quan trọng trong sự phát triển của thư viện thể hiện ở vai trò bao quát làm phong phú vốn tài liệu của thư viện, từ đó hỗ trợ nghiên cứu giảng dạy, học tập và giải trí.

Vốn tài liệu số bao gồm các đối tượng số (digital objects) và các siêu dữ liệu để hỗ trợ cho việc tra cứu và định vị tài nguyên. Các đối tượng số (hay còn gọi là tài liệu điện tử) là bất kỳ tài liệu nào được mã hoá, được lưu trữ trên các vật mang tin và người dùng tin có thể truy cập được thông qua thiết bị lưu trữ điện tử. Nguồn tin số bao gồm dữ liệu trực tuyến (Online) và dữ liệu ở trên vật mang tin vật lý như CD-ROM, DVD, VCD,...

Tài liệu điện tử/số bao gồm tất cả các dạng tài liệu như sách, báo, tạp chí, phim ảnh, bản nhạc, các CSDL,... được bao gói hay được lưu giữ trên các vật mang tin điện tử/số, có nghĩa là tất cả những gì có thể đọc được, truy cập được thông qua máy tính hay mạng máy tính.

Đặc trưng của tài liệu điện tử/số:

Ưu điểm:

- Có mật độ thông tin rất cao, do vậy dung lượng thông tin được lưu trữ trên chúng cũng rất lớn.
- Có khả năng đa truy cập, cho phép người dùng có thể tìm tài liệu đồng thời theo nhiều dấu hiệu, nhiều điểm truy cập khác nhau và truy cập linh hoạt, nhanh chóng.

- Khả năng liên hệ, tiếp cận với các tác giả, tạo ra kênh phản hồi thông tin giữa người dùng tin và người sáng tạo ra thông tin.
- Cho phép lưu giữ thông tin từ nhiều dạng nguồn tin khác nhau như văn bản, âm thanh, hình ảnh tĩnh và hình ảnh động trong cùng một tài liệu.
- Khả năng cho phép nhiều người sử dụng cùng một tài liệu một lúc.
- Khả năng cập nhật dữ liệu nhanh chóng và thường xuyên.

Nhược điểm:

- Tính an toàn thông tin dễ bị vi phạm, do việc sao chép thông tin từ các tài liệu điện tử rất dễ dàng, nhanh chóng.
- Tính ổn định của thông tin thường không đồng nhất, có tài liệu thì rất ổn định, tồn tại lâu dài như các tài liệu ghi trên CD-ROM, DVD-ROM nhưng lại có những tài liệu có đời sống rất ngắn.

2.1.2. Siêu dữ liệu

2.1.2.1. Khái niệm

Siêu dữ liệu được cấu trúc, mã hoá dữ liệu để miêu tả những đặc tính của những thực thể mang thông tin nhằm trợ giúp cho việc mô tả, khám phá, quản trị và truy cập các thực thể được miêu tả.

Trước tiên chúng ta hãy định nghĩa siêu dữ liệu là gì? Một định nghĩa chung nhất cho siêu dữ liệu đó là “*dữ liệu về dữ liệu*”. Định nghĩa này thoạt nghe sẽ cảm thấy cuốn hút, nhưng nó không giúp chúng ta hiểu toàn bộ siêu dữ liệu là gì nhưng thực chất siêu dữ liệu là thông tin được cấu trúc, điều này có nghĩa là nó thuộc phạm trù sáng tạo của con người, và chúng ta không tìm thấy trong tự nhiên.

Ngoài ra còn có định nghĩa khác của Dempsey và Heery, 1997 như sau: “siêu dữ liệu là dữ liệu đi kèm cùng với đối tượng thông tin và nó cho phép những người sử dụng tiềm năng có thể biết trước sự tồn tại cũng như đặc điểm của đối tượng thông tin này”.

Mục đích đầu tiên và cốt yếu nhất của siêu dữ liệu là góp phần mô tả và tìm lại các tài liệu điện tử trên mạng Internet. Sự phát triển mạnh mẽ của Internet đã tạo ra sự bùng nổ của các loại dữ liệu đa dạng ở dạng số, văn bản, hình ảnh, âm thanh, hình ảnh động, tài liệu đa phương tiện. Những tài liệu số này có thể truy cập được trên Internet song việc tìm kiếm chúng một cách hiệu quả và khoa học như với các hệ thống thông tin trực tuyến là hết sức khó khăn. Để góp phần tăng cường chất lượng tìm kiếm các tài liệu số trên mạng Internet, người ta đã đưa ra giải pháp sử dụng siêu dữ. Siêu dữ liệu hỗ trợ phát triển nguồn tin; kiểm tra sự tồn tại của đối tượng thông tin; mô tả ngôn ngữ, vị trí; hỗ trợ người dùng tin đánh giá thông tin mà không phải truy cập trực tiếp đến thông tin.

Siêu dữ liệu trong thư viện số bao gồm các loại sau:

- Siêu dữ liệu hành chính (Administrative): Được dùng để quản lý và quản trị các tài nguyên thông tin. Ví dụ:

- + Thông tin bổ sung
- + Bản quyền và thông tin tái bản
- + Tài liệu về yêu cầu truy cập hợp pháp
- + Thông tin định vị
- + Các tiêu chí số hoá
- + Kiểm soát phiên bản và những khác biệt giữa các đối tượng thông tin tương đồng

+ Thông tin kiểm tra của hệ thống quản lý

- Siêu dữ liệu mô tả (Discriptive): Dùng để mô tả hay nhận diện các tài nguyên thông tin. Ví dụ:

- + Các biểu ghi thư mục
- + Các hỗ trợ tìm kiếm
- + Định chỉ số chuyên biệt
- + Siêu liên kết giữa các tài nguyên
- + Chú giải của người sử dụng

+ Siêu dữ liệu cho hệ thống quản lý biểu ghi được tạo ra bởi người tạo ra biểu ghi

- Siêu dữ liệu bảo quản (Preservation): Đưa ra các thông tin liên quan đến quản lý việc bảo quản tài nguyên thông tin. Ví dụ:

+ Các tài liệu về tình trạng, điều kiện vật lý của tài nguyên thông tin

+ Các tài liệu về công tác bảo quản các phiên bản thông tin dưới dạng vật lý và số.VD: Làm mới dữ liệu (Refreshing); Di trú dữ liệu (Migration)

- Siêu dữ liệu sử dụng (Use): Các thông tin liên quan đến mức độ và loại hình sử dụng tài nguyên thông tin. Ví dụ:

+ Thông tin về phần cứng và phần mềm

+ Thông tin số hoá. VD: Khổ mẫu; Tỷ lệ nén (độ nén); Quy trình phân bố thông tin.

+ Thông tin về thời gian phản hồi của hệ thống

+ Dữ liệu về tính xác thực...

- Siêu dữ liệu kỹ thuật (Technical): Các thông tin liên quan đến cách thức hoạt động của hệ thống cũng như siêu dữ liệu. Ví dụ:

+ Các biểu ghi trình bày (Exhibit)

+ Các thông tin về người sử dụng và việc sử dụng

+ Thông tin về tái sử dụng nội dung và các phiên bản đa phương tiện

Thuộc tính và đặc điểm của siêu dữ liệu

- Xét về nguồn gốc siêu dữ liệu được chia thành: siêu dữ liệu nội sinh (Internal) và siêu dữ liệu ngoại sinh (External). Siêu dữ liệu nội sinh là siêu dữ liệu được cơ quan, tổ chức tạo ra ngay sau khi đối tượng thông tin lần đầu tiên được tạo ra hay được số hoá. Siêu dữ liệu ngoại sinh liên quan đến đối tượng thông tin được tạo ra sau này, do một người khác sau khi được người sáng tạo đầu tiên sáng tạo ra.

- Xét về phương pháp tạo siêu dữ liệu: có thể do máy tính tự động

tạo ra hay do chính con người tạo ra.

- Bản chất của siêu dữ liệu: Có thể do người không phải là chuyên gia thông tin, mà đó là người đầu tiên tạo ra đối tượng thông tin tạo nên.

- Tình trạng: Có thể là siêu dữ liệu ổn định không bao giờ biến đổi từ khi được tạo ra, hoặc siêu dữ liệu biến đổi do thường thay đổi trong quá trình vận hành đối tượng thông tin. Hoặc siêu dữ liệu dài hạn cần thiết cho việc truy cập và sử dụng đối tượng thông tin, hoặc siêu dữ liệu ngắn hạn thông tin về bản chất của quá trình hoạt động thông tin.

2.1.2.2. Vai trò

- Siêu dữ liệu cho phép tra cứu chéo đa ngành các nguồn tài nguyên trên mạng

- Siêu dữ liệu cung cấp một tầng thông tin trung gian mà tại đó người ta có thể lựa chọn để xem hay tìm kiếm trên một lượng thông tin lớn các thông tin toàn văn không thích hợp.

- Các siêu dữ liệu cũng có thể được dùng để lưu trữ thông tin cho các mục đích như:

- + Quản lý các tiến trình hoạt động của một dự án: ví dụ: tên những người thực hiện công việc, ngày tháng, nhiệm vụ đã thực hiện

- + Mô tả các quá trình kỹ thuật được thực hiện trong việc tạo ra các đối tượng đại diện số hoá

2.1.2.3. Các chuẩn siêu dữ liệu phổ biến

- Các chuẩn siêu dữ liệu ra đời bắt đầu từ những năm 90 bởi một số nhóm: chuyên gia máy tính, cán bộ thư viện, chuyên gia tư liệu

- Mục đích ban đầu nhằm giúp những người sử dụng mạng khai thác các nguồn tài nguyên một cách thuận tiện.

a) Chuẩn MARC

MARC (**MA**chine **R**eadable **C**ataloging = Biên mục máy đọc được) là một hệ thống được phát triển bởi Thư viện Quốc hội Hoa Kỳ vào năm

1966 để các thư viện có thể chia sẻ những dữ liệu thư mục máy đọc được (Machine-Readable Bibliographic Data). Có nghĩa rằng các Hệ thống quản trị thư viện tự động phải cần có một dạng thức chung để có thể trao đổi dữ liệu với nhau, dạng thức đó gọi là MARC.

MARC được phát triển đầu tiên từ Hoa kỳ với tên là USMARC, dần dần các quốc gia có điều kiện phát triển tự động hóa đều định riêng chuẩn MARC cho mình chẳng hạn như UKMARC (Anh), AUMARC (Úc),vv...

Quan niệm chuẩn hóa ngày càng được khẳng định với quy mô rộng lớn hơn. Ba chuẩn MARC được biết đến nhiều nhất là:

1. CFC: do một hội nghị của UNESCO phát triển, đã từng họp định kỳ hàng năm để cập nhật, nhưng nay không còn nữa. Cơ quan đại diện là UNESCO. Đây là chuẩn MARC của phần mềm CDS/ISIS

2. UNIMARC: do một hội nghị của IFLA (Hiệp hội thư viện thế giới) phát triển, họp định kỳ hàng năm để cập nhật. Cơ quan đại diện là Chương trình IFLA UBCIM

3. MARC 21: tiền thân là USMARC phối hợp với MARC của Canada. Cơ quan đại diện là Thư viện quốc hội Hoa kỳ và Thư viện quốc gia Canada

MARC 21 đang được dùng rộng rãi khắp nơi trên thế giới bởi lẽ các kho tin khổng lồ trên thế giới, các CSDL trực tuyến tiên tiến nhất trên mạng toàn cầu Internet đều sử dụng MARC 21. Ngày càng có nhiều các chuẩn MARC của các quốc gia như UKMARC (Anh), AUMARC (Úc),vv... chuyển sang MARC 21.

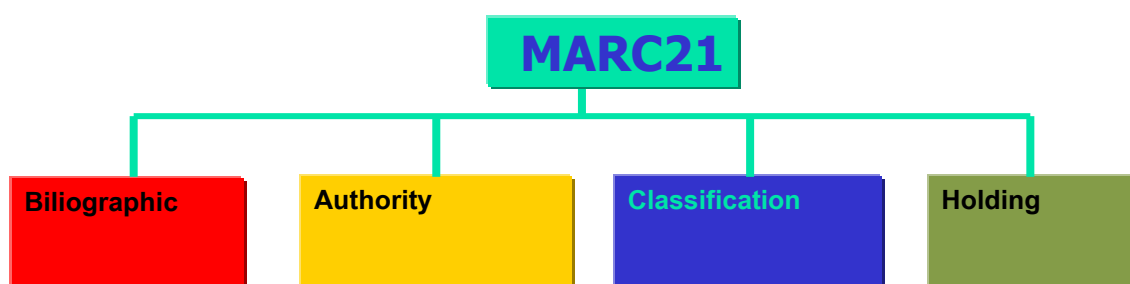
MARC 21 trở thành chuẩn quốc tế cho biên mục máy đọc được, hay chuẩn trao đổi dữ liệu trên phạm vi toàn cầu.

Thành phần của MARC21 gồm:

- Khổ mẫu mô tả thư mục - Bibliographic format: đây là khổ mẫu chính của MARC21 dùng để mô tả các yếu tố thư mục của tài

liệu.

- Khở mẫu kiểm soát thống nhất - Authority format: Khở mẫu này quy định các thức nhập và xây dựng dữ liệu nhất quán cho biểu ghi dùng làm điểm truy cập nhất quán như: tên tác giả, đề mục chủ đề, nhan đề thống nhất,...
- Khở mẫu phân loại - Classification format: Khở mẫu này quy định quy tắc và cấu trúc tổ chức phân loại tài liệu trong biên mục.
- Khở mẫu lưu trữ tài liệu - Holding Data format: khở mẫu này quy định cách thức ghi nhận và kiểm soát tài liệu trong thư viện, phản ánh cụ thể và chi tiết đến từng đơn vị tài liệu của một thư viện hoặc của nhiều thư viện thành viên. Đây là cơ sở để xây dựng mục lục liên hợp.



Hình 2.1 Cấu trúc MARC21

b) Chuẩn Dublin core

Khở mẫu tiêu chuẩn siêu dữ liệu Dublin core (DC) là một tập hợp các yếu tố đơn giản nhưng hữu hiệu trong việc mô tả một loạt nguồn tin trên mạng. DC gồm 15 yếu tố, mà ngữ nghĩa được xác lập của nhiều chuyên gia. Các yếu tố mô tả trong DC đều không bắt buộc và có thể lặp, ngoài ra còn có một số lượng hạn chế các từ hạn định và định ngữ có thể sử dụng để tiếp tục tinh chỉnh ý nghĩa của các yếu tố.

Các yếu tố mô tả: 15 yếu tố

- Nhan đề: tên của tài liệu được đặt bởi tác giả hoặc nhà xuất bản
- Tác giả (người sáng tác): cá nhân hoặc tổ chức sáng tạo ra

nguồn thông tin, ví dụ: nhà văn, họa sĩ, nhà nhiếp ảnh, người minh họa

- Chủ đề và từ khoá: Là chủ đề của nguồn tin. điển hình, chủ đề có thể được hiểu tương đương như từ khoá nhằm mô tả về một chủ đề hoặc một nội dung của nguồn thông tin. Người ta rất khuyến khích việc sử dụng các quyền từ vựng có kiểm soát và các khung phân loại.
- Mô tả: miêu tả nguyên văn về nội dung của nguồn thông tin, có thể bao gồm cả tóm tắt trong trường hợp tài liệu thành văn hoặc mô tả nội dung trong trường hợp không xác định rõ.
- Nhà xuất bản: một thực thể chịu trách nhiệm về việc xây dựng nên hình thức của tài liệu ví dụ: nhà xuất bản, 1 khoa trong trường đại học, hoặc một tổ chức
- Những người cộng tác: là một cá nhân hoặc một tổ chức trên danh nghĩa cũng một người đóng góp trong việc xây dựng nên nguồn thông tin, tuy nhiên, đối với những cá nhân hoặc tổ chức xây dựng nên nguồn tin thì họ chỉ đóng vai trò phụ, thứ yếu. ví dụ: người biên tập, người dịch, người hiệu đính,..
- Thời gian xuất bản: thời gian xuất bản nguồn tin, được diễn tả theo cấu trúc Năm/Tháng/Ngày. Ví dụ: 1994-11-05
- Loại dữ liệu: kiểu mô tả cho nguồn thông tin, ví dụ: Trang chủ, bài thơ, quyền sách, báo cáo kỹ thuật, bài luận, từ điển, nhằm mục đích quốc tế hoá, kiểu mô tả nên nguồn tin nên lấy trong danh sách đã được xây dựng.
- Khổ mẫu (định dạng): Kiểu của dữ liệu tùy thuộc vào từng tài liệu, ví dụ: khối lượng và thời gian của nguồn thông tin. Kiểu dữ liệu được dùng để mô tả cho nguồn thông tin, nhằm mục đích quốc tế hoá. Định dạng nguồn tin nên chọn trong danh sách mới nhất do tổ chức quốc tế xây dựng.

- Yếu tố nhận dạng: Mã nhận dạng: là một từ hoặc con số được sử dụng để nhận dạng nguồn thông tin, ví dụ: các nguồn thông tin trên mạng bao gồm cả URL và URNs. Có một số mã nhận dạng duy nhất như : ISBN, chỉ số sách theo tiêu chuẩn quốc tế, hoặc một số tên tương tự khác cũng có thể đại diện cho yếu tố này.
- Nguồn: những thông tin nguồn dữ liệu
- Ngôn ngữ: Ngôn ngữ của nội dung thông tin.
- Quan hệ: là một yếu tố nhận dạng về nguồn thông tin thứ 2 có liên quan đến nguồn thông tin hiện thời. Yếu tố này được sử dụng để mô tả cho mối quan hệ của các nguồn thông tin, nhằm mục đích quốc tế hoá, mỗi liên hệ này nên chọn trong danh mục của các...
- Phạm vi (mức độ bao phủ): Không và thời gian tạo ra tài liệu.
- Bản quyền: sự xác định về quyền hạn của người quản lý nguồn thông tin.

Chia nhóm trong DC:

- Các đối tượng về nội dung thông tin
 - + Nhan đề
 - + Chủ đề (từ khoá)
 - + Mô tả
 - + Nguồn
 - + Ngôn ngữ
 - + Quan hệ
 - + Phạm vi (mức độ bao phủ)
- Các đối tượng liên quan đến sở hữu trí tuệ
 - + Tác giả (người sáng tác)
 - + Nhà xuất bản
 - + Nhà phân phối, cộng tác

- + Quản lý Bản quyền
- Các đối tượng cá biệt
 - + Thời gian xuất bản
 - + Loại dữ liệu
 - + Khổ mẫu (định dạng)
 - + Yếu tố nhận dạng

Các từ chuẩn giới hạn của Dublin Core

- + Lọc yếu tố mô tả: Làm cho nghĩa của một yếu tố mô tả hẹp hơn hoặc cụ thể hơn
- + Khung mã hoá: nhận diện các quy tắc trợ giúp diễn giải giá trị của một yếu tố mô tả nào đó, bao gồm các từ vựng kiểm soát và các ký hiệu hình thức hay các quy tắc cú pháp
- + Từ vựng cho các yếu tố mô tả: bộ sưu tập; bộ dữ liệu; sự kiện; hình ảnh; các nguồn dữ liệu có tương tác; dịch vụ; phần mềm; âm thanh; văn bản.

Công cụ và đặc điểm Dublin Core

- Công cụ: <http://dublincore.org/tools>
 - + Các tiện ích
 - + Tạo siêu dữ liệu theo mẫu
 - + Công cụ tạo/ thay đổi mẫu
 - + Tự động trích dẫn/ tập hợp lại siêu dữ liệu
 - + Tự động tạo siêu dữ liệu
 - + Chuyển đổi giữa các khổ mẫu siêu dữ liệu
 - + Các môi trường (công cụ) tích hợp
 - + Các phần mềm thương mại sẵn có
- Hướng dẫn sử dụng Dublin Core:
<http://dublincore.org/documents/usageguide>
- + Hướng dẫn sử dụng cho những người không phải là chuyên gia về tạo biểu ghi mô tả DC

- + Cho các chuyên gia tham khảo về DC
- + Giới thiệu và diễn giải về DC
- Đặc điểm Dublin Core
 - + Tạo lập và duy trì biểu ghi một cách dễ dàng
 - + Ngữ nghĩa dễ hiểu
 - + Phạm vi sử dụng quốc tế rộng lớn với các phiên bản đa ngôn ngữ
 - + Khả năng mở rộng các yếu tố thuận tiện

c) Chuẩn XML

Khái niệm về XML

XML (eXtensible Markup Language) ngôn ngữ đánh dấu mở rộng có nguồn gốc từ ngôn ngữ định dạng siêu văn bản HTML (Hyper Text Markup Language), cả hai ngôn ngữ này đều bắt nguồn từ chuẩn ngôn ngữ định dạng văn bản tổng quát có cấu trúc SGML (Structured General Markup Language).

XML là ngôn ngữ được định nghĩa bởi tổ chức mạng toàn cầu (World Wide Web Consortium), thường được viết tắt theo cách chơi chữ là W3G. Đây là tổ chức quốc tế định ra các chuẩn của Web và Internet.

Cấu trúc XML

Một văn bản XML hình thành từ các thẻ (tag) với tên gọi phần tử (element). Khác với ngôn ngữ HTML, số lượng và tên gọi trong XML là không hạn chế. XML là ngôn ngữ tổng quát dùng định nghĩa dữ liệu thông qua các thẻ. Trong HTML các thẻ được định nghĩa và quy định trước. Trong khi đó, với XML ta có thể tùy ý định nghĩa mọi thẻ. Như vậy có thể coi XML như tập cha của ngôn ngữ HTML. Dựa vào một số quy tắc, XML tự tồn tại và phát triển tự thân thành các ngôn ngữ định nghĩa khác.

Điểm quan trọng nhất là XML cho phép dễ dàng xử lý, chuyển tải và trao đổi dữ liệu giữa rất nhiều ứng dụng và tài liệu người dùng với các

định dạng khác nhau. Nếu đã quen với máy tính, hẳn ta biết rằng có rất nhiều định dạng file khác nhau. Việc chuyển đổi dữ liệu giữa chúng quả là nan giải mặc dù có không ít trình ứng dụng hỗ trợ. Ví dụ như file.DOC (văn bản Word), .XLS (bảng tính Excel), .DBF (Lập trình Foxpro, >MDB (lập trình Access), .TXT (file văn bản), .RTF (Rich Text Format) và mới đây là .HTML. Chỉ riêng các file văn bản thôi đã đủ gây khó chịu, nếu bạn nhận được một file Word 2000 mà máy tính của bạn còn dùng Word 7.0, cố gắng lắm cũng chỉ đọc được phần văn bản còn các nội dung khác thường bị biến dạng.

Trong XML, dữ liệu và định dạng được lưu ở dạng văn bản và có thể dễ dàng cấu hình cũng như thay đổi chúng bằng các trình soạn thảo thông thường nếu không có trong tay trình soạn thảo XML chuyên nghiệp. Dữ liệu và các thẻ trong XML không mã hoá, không đòi hỏi bản quyền.

Tháng 12/1997, phiên bản đầu tiên XML 1.0 (Extensible Markup language- Ngôn ngữ đánh dấu mở rộng) ra đời và là chuẩn đơn giản của SGML. Từ đó, nhiều công ty phần mềm đã chạy đua ứng dụng XML vào nhiều ứng dụng. Hàng trăm ngôn ngữ định dạng chuyên dụng dựa trên XML đã ra đời. Điển hình một số tùy biến ngôn ngữ định dạng dựa trên XML cho thấy sức mạnh của XML:

- BITS- Banking Industry Technology Secretariat: Ngôn ngữ văn phòng về kỹ thuật nghiệp vụ ngân hàng
- IFX - Financial Exchange: Trao đổi dữ liệu tài chính.
- BIPS - Banking Internet Payment System: Hệ thống thanh toán qua Internet của nghiệp vụ ngân hàng.
- TIM - Telecommunication Interchange Markup: Định dạng trao đổi viễn thông.
- CBL - Common Business Library: Thư viện kinh doanh phổ thông.

- EbXML - XML: kinh doanh điện tử
- PDML - Product Data Markup Language: Ngôn ngữ định dạng dữ liệu sản phẩm.
- FIX - Financial Information eXchange Protocol: Giao thức trao đổi thông tin tài chính.
- CML - Chemical Markup Language: Ngôn ngữ định dạng trong lĩnh vực hoá học, cho phép biểu diễn các công thức hoá học, hoá trị phân tử ở dạng đồ hoạ.

Ưu điểm của XML:

- Có thể tạo biểu ghi thư tịch một lần và xuất bản chúng theo các dạng khác nhau;
- Hiện thị biểu ghi thư tịch trực tiếp trên trình duyệt Web, search engines (công cụ tìm kiếm) và các hệ thống thư viện tiềm năng khác mà không cần chuyển đổi;
- Biểu ghi thư tịch có thể được chuyển đổi qua lại giữa XML và MARC mà không bị tổn thất.
- Nhiều vấn đề tồn tại trong định dạng MARC được khắc phục, ví dụ như việc kiểm soát tiêu đề chuẩn.

d) Chuẩn MODS

Khổ mẫu MARC một biểu ghi được cấu trúc chuyên biệt ở mức cao cho việc mã hóa thông tin thư tịch. Nó nổi tiếng trong thế giới thư viện và được hậu thuẫn bởi nhiều hệ thống thư viện tại Hoa Kỳ, Canada và nhiều nước khác, đặc biệt trong thế giới các quốc gia nói tiếng Anh. Trong một môi trường nối kết mạng mà ở đó siêu dữ liệu mô tả được chuyển tiếp qua nhiều hệ thống và có thể gắn liền vào trong hoặc với nhiều loại siêu dữ liệu khác, thì quả thật lý tưởng để sử dụng biểu ghi MARC cho mục đích này. Tuy nhiên, vấn đề của MARC đó là việc nhúng dữ liệu này nói chung đòi hỏi sử dụng một cấu trúc dữ liệu XML, và MARC không phải một biểu ghi XML. Thư viện Quốc hội Mỹ đã tạo

ra một cách nhằm biên dịch biểu ghi MARC sang XML, nhưng đã không nhận được sự hưởng ứng, vì một lý do có thể nhận thấy, đó là: biểu ghi MARC lớn và chi tiết hơn là mọi hệ thống có thể cần, và việc sử dụng các thẻ số và mã trường con làm cho mọi người khó hiểu nếu không được đào tạo đầy đủ. Bởi vậy, người ta đã cần một phiên bản MARC nhẹ nhàng và đơn giản hơn mà có thể chấp nhận những phần tử dữ liệu chính từ một biểu ghi MARC và chuyển chúng vào một khổ mẫu XML dễ dàng hiểu hơn. Vậy, tiêu chuẩn MODS đã ra đời (Tiêu chuẩn Mô tả Đối tượng siêu dữ liệu - Metadata Object Description Standard). MODS sử dụng những thẻ mà người ta dễ dàng hiểu để thay thế cho những thẻ bằng 3 chữ số và mã trường con của MARC (vd., “Nhan đề” (“title”) thay thế cho “245”). MODS bỏ qua phần lớn các phần tử dữ liệu của trường cố định, ngoại trừ những mã khổ mẫu có tính vật lý (của mã 007) và nhiều mã dành cho thể loại (của mã 008). MODS cũng giới thiệu một số cách thực hành hiệu quả và cách tân hơn. MODS định nghĩa một cấu trúc gọi là “Tên” (“Name”) đại diện nhiều trường bao gồm cả trường con để chứa đựng thông tin tên cá nhân, công ty và cho tên một hội nghị. Cấu trúc này có thể được sử dụng bất kỳ khi nào một tên xuất hiện, nó có thể là mục dữ liệu chính, phụ hoặc chủ đề

Mặc dù được sinh ra từ MARC21 và chi tiết hơn nhiều tiêu chuẩn Dublin Core, những MODS lại ít quy tắc biên mục hơn MARC21. Tương tự như Dublin Core, không có nhiều trường phụ thuộc và tất cả các trường đều có thể lặp lại. MODS chứa đựng nhiều giá trị từ MARC, song nó cũng có những nhiều điểm khác với MARC21: Không có những khái niệm về “dẫn mục chính” (“main entry”) hoặc “dẫn mục phụ” (“added entry”), tất cả tác giả đơn giản chỉ là “tác giả”; và một biểu ghi có thể có nhiều nhan đề mà không có một “nhan đề chính” (“main title”). Khi những biểu ghi MARC21 được biên dịch tới MODS, thì người ta có một biểu ghi trong XML mang dấu ấn của MARC. Biểu ghi MODS có

thể được tạo ra từ siêu dữ liệu thư tịch mà không khởi đầu từ mục lục thư viện, ví dụ như trích dẫn của bài báo, và nó thường được sử dụng trong cơ sở dữ liệu có sự trộn lẫn của mục lục thư viện và dữ liệu biên mục khác.

e) Chuẩn METS

METS là một tiêu chuẩn để mã hóa mô tả, hành chính, và cấu trúc siêu dữ liệu liên quan đến các đối tượng trong một thư viện kỹ thuật số, thể hiện bằng cách sử dụng ngôn ngữ lược đồ XML. Một tiêu chuẩn METS bao gồm 5 phần chính:

- Siêu dữ liệu mô tả: phần siêu dữ liệu mô tả có thể chỉ ra siêu dữ liệu mô tả bên ngoài tiêu chuẩn METS (Ví dụ: một khổ mẫu MARC trong OPAC hay chương trình hỗ trợ tìm kiếm EAD trong một dịch vụ mạng toàn cầu WWW server), hoặc bao gồm siêu dữ liệu mô tả mã hóa bên trong hoặc cả hai.

- Siêu dữ liệu quản trị: phần siêu dữ liệu quản trị cung cấp thông tin về tạo dựng và lưu trữ các File như thế nào, quyền sở hữu trí tuệ, siêu dữ liệu về nguồn gốc của đối tượng thư viện số, và thông tin về các File gốc cấu tạo nên đối tượng thư viện số (Ví dụ: Mối quan hệ giữa File gốc và File phát sinh, thông tin về chuyển đổi và thay thế). Giống siêu dữ liệu mô tả, siêu dữ liệu quản trị có thể chỉ ra cả siêu dữ liệu quản trị bên ngoài lẫn siêu dữ liệu mã hóa bên trong.

- Nhóm Files: phần nhóm Files liệt kê tất cả các File cấu tạo nên các phiên bản điện tử của đối tượng số. Các phần tử của nhóm file giúp tham chiếu đến các file cấu tạo của phiên bản số.

- Sơ đồ cấu trúc: Sơ đồ cấu trúc là trung tâm chính của 1 file METS. Phần này mô tả cấu trúc thứ bậc của đối tượng thư viện số, tạo liên kết giữa các phần tử của cấu trúc đó với các tệp nội dung và siêu dữ liệu liên quan.

- Cơ chế vận hành: Bộ phận vận hành được sử dụng để liên kết các

hoạt động quản lý với phần nội dung của một đối tượng được mô tả theo tiêu chuẩn METS. Phần này có 1 bộ phận xác định giao diện có thể mô tả cụ thể, chính xác các quy trình hoạt động. Ngoài ra nó còn có một cơ chế vận hành, đó là một bộ phận gồm các mã vận hành có nhiệm vụ duy trì và thực hiện các hoạt động mà bộ phận xác định giao diện đã lập ra.

Tiêu chuẩn METS cung cấp một cơ chế linh hoạt để mã hóa siêu dữ liệu mô tả, quản trị và cấu trúc cho một đối tượng số trong thư viện, và để thể hiện các mối quan hệ phức tạp giữa các hình thức khác nhau của siêu dữ liệu. Bởi vậy nó có thể cung cấp 1 tiêu chuẩn thống nhất cho việc trao đổi các đối tượng số giữa các kho trong thư viện.

2.2 HẠ TẦNG KỸ THUẬT

2.2.1. Phần cứng

Bao gồm các trang thiết bị cần thiết: hệ thống thiết bị mạng, hệ thống máy chủ, hệ thống máy trạm, các thiết bị ngoại vi (máy in, thiết bị mượn/trả ...), thiết bị an ninh thư viện ...

2.2.1.1. Hệ thống thiết bị mạng

Hệ thống hạ tầng mạng là thành phần cơ bản của mạng thông tin thực hiện chức năng kết nối thiết bị như máy chủ, máy trạm, các thiết bị mạng thông qua đường kết nối vật lý. Hệ thống hạ tầng mạng bao gồm thiết bị kết nối và môi trường truyền.

Các thiết bị kết nối phổ dụng hiện nay bao gồm: hub, switch, router, firewall cứng ...

- Hub hay bộ chuyển mạch đơn: Là một trong những yếu tố quan trọng nhất của mạng nội bộ, đây là điểm kết nối dây trung tâm của mạng, tất cả các trạm trên mạng được kết nối thông qua HUB.



- Switch hay bộ chuyển mạch: là bộ phận tối quan trọng trong



mạng, chức năng chính của switch là trung chuyển dữ liệu trong mạng.

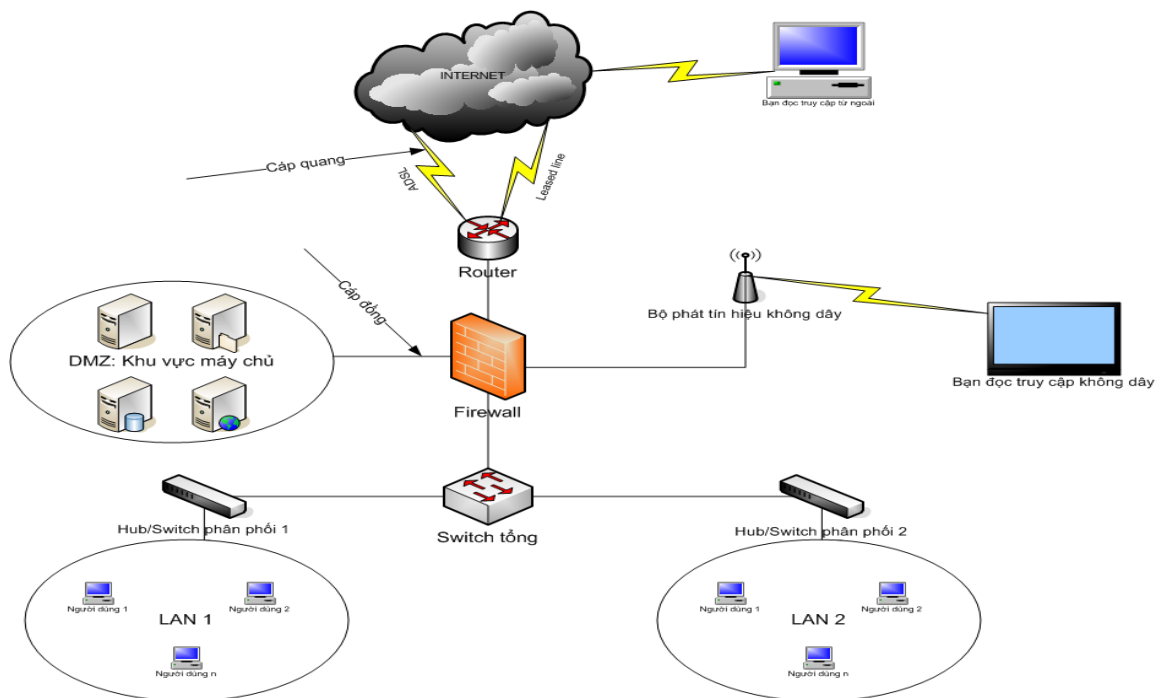
- Router hay bộ định tuyến: khác hơn các bộ chuyển mạch, một bộ định tuyến là thiết bị nối giữa các mạng. Nó có khả năng tạo nối kết với các nút mạng không có chung giao thức.



- Firewall cứng hay tường lửa cứng: là một thiết bị phần cứng hoạt động trong một môi trường máy tính nối mạng để ngăn chặn một số liên lạc bị cấm bởi chính sách an ninh của cá nhân hay tổ chức.



Môi trường truyền: môi trường truyền có thể là các loại dây dẫn (cáp đồng, cáp quang ...), sóng điện từ (đối với các mạng không dây); nó là một trong những thành phần không thể thiếu của bất kỳ một hệ thống mạng nào. Mục tiêu đầu tiên của môi trường truyền là phải vận chuyển dòng dữ liệu một cách nhanh nhất và ổn định nhất, mục tiêu thứ hai là phải triển khai nhanh nhất.



Hình 2.2 Mô hình mạng cơ bản trong thư viện

2.2.1.2. Hệ thống máy chủ

Để đảm bảo cung cấp các dịch vụ phục vụ cho hệ thống mạng, cần có hệ thống máy chủ. Máy chủ là một máy tính hay một thiết bị trên mạng, chức năng chủ yếu của nó là quản lý các tài nguyên của mạng. Các máy chủ là các máy được sử dụng để cho nhiều máy tính khác nhau cùng truy cập, vì vậy thông thường chúng là các máy tính có cấu hình phần cứng rất mạnh, và có các hệ thống phần mềm chuyên biệt để có thể phục vụ cho nhiều người cùng một lúc (các phần mềm này thường gọi là các phần mềm phía server). Các máy chủ có thể được sử dụng để cung cấp nhiều dịch vụ cùng một lúc, chẳng hạn như vừa truyền tải file, vừa cung cấp các trang Web... nhưng để đảm bảo cho các dịch vụ luôn hoạt động tốt, các nhà quản lý thường chuyên biệt hoá các máy chủ theo từng mục đích khác nhau. Chẳng hạn máy chủ Web (Web server) chuyên dùng để quản lý Web, máy chủ FTP (FTP server) chuyên dùng trong việc quản lý các file được upload hay download, máy chủ SMTP (SMTP server) chuyên dùng vào việc trao đổi các thông tin thư điện tử (E-mail) giữa các máy chủ với nhau....



Hình 2.3 Một số hình ảnh minh họa máy chủ

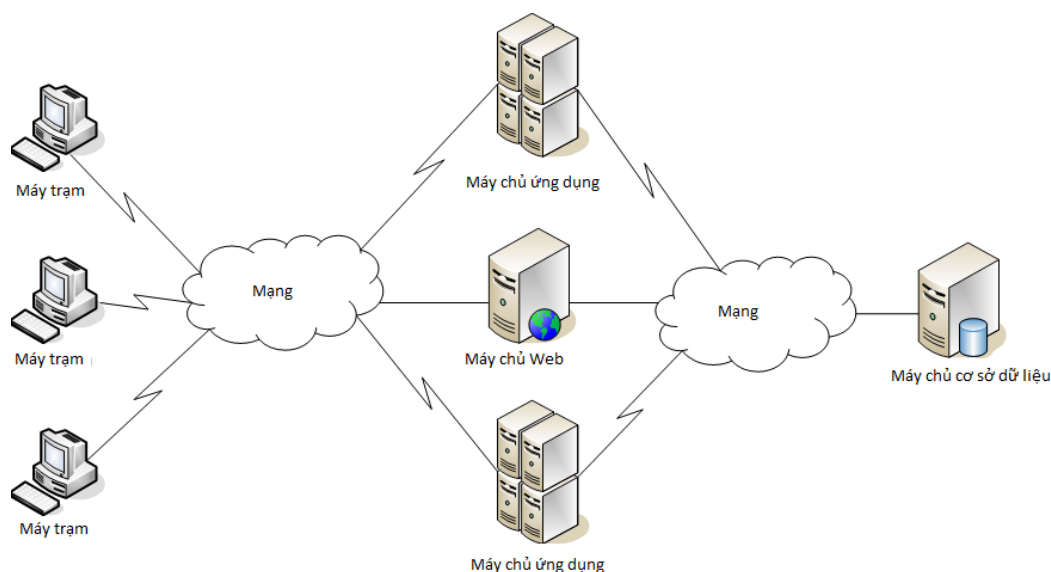
Máy chủ ứng dụng (Application Server): Là một chương trình mà điều khiển tất cả các hoạt động ứng dụng giữa những người sử dụng và các ứng dụng tầng cuối của một tổ chức hoặc là các cơ sở dữ liệu. Các dịch vụ ứng dụng điển hình được sử dụng cho các ứng dụng liên hợp giao dịch cơ bản. Để hỗ trợ yêu cầu cao, một dịch vụ ứng dụng phải có

sự dư thừa bên trong, điều khiển cho khả năng xuất hiện cao, trình diễn mức độ cao, phân bố các dịch vụ ứng dụng và hỗ trợ liên kết truy cập cơ sở dữ liệu.

- Máy chủ ứng dụng (Application Server): Các ứng dụng được cài đặt trên một mạng để định tuyến các yêu cầu sử dụng các ứng dụng, tài nguyên dùng chung từ các máy tính trạm của mạng đó. Ví dụ như máy chủ ứng dụng in ấn (Printer Server): Các máy chủ điều khiển in tập tin yêu cầu và gửi tập tin đó tới máy in được yêu cầu - Một dịch vụ in cho phép nhiều người cùng sử dụng một máy in trên mạng

- Máy chủ cơ sở dữ liệu (Database Server): Một dịch vụ cơ sở dữ liệu là một ứng dụng cơ bản trên mô hình kiến trúc máy chủ / máy trạm. Ứng dụng được chia làm hai phần: một phần chạy trên một máy trạm (nơi mà người sử dụng tích lũy và hiển thị thông tin cơ sở dữ liệu) và phần còn lại chạy trên máy chủ, nơi có nghĩa vụ như là kết nối dữ liệu và lưu trữ - được thực hiện.

- Máy chủ Web (Web Server): Ở phần lõi của nó, một dịch vụ web phục vụ nội dung tĩnh cho một trình duyệt bằng cách tải một tập tin từ đĩa và chuyển nó lên mạng, tới một người sử dụng trình duyệt web. Sự trao đổi hoàn toàn này được thực hiện gián tiếp thông qua một trình duyệt và một máy chủ kết nối tới một thiết bị khác sử dụng HTTP. Bất kỳ máy tính nào cũng có thể vào trong một dịch vụ web bằng cách cài đặt phần mềm dịch vụ và kết nối internet.

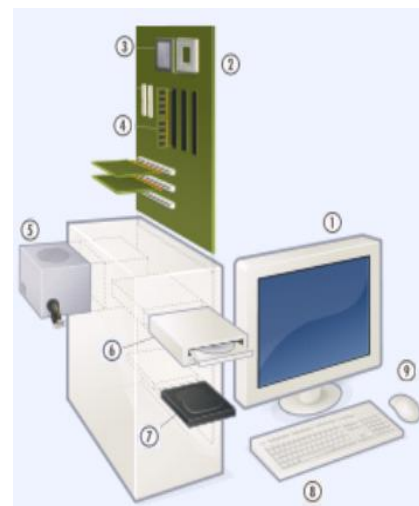


Hình 2.4 Mô hình làm việc giữa máy yêu cầu (máy trạm) và máy phục vụ (máy chủ)

2.2.1.3. Hệ thống máy trạm

Hệ thống máy trạm là hệ thống các máy tính hỗ trợ tin học hóa các công tác nghiệp vụ trong thư viện, hệ thống máy trạm bao gồm một số loại thông dụng như sau: máy trạm (workstation), máy tính cá nhân (personal computer), máy tính xách tay (notebook computer, laptop computer). Căn cứ theo chức năng, nhiệm vụ có thể chia máy trạm thành một số loại như sau:

- Máy trạm nghiệp vụ: Phục vụ công tác của các cán bộ thư viện như nghiệp vụ bổ sung, biên mục cũng như biên tập thông tin.
- Máy trạm phòng đọc multimedia:
Sử dụng phục vụ bạn đọc khai thác thông tin thư viện số cũng như khai thác ứng dụng multimedia.
- Máy trạm quản trị: Phục vụ công tác quản trị mạng của cán bộ quản trị thư viện.
- Máy trạm tra cứu: Phục vụ công tác tra cứu cho bạn đọc hoặc cán bộ thư viện



tra cứu tài liệu dựa trên các tiêu chí

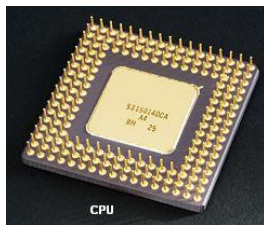
khác nhau. Ngày nay, khi khoa học kỹ thuật phát triển, người ta sử dụng các kiosk tra cứu thông tin thay cho các máy trạm tra cứu, thực chất các kiosk tra cứu thông tin cũng là một máy trạm tra cứu với nhiều tính năng đặc biệt như: màn hình cảm ứng, tích hợp CPU và màn hình ...

Các thành phần chính của phần cứng máy trạm bao gồm

1. Màn hình (Monitor)
2. Bo mạch chủ (Motherboard hay mainboard)
3. Bộ xử lý CPU (Central Processing Unit)
4. Bộ nhớ (RAM: Random Access Memory)
5. Nguồn máy tính (Power Supply Unit hay PSU), vỏ máy tính (case)
6. Ổ đĩa quang (Optical disk) (Ví dụ: CD, DVD ...)



Bo mạch chủ



CPU



RAM



7. Ổ đĩa cứng hay còn gọi là ổ cứng (HDD: Hard disk drive)
8. Bàn phím máy tính (Keyboard)
9. Chuột (Mouse)

2.2.1.4. Các thiết bị ngoại vi

- Máy in: là một thiết bị dùng để thể hiện ra các chất liệu khác nhau các nội dung được soạn thảo hoặc thiết kế sẵn. Hệ thống máy in trong thư viện chủ yếu sử dụng là máy in dùng trong văn phòng, gồm nhiều loại và công nghệ khác nhau:

- + Máy in laser: máy in sử dụng công nghệ la de



(laser) là loại máy in dùng in ra giấy, hoạt động dựa trên nguyên tắc dùng tia la de để chiếu lên một trống từ, trống từ quay qua ống mực (có tính chất từ) để mực hút vào trống, giấy chuyển động qua trống và mực được bám vào giấy, công đoạn cuối cùng là sấy khô mực để mực bám chặt vào giấy trước khi ra ngoài. Đây là loại máy in được sử dụng phổ biến hiện nay.

+ Máy in kim: máy in kim sử dụng các kim để chấm qua một băng mực làm hiện mực lên trang giấy cần in. Máy in kim đã trở thành lạc hậu do các nhược điểm: in rất chậm, độ phân giải của bản in rất thấp và khi làm việc chúng rất ồn.



+ Máy in phun: máy in phun hoạt động theo nguyên lý phun mực vào giấy in. Mực in được phun qua một lỗ nhỏ theo từng giọt với một tốc độ lớn (khoảng 5.000 lần /giây) tạo ra các điểm mực đủ nhỏ để thể hiện bản in sắc nét, đa số các máy in phun thường là các máy in màu.



So sánh trong các thể loại máy in thì máy in phun thường có chi phí trên mỗi bản in lớn nhất nên giá thành máy in phun thường thấp hơn máy in laser nhưng hộp mực cho máy in phun lại có giá cao, số lượng bản in trên bộ hộp mực thấp.

- Máy quét (Scanner): Là thiết bị ngoại vi dùng để chuyển đổi các tài liệu từ dạng tương tự sang dạng số phục vụ công tác lưu trữ, tra cứu, tìm kiếm và xây dựng các nguồn tin số hóa của thư viện



- Máy ảnh kỹ thuật số: là một máy điện tử dùng để thu và lưu giữ hình ảnh một cách tự động thay vì phải dùng phim ảnh giống như máy chụp ảnh thường. Những máy chụp ảnh số đời mới thường có nhiều chức năng như: ghi âm, quay phim



- Máy quay phim kỹ thuật số: là loại máy mà mục đích chính là để thu ảnh động, phổ biến hiện nay là máy quay phim kết hợp (camcorder) thường có sẵn microphone để ghi âm và một màn hình tinh thể lỏng nhỏ để xem trong khi ghi hình và phát lại. Máy quay phim kỹ thuật số là một trong những thiết bị quan trọng trong quá trình xây dựng kho tài nguyên thông tin số của thư viện.



2.2.1.5. Thiết bị mã vạch, từ

- Máy quét mã vạch: là một máy thu nhận hình ảnh của mã vạch in trên các bề mặt và chuyển thông tin chứa trong mã vạch đến máy tính hay các thiết bị cần thông tin này. Nó thường có một nguồn sáng kèm theo thấu kính để hội tụ ánh sáng lên mã vạch rồi thu ánh sáng phản xạ về một cảm quang chuyển hóa tín hiệu ánh sáng thành tín hiệu điện. Máy quét mã vạch được sử dụng trong công tác lưu thông và quản lý tài liệu tại các phòng chức năng:



- + Quét thẻ bạn đọc
- + Quét tài liệu

- Máy gom dữ liệu di động (thiết bị kiểm kê cầm tay): sử dụng trong công tác kiểm kê tài liệu tự động. Nhân viên thư viện sử dụng thiết bị kiểm kê để lướt qua các tài liệu có trên giá, thiết bị sẽ quét mã vạch của tài liệu sau đó xử lý và lưu trữ vào bộ nhớ trong của thiết bị. Cuối cùng được kết nối với hệ thống máy tính để xử lý dữ liệu đã được thu thập.

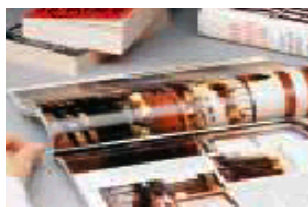
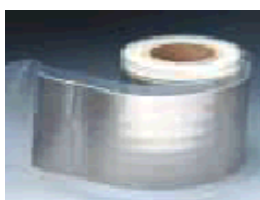


- Máy in mã vạch: là một loại máy in chuyên dụng, thường được kết nối với máy tính như một thiết bị ngoại vi để in mã vạch. Các máy in này thường được hỗ



trợ bằng phần mềm để người dùng lựa chọn kiểu dáng nhãn mã vạch, nội dung kèm theo, độ phân giải, loại và kích thước mã vạch cần in. Việc lựa chọn máy in mã vạch thường phụ thuộc các yêu cầu về công suất in ấn, điều kiện làm việc của máy in, chất lượng mã vạch để phù hợp với mục đích sử dụng. Trong thư viện máy in mã vạch được sử dụng in mã vạch dán cho tài liệu trong thư viện.

- Băng từ: là một loại băng từ trong suốt, đặc biệt dán lên sách, băng audio cassette, video cassette, CD ... để đảm bảo các tài liệu này không thể bị mang ra khỏi thư viện trừ khi đã được đăng ký mượn



- Thiết bị khử từ / nhiễm từ: dùng để khử từ khi độc giả đã đăng ký mượn, sau đó độc giả có thể mang sách ra ngoài mà không bị các hệ thống an ninh thư viện báo động. Nhiễm từ để tạo lại hệ thống từ khi tài liệu được mang trả lại thư viện



- Thiết bị nhiễm từ cho băng: thiết bị này làm nhiệm vụ nhiễm từ cho các tài liệu từ như băng audio/video cassette có dán băng từ tại quầy nhân viên



2.2.1.6. Thiết bị an ninh thư viện

a) Cổng từ

Là một thiết bị nhằm kiểm soát, phát hiện độc giả mang tài liệu ra khỏi thư viện khi chưa được phép của thư viện. Đây là một thiết bị điện từ trường nhưng an toàn tuyệt đối với các tài liệu lưu trữ dưới dạng từ chẳng hạn như băng cassette, băng video, đĩa mềm máy tính ...



Một số chức năng chính của cổng từ:

- Bảo vệ tuyệt đối an toàn cho tất cả các tài liệu thư viện đã được dán băng từ
- Phát ra các tín hiệu cảnh báo cả âm thanh và hình ảnh cho phép nhận biết chính xác độc giả nào đã mang qua tài liệu chưa được đăng ký mượn
- Có bộ đếm tự động ghi nhận và hiển thị lưu lượng người vào/ra thư viện
- Cho phép lựa chọn dạng âm thanh cảnh báo
- Có thể kết hợp với hệ thống Video bảo vệ đem lại sự bảo vệ suốt ngày đêm

b) Camera quan sát

Ngoài thiết bị cổng từ, hệ thống camera quan sát là hệ thống quan trọng, nhằm giám sát, quản lý và đảm bảo an toàn cho các kho mở.

Hệ thống Camera quan sát còn được gọi là hệ thống truyền hình mạch kín (CCTV) nhờ tính đồng bộ của hệ thống bao gồm từ việc thu hình ảnh, truyền hình ảnh và xử lý hình ảnh.

Hệ thống Camera quan sát được sử dụng để hỗ trợ cho việc giám sát vì mục đích an ninh, hoạt động của các cơ quan và mục đích sử dụng khác...

Ngày nay, hệ thống CCTV đã được tích hợp với các công nghệ mới giúp cho việc lưu trữ, xem lại cũng như sao lưu dữ liệu trở nên cực kỳ dễ dàng. Hệ thống còn có thể tích hợp vào mạng LAN hoặc Internet giúp

người điều khiển có thể thao tác từ xa vào hệ thống.

Các cấu thành cơ bản của một hệ thống CCTV gồm:

1. Camera: Camera là thành phần quan trọng nhất của hệ thống. Tùy theo mục đích giám sát, các khu vực cần theo dõi sẽ được lắp đặt các loại camera thích hợp:

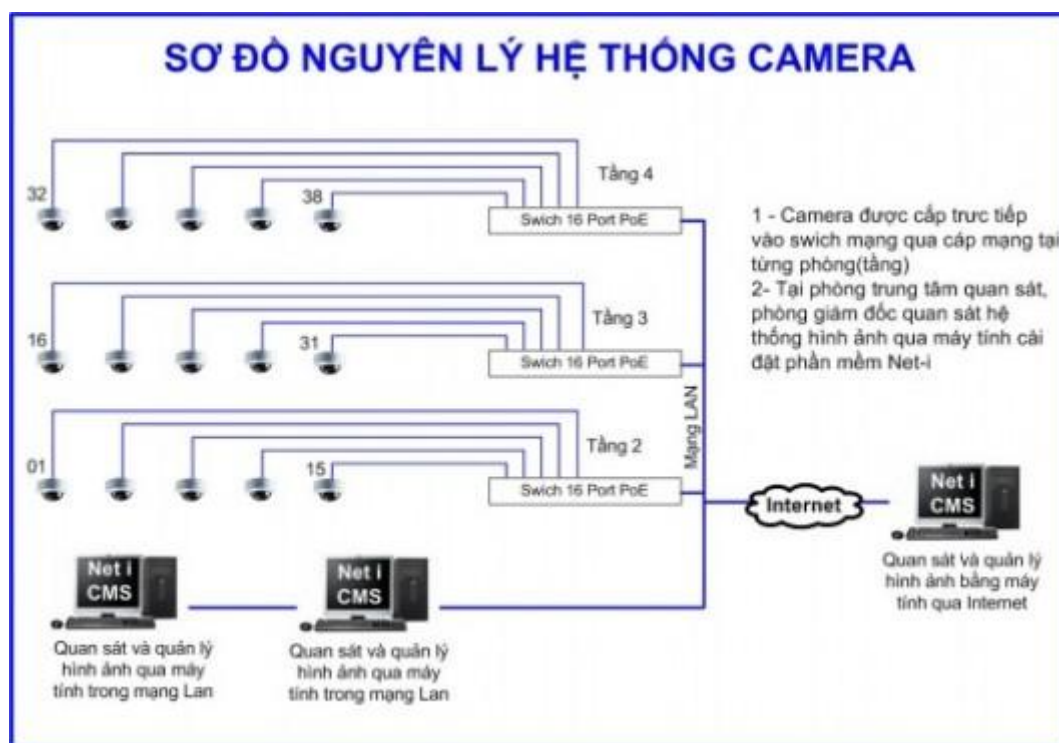
- Camera màu: Quan sát trong điều kiện ánh sáng bình thường
- Camera hồng ngoại, Camera day/night: Quan sát trong điều kiện ánh sáng yếu và trời tối.
- Camera quay quét: Camera có thể điều khiển giúp quan sát linh hoạt.
- Camera IP: Camera sử dụng giao thức TCP/IP để truyền tải hình ảnh, thuận tiện cho các mô hình giám sát phân tán.

2. Thiết bị lưu trữ hình ảnh (DVR-Digital Video Recorder): DVR là thành phần không thể thiếu của hệ thống CCTV giúp người quản lý có thể lưu trữ và xem lại dữ liệu hình ảnh các thời điểm nhất định. Nó cũng là thành phần trung gian liên kết giữa hệ thống CCTV với các hệ thống khác (hệ thống báo động, hệ thống kiểm soát vào ra, kết nối internet ...). Tùy theo mục đích sử dụng ta có các loại thiết bị ghi hình khác nhau, nhưng có thể phân vào các nhóm chính như sau:

- DVR: Đầu ghi hình chuyên dùng cho các hệ thống mở có thể có nhiều kết nối với các hệ thống khác và yêu cầu điều khiển hệ thống từ xa.
- Card ghi hình: Chuyên dùng để tích hợp hệ thống camera cùng với máy tính PC.
- NVR: Thiết bị ghi hình qua mạng.
- VCR: Thiết bị ghi hình sử dụng băng từ.

3. Thiết bị điều khiển: Thiết bị điều khiển là các thiết bị chuyên dùng để điều khiển các camera (camera quay quét, camera thân dài có đế quay quét...):

- Bàn phím điều khiển: chuyên dùng điều khiển camera quay quét, camera zoom, camera thân dài có đế quay...
- Ma trận Camera: Tích hợp cùng bàn phím điều khiển giúp điều khiển linh hoạt, thường sử dụng với các hệ thống lớn.



Hình 2.5 Sơ đồ nguyên lý hệ thống camera

2.2.1.7. Hệ thống RFID

a) Giới thiệu công nghệ RFID trong hoạt động thư viện

RFID (Radio Frequency Identification là công nghệ định danh bằng sóng Radio) là một sự kết hợp giữa công nghệ dựa trên sóng radio và chip điện tử. Đây là một công nghệ tiên tiến để kiểm soát tài liệu, nó có nhiều ưu điểm vượt trội so với công nghệ mã vạch. Khác với công nghệ mã vạch là công nghệ định danh trực diện (line-of-sight technology), nghĩa là để nhận dạng đối tượng, máy đọc cần phải tiếp xúc trực tiếp đối tượng ở khoảng cách gần. Đối với công nghệ RFID, có thể xác định đối tượng ở khoảng cách xa từ vài mét tới hàng trăm mét trong môi trường không gian 3 chiều (3D).

Ngày nay, các thư viện đang đối diện với những thách thức như sự

gia tăng không ngừng mật độ tại các điểm lưu thông và vốn tài liệu thư viện. Các thủ thư làm việc tại quầy lưu thông ngoài việc đảm bảo hiệu quả sử dụng nguồn tài nguyên thư viện, công việc hàng ngày của họ còn là tiếp xúc bạn đọc và cung cấp dịch vụ bạn đọc chất lượng cao thỏa mãn mọi nhu cầu của khách thăm quan cũng như bạn đọc của thư viện.

Công nghệ RFID đã và đang là một trong những giải pháp hữu hiệu để giải quyết các thách thức nêu trên. Với tính năng “3 trong 1”, lưu thông - an ninh - kiểm kê, RFID không những tối ưu hóa quỹ thời gian của cán bộ thư viện mà đặc biệt là đem lại sự thuận tiện và đảm bảo tính riêng tư của bạn đọc khi họ sử dụng quầy mượn trả tự động.

Ưu điểm vượt trội của RFID chính là tính năng kiểm kê khi cán bộ thư viện chỉ cần đi dọc theo giá sách mà không cần phải nhắc xuống (rời) đặt lên bất kỳ quyển sách nào và tính năng lưu thông nhiều tài liệu cùng một lúc (ví dụ: một chồng sách gồm 10 quyển, vài đĩa CD-ROM và băng video hoặc cát xet chỉ cần một lần quét duy nhất tại quầy lưu thông).

Ứng dụng RFID trong thư viện đã và đang đem đến những lợi trước mắt và lâu dài cho quy trình quản lý thư viện, “truy tìm dấu vết” của các tài liệu xếp sai vị trí, tự động mượn trả, gia tăng an ninh thư viện.

b) Các thành phần và chức năng của một hệ thống RFID

Một hệ thống RFID trong thư viện cơ bản bao gồm 2 thành phần chính:

- Phần cứng: Công kiểm soát vào/ra bằng công nghệ RFID, Trạm lập trình, Trạm lưu thông, Thiết bị kiểm kê cầm tay và Chíp/Thẻ RFID (RFID tag)
- Phần mềm: gồm có phần mềm trung gian (middle ware) và phần mềm ứng dụng (trong lĩnh vực thư viện nó là các phần mềm quản trị thư viện).

Phần cứng

1. Thẻ RFID (RFID tag)

Thẻ RFID được cấu tạo mềm mỏng có chứa chip vi xử lý và ăngten (đối với loại thẻ thụ động). Nó có thể đọc, ghi dữ liệu, và thậm chí có chứa cả thông tin về bảo mật. Thẻ này có thể dán vào các vật cần quản lý như sách, băng, đĩa v.v...

Thẻ này có kích thước mỏng như tờ giấy và có thể chứa dữ liệu ít nhất là 16 bits. Trong thẻ thường bao gồm các thông tin:

- Số ID (số nhận dạng thẻ)
- Nhan đề tài liệu
- Tác giả
-

Căn cứ vào khả năng cung cấp năng lượng phục vụ việc truyền tín hiệu, có thể chia thẻ RFID làm 3 loại:

- Thẻ thụ động (Passive tag): là loại thẻ không có nguồn cung cấp điện bên trong phục vụ cho việc truyền tín hiệu. Bên trong mỗi thẻ này có chứa một ăngten nhỏ. Nó sử dụng nguồn điện nhờ sóng radio phát ra từ ăngten của máy đọc. Nguồn điện này thông qua ăngten của thẻ sẽ kích hoạt thẻ hoạt động để gửi thông tin có trên thẻ ngược trở lại tới máy đọc. Phạm vi đọc của thẻ khá ngắn, khoảng từ 10 cm đến một vài mét phụ thuộc vào thiết kế/kích thước của anten và tần số sóng radio sử dụng. Do không có pin nên thẻ thụ động khá nhỏ 0.15 mm × 0.15 mm (mỏng bằng tờ giấy) do vậy có thể dễ dàng gắn vào các vật dụng. Giá thành của loại này vì thế rẻ hơn so với các thẻ khác. Dung lượng bộ nhớ của thẻ khoảng từ 16 - hơn 512 bits.

Thẻ thụ động thường hoạt động ở tần số thấp, cao hoặc siêu cao. Tần số sẽ xác định các đặc tính hoạt động của thẻ, bao gồm khoảng cách mà bộ đọc có thể đọc được thẻ. Phạm vi đọc được của các loại thẻ phổ biến như sau:

Tần số	Thấp (Low) 128 KHz	Cao (High) 13.56MHz	Siêu cao (Ultra-high) 915MHz
Phạm vi	0- 152mm	0- 0.91m	0- 4.5m

Hiện tại, các hệ thống thư viện sử dụng thẻ ở tần số cao (HF) vì chức năng và phạm vi đọc của thẻ. Phạm vi đọc gần sẽ tiện lợi trong việc kiểm tra bằng các thiết bị mượn trả tự động và công an ninh nhưng lại không đọc tốt tài liệu để sát thành giá kệ. Tuy nhiên, một số thư viện lại quan tâm đến những thẻ hoạt động ở tần số siêu cao (UHF) với phạm vi đọc xa hơn, ở tần số này sẽ dễ dàng hơn trong công tác quản lý tài liệu trên giá, lối đi giữa các công an ninh sẽ được nới rộng hơn. Thẻ tần số siêu cao (UHF) có thể phát huy một số ưu điểm ở một vài ứng dụng nhất định nơi mà thẻ tần số cao (HF) không có khả năng thực hiện.

Những thẻ dùng trong thư viện phải được làm bằng những vật liệu tốt và phải trải qua những quy trình chế tạo nghiêm ngặt để đảm bảo rằng vòng đời của nó phải dài tương ứng với vòng đời của sản phẩm mà nó bảo vệ. Trong hầu hết các trường hợp, loại thẻ này có giá cao hơn những loại tem sử dụng trong hệ thống cung cấp hàng hóa.

- Thẻ bán thụ động (Semi-passive tag): thẻ này gần giống với thẻ thụ động ngoại trừ nó có một quả pin nhỏ để cung cấp điện. Do có nguồn điện cung cấp vì vậy ăngten có trong thẻ không cần thực hiện chức năng nhận tín hiệu như thẻ thụ động. Ăngten này được sử dụng cho việc phản hồi thông tin do vậy việc truyền thông tin nhanh hơn so với thẻ thụ động. Ngoài ra thẻ bán thụ động cũng dễ đọc được trong các môi trường có kim loại hoặc chất lỏng hơn so với thẻ thụ động.
- Thẻ chủ động (Active tag): khác với thẻ thụ động và bán thụ động, thẻ chủ động có nguồn cung cấp điện riêng do vậy nó có thể tự phát sóng và truyền thông tin. Do có nguồn điện tự cung cấp vì vậy mà

nó hoạt động hiệu quả hơn trong các môi trường “kì” với sóng radio như môi trường nước, kim loại. Nhiều thẻ chủ động cho phép đọc ở khoảng cách khá xa tới vài trăm mét và tuổi thọ của pin có thể lên tới 10 năm. Thẻ active còn cho phép lưu trữ được thông tin nhiều hơn so với 2 loại thẻ trên (khoảng 512 Kb hoặc hơn). Kích cỡ của thẻ chủ động chỉ bằng 1 đồng xu. Tuy nhiên, giá của nó khá cao. Nó được ứng dụng chủ yếu để quản lý các loại hàng hóa đắt tiền ví dụ như quân đội Mỹ sử dụng loại thẻ này để theo dõi các công te nơ trong cảng.

Căn cứ vào số lần đọc và ghi của thẻ, có thể chia thẻ RFID có 3 loại: thẻ cho phép đọc viết nhiều lần, thẻ chỉ đọc và thẻ viết một lần và đọc nhiều lần (WORM)

- Thẻ cho phép đọc và viết: với thẻ này người dùng có thể thêm thông tin hoặc viết đè lên thông tin đang có trong thẻ nhiều lần tới khi thẻ không sử dụng được. Trong mỗi thẻ này có một số serial và không cho phép được xóa hoặc viết đè. Hầu hết các thư viện hiện nay sử dụng thẻ không năng cho phép đọc và ghi.

- Thẻ chỉ đọc: chứa thông tin do nhà sản xuất đưa vào trong quá trình sản xuất thẻ. Người dùng không thể thêm hoặc thay đổi thông tin trong thẻ.

- Thẻ viết một lần và đọc nhiều lần (còn gọi là thẻ WORM): với thẻ này người dùng chỉ có thể thêm số serial của thẻ trong lần đầu tiên, lần tiếp sau nó chỉ cho phép đọc.



2. Thiết bị kiểm kê cầm tay

Thiết bị kiểm kê cầm tay được sử dụng trong quá trình kiểm kê sách hiện đang xếp giá trong thư viện. Thông qua đó có thể biết được số lượng sách đã cho mượn, số lượng sách đang xếp giá. Nhân viên thư viện sử dụng thiết bị



kiểm kê cầm tay để lướt qua các tài liệu có trên giá. Ăngten trong thiết bị sẽ nhận tín hiệu của các chip có trong các tài liệu này và truyền thông tin đến bộ đọc, bộ đọc xử lý dữ liệu và truyền các thông tin thu được đến màn hình để hiển thị tương ứng.

3. Trạm lập trình, trạm lưu thông

Trạm lập trình hoạt động như 1 trạm thu phát trao đổi và ghi dữ liệu giữa thẻ RFID và hệ thống quản lý thư viện số tích hợp. Cán bộ thư viện sử dụng trạm lập trình để ghi thông tin mới của tài liệu cho các thẻ trắng (thẻ chưa có dữ liệu), hoặc có thể nhập các thông tin từ hệ thống mã vạch (barcode). Điều này là rất tiện lợi và tiết kiệm thời gian cho các thư viện đang sử dụng hệ thống quản lý bằng mã vạch. Tất cả các thông tin lập trình này cũng sẽ được ghi lại vào cơ sở dữ liệu chung của thư viện và được quản lý bằng hệ thống thư viện số tích hợp

Trạm lưu thông là hệ thống giúp thủ thư thực hiện các thao tác mượn/trả sách, kiểm tra thông tin cho bạn đọc. Bạn đọc có thể vào thư viện, chọn xem sách, và khi có nhu cầu mượn thì sẽ mang ra quầy lưu thông được bố trí gần cửa ra vào. Cán bộ thủ thư sẽ cho nhận diện sách bằng ăngten RFID để trên bàn (gắn trong trạm lưu thông). Ăng ten có khả năng nhận diện lên tới 16 tài liệu cùng 1 lúc. Tất cả các tài liệu đặt trên ăng ten sẽ được nhận diện và hiển thị trên màn hình máy tính (kết nối với trạm lưu thông), bao gồm tất cả các thông tin về tài liệu như tựa đề, tác giả, nhà xuất bản, mã số ... Nhân viên thư viện sau khi đã kiểm tra thẻ bạn đọc chỉ việc ấn nút cho mượn trên màn hình máy tính đối với bạn đọc này cho tất cả các tài liệu đó. Chức năng kích hoạt chống trộm sẽ tự động được bỏ đi và bạn đọc có thể mang tài liệu ra khỏi thư viện mà công an ninh không báo động. Ngược lại với quy trình mượn tài liệu, trong quy trình trả tài liệu, bạn đọc sẽ mang tài liệu đã mượn trước đó của thư viện để trả tại quầy lưu thông. Cán bộ thủ thư tiến hành đặt các tài liệu lên trên ăng ten của trạm lưu thông để trạm tự động nhận diện các

tài liệu được mang trả và nổi thông tin của bạn đọc thông qua thẻ bạn đọc. Sau đó cán bộ thủ thư chỉ việc ấn nút trả tài liệu trên màn hình máy tính. Sách đã ở trong tình trạng được trả vào thư viện và tự động được kích hoạt chức năng chống trộm. Tất cả các thông tin trong quá trình lưu thông sẽ được ghi lại vào cơ sở dữ liệu chung của thư viện và được quản lý bằng phần mềm thư viện số tích hợp.

4. Cổng kiểm soát vào ra bằng công nghệ RFID (Cổng an ninh)

Tài liệu sẽ được bảo vệ an ninh bằng thẻ RFID. Cổng an ninh sẽ hoạt động theo nguyên lý tần số sóng vô tuyến (Radio Frequency), cổng sẽ kiểm tra tất cả các tài liệu chưa check-out (đã check-in), cổng sẽ phát ra tín hiệu kích hoạt thẻ RFID đang đi vào vùng hoạt động, thẻ RFID sẽ phản hồi lại một tín hiệu, nếu tài liệu chưa được làm thủ tục check-out, cổng an ninh sẽ phát báo động bằng âm thanh, ánh sáng.



Ngoài các thiết bị cơ bản kể trên, còn một số các thiết bị mở rộng khác cho thư viện như: Hệ thống mượn/ trả sách tự động, hệ thống phân loại sách tự động



Hình 2.6 Thiết bị mượn, trả sách tự động

Phần mềm

1. Phần mềm trung gian (Middleware)

Phần mềm trung gian là các phần mềm được sử dụng để nhận và xử

lý các dữ liệu thô nhận được từ các trạm lập trình để chuyển tới các phần mềm thư viện số tích hợp. Đây là một thành phần không thể thiếu trong hệ thống RFID. Phần mềm này thường được xây dựng và cung cấp bởi các nhà cung cấp thiết bị RFID.

2. Phần mềm ứng dụng (Application software)

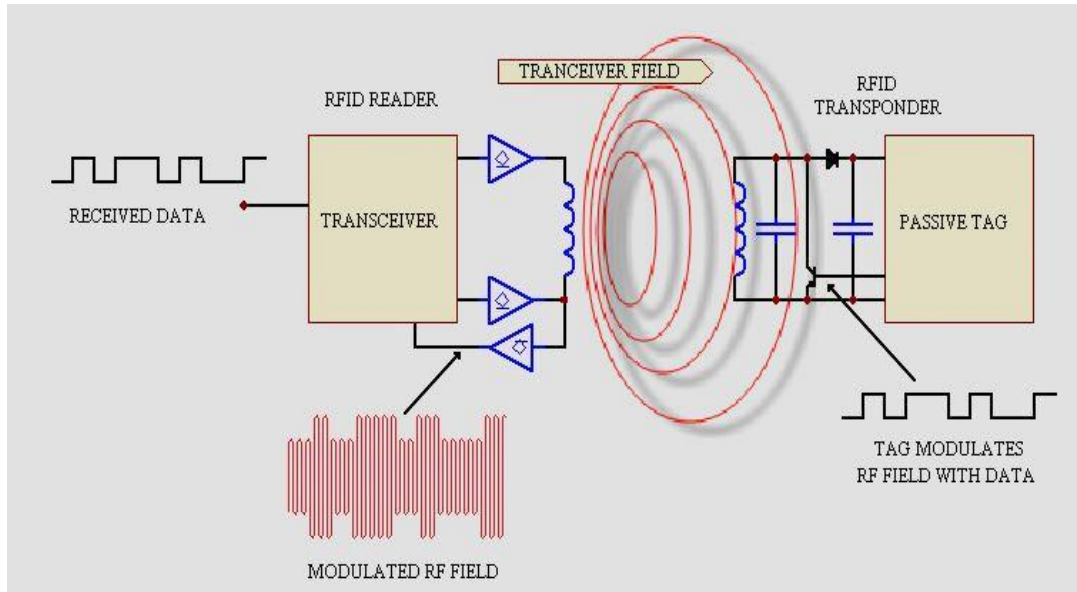
Là các phần mềm được sử dụng để xử lý và tự động hóa các công việc của một cơ quan, tổ chức. Trong lĩnh vực thư viện, phần mềm ứng dụng chính là các phần mềm thư viện số tích hợp trong đó có lưu thông tin về các tài liệu, bạn đọc, quá trình mượn/trả, hệ thống kho... mà thư viện quản lý. Phần mềm này sẽ nhận dữ liệu đã được xử lý từ phần mềm trung gian để phân tích.

Để giao tiếp giữa hệ thống quản trị thư viện tích hợp với các ứng dụng của RFID người ta sử dụng chuẩn SIP2 (Standard Interface Protocol, version 2) do 3M xây dựng hoặc NCIP (ANSI/NISO Z39.83 – 2002 Circulation Interchange), trong đó SIP 2 được sử dụng phổ biến hơn. Mỗi nhà cung cấp ứng dụng RFID sử dụng các chuẩn khác nhau cho ứng dụng của mình.

Hệ thống RFID sử dụng nhiều tần số khác nhau, nhưng nói chung chủ yếu có 3 mức: tần số thấp (khoảng 125 Khz), tần số cao (13.56Mhz), và tần số siêu cao (khoảng 860 – 960 Mhz). Trong lĩnh vực thư viện người ta sử dụng tần số sóng radio cao (13.56 Mhz). Vì vậy để máy trạm lưu thông và thẻ có thể giao tiếp được với nhau thì hai thiết bị này phải cùng tần số.

c) Nguyên tắc hoạt động của hệ thống RFID

Khi tài liệu có gắn thẻ RFID đi qua vùng ăngten phát sóng radio thì thẻ sẽ được tự động kích hoạt và gửi thông tin trở lại ăngten. Ăngten truyền tín hiệu tới máy đọc để giải mã thông tin rồi chuyển tới các phần mềm trung gian và ứng dụng để xử lý. Người sử dụng sẽ nhận được thông tin đã được xử lý thông qua màn hình máy tính.



Hình 2.7 Mô hình nguyên lý hoạt động của hệ thống RFID

d) Mô hình hệ thống RFID trong thư viện

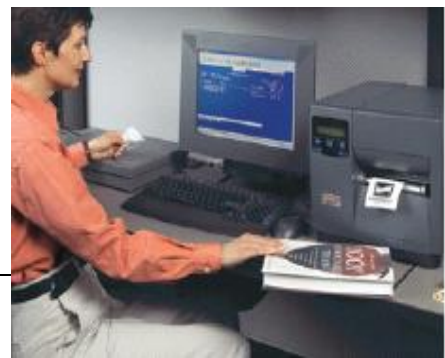
Bao gồm các công đoạn nhập thông tin vào thẻ, Mượn/trả tài liệu, Phân loại tài liệu tự động, Kiểm kê kho. Thông tin ở mỗi công đoạn này đều được cập nhật vào hệ thống quản trị thư viện số tích hợp.



Hình 2.8 Mô hình hệ thống RFID trong thư viện

1. Nhập thông tin vào thẻ

Mỗi tài liệu cần quản lý sẽ được dán một thẻ RFID. Trong thẻ RFID chứa các thông tin



về đối tượng mà nó được dán lên. Khác với công nghệ mã vạch, thông tin lưu trên mỗi mã vạch chỉ có duy nhất một thông tin đó là số đăng ký cá biệt của tài liệu hoặc số nhận dạng tài liệu. Thông tin lưu trên các thẻ RFID có thể là số hiệu biểu ghi, nhan đề tài liệu, tác giả, số đăng ký cá biệt,... tùy thuộc vào mục đích quản lý tài liệu của thư viện và khả năng lưu trữ của thẻ.

Các thiết bị để nhập thông tin bao gồm: máy tính chứa phần mềm ứng dụng, trạm lập trình và tài liệu đã được dán thẻ. Để đưa các thông tin này vào thẻ, người ta nhập các thông tin thư mục lên phần mềm chuyên dụng hoặc lấy thông tin đã có sẵn trong cơ sở dữ liệu của phần mềm thư viện số tích hợp. Máy tính có chứa phần mềm chuyên dụng kết nối với trạm lập trình sẽ truyền dữ liệu vào thẻ dán trên tài liệu.

2. Mượn/trả sách

Khi ứng dụng công nghệ RFID, quá trình mượn/trả tài liệu có thể được tiến hành bằng 2 cách: Mượn/Trả tại bàn hoặc Mượn /trả tự động.

(1) Mượn/Trả tại bàn: khi tiến hành mượn/trả một tài liệu bạn đọc tới trực tiếp bàn của thủ thư. Quá trình mượn/trả gần giống với mượn/trả sử dụng công nghệ mã vạch. Thủ thư sẽ dùng máy đọc để nhận biết thông tin trên thẻ bạn đọc và trên tài liệu để ghi nhận một phiên mượn vào phần mềm ứng dụng.

(2) Mượn/Trả tự động: đây là một ưu điểm vượt bậc so với công nghệ sử dụng mã vạch, khi mượn/trả một tài liệu bạn đọc sẽ không cần đến sự hỗ trợ của thủ thư. Cách này thường được áp dụng đối với kho sách tổ chức theo hình thức kho mở. Để tiến hành mượn một tài liệu, bạn đọc chỉ cần mang thẻ bạn đọc và tài liệu tới các máy mượn sách. Thao tác đầu tiên là đưa thẻ vào máy để máy nhận biết thông tin của người mượn sau đó bạn đọc để sách lên máy



để anten của máy đọc kích hoạt thẻ gửi thông tin về tài liệu tới bộ đọc. Thông tin về bạn đọc và tài liệu sẽ được các phần mềm trung gian và phần mềm thư viện số xử lý sau đó hiện thông tin lên màn hình để người mượn theo dõi. Nếu thông tin mượn trên màn hình là đúng thì người mượn thực hiện thao tác chấp nhận mượn và sẽ nhận được một tờ giấy biên nhận do máy tính in ra.

Khi tiến hành trả sách bạn đọc chỉ cần đi tới trạm trả sách tự động và cho sách vào khoang trả. Trong trạm trả sách có gắn thiết bị đọc tích hợp anten sẽ nhận dữ liệu có từ thẻ gắn trong sách để chuyển dữ liệu tới phần mềm thư viện số. Nếu quyền sách đó được quyền mượn, một phiên mượn được kết thúc và bạn đọc có thể mang sách đi qua cổng an ninh. Một trạm trả sách tự động có thể thực hiện được khoảng 20.000 lượt trả mỗi tháng.

3. Kiểm kê tự động

Khi tiến hành kiểm kê, thủ thư sử dụng thiết bị kiểm kê di động cho phép lưu dữ liệu kiểm kê. Thủ thư sẽ đi đến từng giá sách và quét lên từng quyền sách. Việc tập hợp thông tin từ sách sử dụng công RFID khá nhanh và không cần phải di sát thiết bị kiểm kê vào từng quyền sách hoặc di



chuyển sách ra khỏi giá sách. Với việc sử dụng công nghệ không dây, kết nối dữ liệu kiểm kê với trực tiếp máy chủ sử dụng phần mềm thư viện số tích hợp, nó cho phép không chỉ cập nhật thông tin về kiểm kê mà còn cho phép biết được ngay vị trí đúng của tài liệu trên giá sách.

4. Chống trộm:

Hệ thống công an ninh sử dụng công nghệ RFID có gắn một bộ cảm ứng phát ra sóng radio khi thẻ RFID đi qua vùng phủ sóng của nó và nhận thông tin từ thẻ chuyển tới phần mềm trung gian và phần mềm ứng dụng để xử lý. Trường hợp tài liệu mượn chưa hoàn thành thủ tục mượn

tại bàn thủ thư hoặc tại hệ thống mượn tự động, nghĩa là, thông tin tài liệu mượn chưa được hệ thống thư viện số chấp nhận mượn thì khi bạn đọc mang tài liệu ra khỏi thư viện, cổng an ninh sẽ phát ra tín hiệu báo động.

5. Phân loại tài liệu tự động

Trong mỗi thẻ RFID có chứa thông tin về môn loại và kết hợp với hệ thống phân loại tự động giúp cho việc phân loại và sắp xếp tài liệu ở mức sơ bộ. Nó hỗ trợ đắc lực cho các thư viện tổ chức theo kho mở.

e) Ưu điểm và nhược điểm của hệ thống RFID

Ưu điểm

- Không cần tiếp xúc trực diện với tài liệu: khác với công nghệ mã vạch, để nhận dạng tài liệu cần phải tiếp xúc trực tiếp giữa tài liệu và thiết bị đọc. Đối với công nghệ RFID, cho phép máy đọc có thể nhận dạng được tài liệu ở khoảng cách khá xa từ vài mét đến 100 mét tùy từng loại thẻ.

- Kết hợp giữa chức năng an ninh và chức năng nhận dạng tài liệu (thông qua các thông tin có trong thẻ: số đăng ký cá biệt, môn loại,...): đối với hệ thống sử dụng mã vạch, mỗi nhãn mã vạch chỉ cho phép nhận dạng tài liệu còn để chống trộm tài liệu thì người ta phải sử dụng chỉ từ. Như vậy, mỗi tài liệu đều được gắn cả nhãn mã vạch và chỉ từ. Đối với các hệ thống RFID, mỗi thẻ RFID đã đảm nhiệm được cả 2 chức năng này: chức năng an ninh và nhận dạng tài liệu.

- Mượn/Trả nhanh chóng: khả năng đọc thông tin từ thẻ RFID nhanh hơn nhiều so với sử dụng barcode vì vậy nó làm cho thao tác mượn trả tài liệu cũng nhanh hơn. Nó hỗ trợ tối đa việc tự động hóa quá trình mượn/trả tài liệu: hỗ trợ mượn/trả không có sự can thiệp của thủ thư. Bạn đọc có thể mượn cùng một lúc với nhiều tài liệu một cách dễ dàng, việc này ở hệ thống sử dụng mã vạch không thể nào thực hiện được.

- Phân loại tự động tài liệu: mỗi thẻ RFID cho phép lưu nhiều thông tin khác nhau trong đó có môn loại của tài liệu. Vì vậy, các nhà sản xuất thiết bị RFID đã chế tạo ra loại máy giúp cho việc phân loại tài liệu tự động sơ bộ.

- Khả năng chống trộm tốt: các hệ thống an ninh sử dụng cổng từ, khả năng phát hiện tài liệu chỉ trong một khoảng cách ngắn và chỉ trong không gian 2 chiều. Do vậy, nếu một bạn đọc khi cầm sách cao hơn chiều cao của cổng từ thì cổng từ sẽ không phát hiện được. Với hệ thống an ninh sử dụng công nghệ RFID, nó có khả năng phát hiện tài liệu với khoảng cách xa và trong không gian 3 chiều vì vậy khả năng chống trộm của nó an toàn và đáng tin cậy hơn hệ thống an ninh sử dụng cổng từ.

- Kiểm kê nhanh chóng: với hệ thống RFID, nó có khả năng quét và nhận thông tin từ các quyển sách một cách nhanh chóng mà không cần phải đi sát máy gom di động vào sách hoặc di chuyển sách ra khỏi giá. Với việc sử dụng công nghệ không dây, kết nối dữ liệu kiểm kê với trực tiếp máy chủ sử dụng phần quản trị thư viện, nó cho phép không chỉ cập nhật thông tin về kiểm kê mà còn cho phép biết được ngay vị trí đúng của tài liệu trên giá sách.

- Độ bền của thẻ cao: độ bền của thẻ RFID cao hơn so với mã vạch bởi vì nó không tiếp xúc trực tiếp với các thiết bị khác. Các nhà cung cấp RFID cho rằng mỗi thẻ RFID có thể sử dụng ít nhất được 100.000 lượt mượn/trả trước khi nó bị hỏng.

Nhược điểm

- Chi phí đầu tư cao: đây là nhược điểm lớn nhất cản trở việc ứng dụng RFID. Thẻ RFID có giá khoảng 0.8-1\$/thẻ thụ động, trong khi đó một nhãn mã vạch chuyên dụng khoảng 100 đồng, nghĩa là gấp hơn 150 lần. Ngoài ra, các trang thiết bị khác như máy đọc, cổng an ninh, thiết bị kiểm kê cầm tay, phần mềm, hệ thống chuyển đổi từ hệ thống mã vạch sang hệ thống RFID còn khá cao.

- Dễ bị ăn cắp thẻ: đối với hệ thống an ninh sử dụng công nghệ từ, thông thường nhãn từ sẽ được gắn vào trong gáy sách vì vậy bạn đọc muốn bóc nhãn từ này ra cũng sẽ gặp khó khăn hơn. Các thẻ RFID được gắn lộ ra ngoài ở phần bên trong của trang bìa sau của sách vì vậy nó rất dễ bị bạn đọc bóc ra.

- Không đọc được: máy đọc sẽ không thể nhận được tín hiệu từ tài liệu nếu nó được bọc kín bên trong bởi lớp kim loại. Đây là điểm yếu của của RFID mà dễ bị kẻ trộm khai thác.

2.2.1.8. Hệ thống lưu trữ dữ liệu

Cùng với sự phát triển CNTT và sự bùng nổ về dữ liệu, một nhu cầu xuất hiện là việc bảo quản, lưu trữ các số liệu một cách an toàn và hiệu quả. Chính vì vậy các giải pháp lưu trữ dữ liệu hiện đại đã ra đời nhằm đáp ứng được nhu cầu của người dùng.

Do dung lượng dữ liệu gia tăng không ngừng, yêu cầu ngày càng cao về hiệu năng truy xuất, tính ổn định và sự sẵn sàng của dữ liệu; việc lưu trữ đã và đang trở nên rất quan trọng. Lưu trữ dữ liệu không còn đơn giản là cung cấp các thiết bị lưu trữ dung lượng lớn mà còn bao gồm cả khả năng quản lý, chia sẻ cũng như sao lưu và phục hồi dữ liệu trong mọi trường hợp.

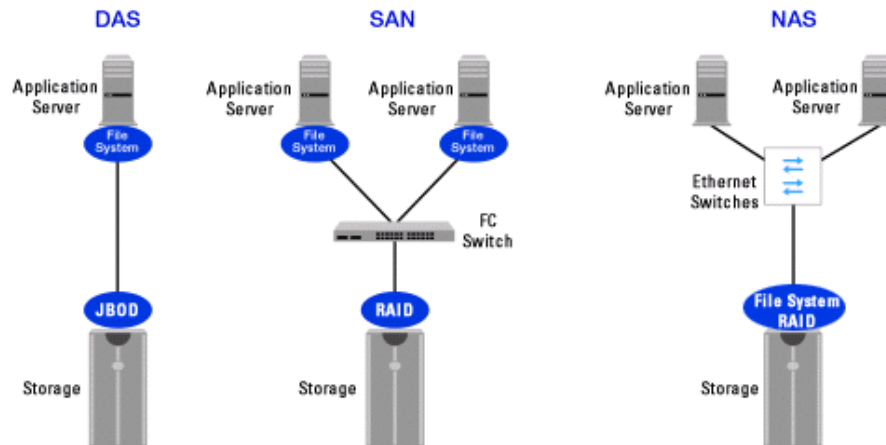
Để lưu trữ dữ liệu người ta có thể dùng nhiều thiết bị khác nhau, nhiều công nghệ khác nhau. Các kho dữ liệu có thể là dùng đĩa cứng, dùng băng từ, dùng đĩa quang... Tùy theo yêu cầu cụ thể của bài toán đặt ra mà lựa chọn công nghệ và thiết bị cho phù hợp. Theo cơ chế lưu trữ, hiện nay có một số loại hình lưu trữ dữ liệu cơ bản như:

- DAS (Direct Attached Storage): lưu trữ dữ liệu qua các thiết bị gắn trực tiếp

- NAS (Network Attached Storage): lưu trữ dữ liệu vào thiết bị lưu trữ thông qua mạng IP

- SAN (Storage Area Network): lưu trữ dữ liệu qua mạng lưu trữ chuyên dụng riêng.

Mỗi loại hình lưu trữ dữ liệu có những ưu nhược điểm riêng và được dùng cho những mục đích nhất định. Dưới đây là mô hình lưu trữ dữ liệu tổng quát.



Hình 2.9 Mô hình lưu trữ dữ liệu tổng quát.

a) Giải pháp lưu trữ truyền thống - DAS

DAS (Direct attached storage) là cơ chế lưu trữ với thiết bị gắn trực tiếp vào máy chủ. Đây được coi là công nghệ lưu trữ truyền thống được nhiều doanh nghiệp sử dụng. Với cơ chế DAS, mỗi máy chủ sẽ có một hệ thống lưu trữ và phần mềm quản lý lưu trữ riêng biệt.

Ưu điểm của giải pháp DAS là khả năng dễ lắp đặt, chi phí thấp, hiệu năng cao.

Tuy nhiên, nhược điểm của DAS là khả năng mở rộng hạn chế. Thực tế DAS làm việc rất tốt với một server nhưng khi dữ liệu tăng, số lượng máy chủ cũng tăng sẽ tạo nên những vùng dữ liệu phân tán và gián đoạn. Khi đó, nhà quản trị sẽ phải bổ sung hay thiết lập lại dung lượng, và công việc bảo trì sẽ phải thực hiện trên từng server. Điều đó sẽ làm tăng chi phí lưu trữ tổng thể cho doanh nghiệp và sẽ càng khó khăn hơn khi muốn sao lưu hay bảo vệ một hệ thống kho lưu trữ dữ liệu đang nằm rải rác và phân tán như vậy.

b) Giải pháp lưu trữ theo công nghệ NAS

NAS (Network Attached Storage) là phương pháp lưu trữ dữ liệu sử dụng các thiết bị lưu trữ đặc biệt gắn trực tiếp vào mạng LAN như một thiết bị mạng bình thường (tương tự máy tính, switch hay router). Các thiết bị NAS cũng được gán các địa chỉ IP cố định và được người dùng truy nhập thông qua sự điều khiển của máy chủ. Trong một số trường hợp, NAS có thể được truy cập trực tiếp không cần có sự quản lý của máy chủ.

Trong môi trường đa hệ điều hành với nhiều máy chủ khác nhau, việc lưu trữ dữ liệu, sao lưu và phục hồi dữ liệu, quản lý hay áp dụng các chính sách bảo mật đều được thực hiện tập trung.

Ưu điểm của NAS

- Khả năng mở rộng: khi người dùng cần thêm dung lượng lưu trữ, các thiết bị lưu trữ NAS mới có thể được bổ sung và lắp đặt vào mạng.
- NAS cũng tăng cường khả năng chống lại sự cố cho mạng. Trong môi trường DAS, khi một máy chủ chứa dữ liệu không hoạt động thì toàn bộ dữ liệu đó không thể sử dụng được. Trong môi trường NAS, dữ liệu vẫn hoàn toàn có thể được truy nhập bởi người dùng. Các biện pháp chống lỗi và dự phòng tiên tiến được áp dụng để đảm bảo NAS luôn sẵn sàng cung cấp dữ liệu cho người sử dụng.

Nhược điểm của NAS

- Với việc sử dụng chung hạ tầng mạng với các ứng dụng khác, việc lưu trữ dữ liệu có thể ảnh hưởng đến hiệu năng của toàn hệ thống (làm chậm tốc độ của LAN), điều này đặc biệt đáng quan tâm khi cần lưu trữ thường xuyên một lượng lớn dữ liệu.
- Trong môi trường có các hệ cơ sở dữ liệu thì NAS không phải là giải pháp tốt vì các hệ quản trị cơ sở dữ liệu thường lưu dữ liệu dưới dạng block chứ không phải dưới dạng file nên sử dụng NAS sẽ không cho hiệu năng tốt.

c) Giải pháp SAN

SAN (Storage Area Network) là một mạng riêng tốc độ cao dùng cho việc truyền dữ liệu giữa các máy chủ tham gia vào hệ thống lưu trữ cũng như giữa các thiết bị lưu trữ với nhau. SAN cho phép thực hiện quản lý tập trung và cung cấp khả năng chia sẻ dữ liệu và tài nguyên lưu trữ. Hầu hết mạng SAN hiện nay dựa trên công nghệ kênh cáp quang, cung cấp cho người sử dụng khả năng mở rộng, hiệu năng và tính sẵn sàng cao.

Hệ thống SAN được chia làm hai mức: mức vật lý và logic

- Mức vật lý: mô tả sự liên kết các thành phần của mạng tạo ra một hệ thống lưu trữ đồng nhất và có thể sử dụng đồng thời cho nhiều ứng dụng và người dùng.

- Mức logic: bao gồm các ứng dụng, các công cụ quản lý và dịch vụ được xây dựng trên nền tảng của các thiết bị lớp vật lý, cung cấp khả năng quản lý hệ thống SAN.

Ưu điểm của hệ thống SAN

- Có khả năng sao lưu dữ liệu với dung lượng lớn và thường xuyên mà không làm ảnh hưởng đến lưu lượng thông tin trên mạng.

- SAN đặc biệt thích hợp với các ứng dụng cần tốc độ và độ trễ nhỏ ví dụ như các ứng dụng xử lý giao dịch trong ngành ngân hàng, tài chính.

- Dữ liệu luôn ở mức độ sẵn sàng cao.

- Dữ liệu được lưu trữ thống nhất, tập trung và có khả năng quản lý cao. Có khả năng khôi phục dữ liệu nếu có xảy ra sự cố.

- Hỗ trợ nhiều giao thức, chuẩn lưu trữ khác nhau như: iSCSI, FCIP, DWDM...

- Có khả năng mở rộng tốt trên cả phương diện số lượng thiết bị, dung lượng hệ thống cũng như khoảng cách vật lý.

- Mức độ an toàn cao do thực hiện quản lý tập trung cũng như sử dụng các công cụ hỗ trợ quản lý SAN.

Nhược điểm của hệ thống SAN

- Giá thành cao.

Do đó, SAN thường được sử dụng ở những trung tâm dữ liệu lớn vì mang một số đặc điểm nổi bật như: Giảm thiểu rủi ro cho dữ liệu, khả năng chia sẻ tài nguyên rất cao, khả năng phát triển dễ dàng, thông lượng lớn, hỗ trợ nhiều loại thiết bị, hỗ trợ và quản lý việc truyền dữ liệu lớn và tính an ninh dữ liệu cao.

Hơn nữa, SAN tăng cường hiệu quả hoạt động của hệ thống bằng việc hỗ trợ đồng thời nhiều hệ điều hành, máy chủ và các ứng dụng, có khả năng đáp ứng nhanh chóng với những thay đổi về yêu cầu hoạt động của một tổ chức cũng như yêu cầu kỹ thuật của hệ thống mạng

Ngày nay, nhiều thư viện ở Việt Nam đã trang bị hệ thống SAN để lưu trữ dữ liệu

2.2.2. Phần mềm ứng dụng

Có thể khẳng định rằng, phần mềm thư viện ra đời và trở thành công cụ quan trọng nhất để cấu thành nên một thư viện hiện đại. Phần mềm thư viện có khả năng tự động hóa hoàn toàn các chức năng và quy trình nghiệp vụ trong thư viện, là công cụ hiệu quả để xây dựng các cơ sở dữ liệu thư mục, dữ liệu số, kiểm soát chất lượng biên mục các dữ liệu được nhập vào... Ngoài ra, nó còn giúp giảm thiểu các công việc trùng lặp, tái sử dụng các kết quả của những bộ phận có liên quan. Vai trò đặc trưng của phần mềm quản trị thư viện được thể hiện ở một số điểm sau:

- Tính quản trị cao: Là công cụ giúp nhà quản lý nắm bắt được tổng thể nguồn tài nguyên của đơn vị, từ đó xây dựng kế hoạch, chính sách phù hợp để phát triển thư viện
- Độ chính xác và chuẩn hóa: Phần mềm thư viện được xây dựng gắn liền với các chuẩn nghiệp vụ thư viện, tất cả các dữ liệu được nhập vào đều được sự kiểm soát chặt chẽ của hệ thống, từ đó đảm bảo độ chính xác và chuẩn hóa

- Tính tự động hóa: Phần mềm thư viện có khả năng tự động hóa các quy trình hoạt động trong thư viện, giúp giảm thiểu các hoạt động trùng lặp, giảm thiểu công sức, thời gian cho nhân viên
- Tìm kiếm thông tin: Dữ liệu nhập vào theo quy trình nhất định và có sự kiểm soát chặt chẽ của hệ thống. Từ đó đảm bảo hoạt động tìm kiếm được diễn ra nhanh chóng, đưa ra kết quả chính xác, phù hợp yêu cầu
- Lưu trữ thông tin: Tài liệu được tổ chức, sắp xếp và lưu trữ theo khoa học, có sự kiểm soát...
- Có thể quản lý tất cả các thông tin trong hoạt động giao dịch (vốn tài liệu, giao dịch mượn trả, bạn đọc, ngân sách...)
- Tính tương tác: Khả năng tương tác giữa bạn đọc và hệ thống được diễn ra một cách nhanh chóng, linh hoạt, chính xác, là cầu nối giữa thư viện và người dùng.

Phần mềm ứng dụng trong thư viện hiện nay được chia thành 3 loại cơ bản:

- Phần mềm quản trị thư viện số tích hợp (Hay phân hệ tự động hóa trong bộ phần mềm thư viện số)
- Phần mềm tạo lập, quản lý các bộ sưu tập số và các nguồn tài nguyên số (Hay phân hệ số trong bộ phần mềm thư viện số).
- Hệ thống phát hiện và tìm kiếm dữ liệu tập trung (Công tìm kiếm tập trung)

Việc kết hợp phân hệ tự động hóa và phân hệ số sẽ giúp các thư viện hình thành bộ phần mềm thư viện số. Đánh giá và lựa chọn bộ phần mềm thư viện số luôn là vấn đề phức tạp, cần kết hợp chặt chẽ nguyên tắc tiếp cận hệ thống với nguyên tắc lịch sử phát triển. Dưới đây đề cập một số tiêu chí cơ bản để đánh giá, lựa chọn bộ phần mềm thư viện số

2.2.2.1. Yêu cầu về công nghệ nền tảng

- Hệ quản trị CSDL: Phần mềm phải hoạt động được trên đa hệ quản trị CSDL: Oracle, MS SQL Server, PostgreSQL ...
- CSDL lớn: Quản lý được cơ sở dữ liệu lớn (hơn 1 triệu bản ghi) với tốc độ tra cứu nhanh.
- Hệ điều hành: Chạy trên đa hệ điều hành: Windows, Unix ...
- Hỗ trợ đa ngữ và đa mã tiếng Việt: Quản lý dữ liệu đa ngữ bằng mã UNICODE và cung cấp giao diện làm việc theo nhiều bảng mã tiếng Việt (Unicode, ABC, VNI,..), và ngôn ngữ khác nhau (Anh, Pháp, Nga).
- Giao diện trên Web: Toàn bộ giao diện của tất cả các phân hệ phần mềm được xây dựng trên Web, sẵn sàng cho kết nối diện rộng với Internet, tuân thủ chặt chẽ các chuẩn của giao thức TCP/IP.
- Mã hoá: Tuân thủ chặt chẽ các yêu cầu về bảo mật, với khả năng hoạt động trên giao thức mã hoá đường truyền SSL (Secure Socket Layer), sử dụng các thuật toán mã hoá mạnh.
- Xác thực: Sử dụng khả năng xác thực máy chủ và máy trạm (client & server certificate), cho phép quản lý các hoạt động của người dùng trong hệ thống từ nhiều mức: trạm làm việc, người sử dụng, tính năng được sử dụng.
- Tra cứu toàn văn: Tích hợp với mọi dạng dữ liệu số hóa với khả năng đánh chỉ mục
 - và tìm kiếm toàn văn tiếng Việt mạnh.
 - Hỗ trợ Việt ngữ triệt để: Hỗ trợ tiếng Việt đầy đủ với khả năng sắp xếp, tìm kiếm phân biệt hoặc không phân biệt hoa thường.
 - Hỗ trợ các công nghệ: Mã vạch, từ, RFID trong công tác quản lý và lưu thông tài liệu

2.2.2.2. Yêu cầu về chuẩn thư viện

- Chuẩn MARC: Hỗ trợ khung biên mục MARC21 và MARC21

Việt Nam.

- Chuẩn ISO 2709: Hỗ trợ khuôn dạng trao đổi dữ liệu theo ISO 2709.
- Chuẩn Z39.50: Hỗ trợ chuẩn tra cứu liên thư viện theo giao thức Z39.50 (cả client và server).
- Chuẩn biên mục: Hỗ trợ các chuẩn biên mục ISBD, AACR-2, TCVN 4743/89
- Khung phân loại: Hỗ trợ các khung phân loại: BBK, UDC, DDC, LC, khung đề mục quốc gia, subject headings.
- Chuẩn OCLC: Hỗ trợ chuẩn OCLC 4 figure cutter table, OCLC sanborn 4 figure cutter table, chuẩn cutter của TVQG cho nhan đề/tác giả tiếng Việt
- Chuẩn ISO 10161: cho nghiệp vụ mượn liên thư viện
- Chuẩn METS: cho biên mục tài liệu số

2.2.2.3. Yêu cầu về các chức năng của phần mềm

Phần mềm phải bao gồm các chức năng tối thiểu sau:

Phân hệ tự động hóa

a) Module Bổ sung

- Cho phép theo dõi, kiểm tra và giám sát các nguồn tài liệu được bổ sung đồng thời quản lý việc cập nhật tài liệu lên cơ sở dữ liệu của thư viện một cách nhanh chóng và hiệu quả:

- Quản lý bổ sung theo thể loại tài liệu: tài liệu dạng bản in (sách, báo tạp chí...), cung cấp khả năng quản lý và kiểm tra cho từng đơn đặt hàng, nhưng đồng thời cung cấp thông tin quản lý tất cả các nguồn quỹ. Quan trọng là việc tạo và duy trì sự liên kết giữa biểu ghi thư mục với tất cả các module, như vậy sự thay đổi tình trạng của một đơn đặt hàng có thể tự động tác động đến chức năng biên mục và tìm kiếm, điều đó có thể chỉ ra nhan đề đó đã được đặt hàng hay chưa.

b) Module Biên mục

Cần phải đáp ứng các yêu cầu cụ thể như sau:

Chuẩn hoá:

- Phần mềm phải cung cấp một danh sách các trường chuẩn theo MARC 21/MARC 21 VN được mô tả song ngữ Anh/Việt.
- Những trường này có thể được tùy ý lựa chọn đưa vào các mẫu biên mục.
- Mọi trường trong chuẩn MARC đều được hỗ trợ một cách đầy đủ với indicators và các trường con.

Biên mục tiện lợi:

- Hỗ trợ mọi khâu biên mục: như tạo mới, chỉnh sửa, xóa, duyệt xem, tái sử dụng, gắn kết, xuất nhập các bản ghi biên mục.
- Trợ giúp biên mục theo từng trường con: Cho phép biên mục một trường theo từng trường con (subfield codes) được chuẩn MARC gán cho trường đó.
- Cho phép dùng lại biểu ghi: dùng lại giá trị của một bản ghi có sẵn để biên mục một biểu ghi mới.
- Số định danh cho xếp giá kho mở: số cutter và năm xuất bản cho ấn phẩm trong kho mở.
- Tự động sinh giá trị số cutter theo thông tin đưa vào: Chương trình có khả năng sinh số cutter.

Tùy biến khung biên mục:

- Bổ sung các trường phi MARC: có thể gán thêm các trường mới, trường con mới.
- Tạo và chỉnh sửa các mẫu biên mục: Có thể chỉnh sửa lại các mẫu biên mục có sẵn hoặc tạo mẫu mới một cách nhanh chóng và dễ dàng.
- Phân quyền: Quyền tạo, thay đổi và huỷ một biểu ghi biên mục được gán cho từng người dùng cụ thể.

Khả năng trao đổi dữ liệu với các Phần mềm/phần mềm thư viện khác:

- Module cho phép xuất/nhập dữ liệu biên mục theo khung MARC 21 ở một trong các dạng ISO 2709 hoặc tagged; dạng XML.

- Xuất khẩu có chọn lọc: có thể thiết đặt một số tiêu chí lọc dữ liệu để xuất khẩu một nhóm bản ghi cụ thể nào đó.

- Nhập khẩu trực tuyến: Cho phép kết nối đến một hoặc nhiều máy chủ Z39.50 trên mạng Internet để nhập khẩu trực tiếp biểu ghi vào cơ sở dữ liệu.

- Nhập khẩu tích hợp: Khi nhập khẩu biểu ghi bên ngoài vào một biểu ghi đã tồn tại, kết quả phải là tổ hợp chung của 2 biểu ghi.

c) Module tra cứu trực tuyến OPAC (Online Public Access Catalog)

Hoàn toàn dựa trên công nghệ Web, OPAC cho phép người sử dụng có thể truy cập thông tin của thư viện thông qua các trình duyệt Web ở mọi nơi mọi lúc. OPAC cung cấp những phương thức tìm kiếm đa dạng, tốc độ nhanh, thoả mãn mọi yêu cầu tìm tin của người sử dụng. Opac cung cấp các tính năng cơ bản như:

- Cho phép tra cứu mọi thông tin tại chỗ cũng như truy nhập từ xa thông qua Internet. Cho phép khai thác thông tin từ các CSDL trực tuyến trên mạng qua Z39.50. Cung cấp khả năng tra cứu liên thư viện mạnh mẽ, có khả năng tìm kiếm đồng thời trên hàng trăm thư viện.

- Cho phép tìm tin theo nhiều dấu hiệu tìm kiếm khác nhau bằng các công cụ tìm tin đáp ứng các chuẩn quốc tế về tìm tin như sử dụng toán tử logic, toán tử lân cận, toán tử chặt cụt, toán tử so sánh, các dấu ngoặc, cùng với khả năng viết các biểu thức tìm tin phức hợp thoả mãn những yêu cầu tìm tin đa dạng của người sử dụng. Tìm tin không phân biệt chữ Việt hoa, chữ Việt thường.

- Cho phép đặt lọc phạm vi tìm kiếm dữ liệu thư mục theo:

- + Cơ sở dữ liệu

- + Loại tài liệu

- + Ngôn ngữ

- + Nơi xuất bản
 - + Năm xuất bản (lớn hơn, nhỏ hơn và trong khoảng)
 - Phạm vi được đặt lọc sẽ có tác dụng trong tất cả các phương thức tìm kiếm.
 - + Cung cấp 4 phương thức tìm kiếm khác nhau phục vụ tốt tất cả những yêu cầu tìm kiếm dữ liệu từ đơn giản đến phức tạp
 - + Cung cấp giao diện với khả năng lựa chọn ngôn ngữ hiển thị như Việt, Anh. Có thể dễ dàng tạo thêm giao diện ngôn ngữ mới.
 - + Cho phép tùy biến nội dung hiển thị dữ liệu thư mục: ISBD, Thẻ, Marc21....
 - + Có khả năng in và sắp xếp các kết quả tìm kiếm theo nhiều dấu hiệu khác nhau.
 - + Cung cấp thông tin về các đầu mục (tài liệu) đang được quản lý lưu thông và tình trạng của chúng.
 - + Hỗ trợ các dịch vụ với khách hàng như: xem thông tin, yêu cầu gia hạn, yêu cầu tài liệu...
 - + Hỗ trợ các ký tự thay thế như * thay cho một chuỗi, ? thay cho một ký tự
 - + Sử dụng nháy kép để yêu cầu tìm kiếm cả cụm từ chính xác.
 - + Hỗ trợ việc xây dựng các biểu thức tìm kiếm với số lượng điều kiện kết hợp không giới hạn, cho phép sử dụng ký tự thay thế và các dấu ngoặc, thỏa mãn mọi yêu cầu tìm kiếm thông tin của độc giả.
 - + Tốc độ tìm kiếm nhanh, chính xác thỏa mãn mọi yêu cầu tìm tin chuyên nghiệp với CSDL lớn hàng trăm nghìn đến hàng triệu biểu ghi.
 - + Cung cấp đầy đủ thông tin và các chức năng tương tác giữa độc giả và thư viện: thông tin độc giả, tình trạng tài liệu, đặt mượn...
- d) Module Quản lý lưu thông
- Trợ giúp thực hiện việc quản lý các bộ sưu tập trong thư viện một cách tiện lợi, công bằng và hợp lý. Bao gồm việc lưu thông của các sưu

tập thông thường cũng như các sưu tập nghe nhìn đa phương tiện và các dạng mô hình trực quan, tài liệu mượn từ thư viện khác và quản lý bàn tài liệu đặt trước.

- Môi trường làm việc Winform.
- Tự động hoá việc lưu thông tài liệu với các chức năng: cho mượn, nhận trả, gia hạn, phạt và thu phí mượn tài liệu, quản lý thực hiện các chính sách cho mượn.
- Dễ dàng sử dụng với giao diện Winform, dễ thực hiện các chức năng, có khả năng mở nhiều cửa sổ chức năng một lúc.
- Thiết lập và thông báo thời gian mở cửa theo định kỳ hoặc đột xuất
- Thực hiện chức năng bảo trì hồ sơ độc giả và đầu mục, cho phép cập nhật thời gian thực tình trạng của độc giả và tài liệu.
- Thiết lập các chính sách cho mượn theo ma trận các điều kiện từ đơn giản đến chi tiết theo đặc thù của thư viện thỏa mãn mọi yêu cầu quản lý.
- Quản lý tài khoản độc giả, cho phép ghi nhận tự động các khoản phạt phát sinh, ghi nhận các khoản phí dịch vụ, cho phép ghi nhận các khoản trả trước hoặc đặt cọc.
- Tích hợp dễ dàng với các thiết bị in, quét mã vạch, thẻ nhựa, thiết bị công nghệ từ tính và RFID.

Quản lý hồ sơ người dùng:

- Quản lý mọi thông tin cần thiết: Ảnh, ngày sinh, trình độ, địa chỉ, nhóm ngành nghề, cơ quan; niên khóa, khoa, trường... và thông tin về thẻ (số thẻ, loại thẻ, ngày cấp, ngày hết hạn).

Khả năng xử lý thông tin người dùng:

- Có khả năng hỗ trợ các nghiệp vụ gia hạn thẻ, rút hạn thẻ, xóa thẻ, in thẻ tạo thuận tiện cho thư viện trong việc quản lý bạn đọc.

In thẻ thư viện:

- Tích hợp mã vạch: Thẻ có thể được in ra kèm theo mã vạch. Đảm bảo tương thích với mọi thiết bị phần cứng đọc mã vạch.

- In thẻ theo mẫu định sẵn (template): khả năng định dạng nhiều mẫu thẻ khác nhau.

- Có thể in kèm ảnh hoặc không kèm ảnh.

Phân loại người dùng theo nhóm:

- Chính sách thích hợp cho các đối tượng bạn đọc khác nhau (sinh viên, giảng viên, cán bộ, lãnh đạo thư viện, nhà trường): Cho phép quy định các quyền và hạn chế với từng nhóm thông qua các tham số.

- Hỗ trợ hoạt động lưu thông đa điểm: Cho phép chỉ rõ danh sách các điểm lưu thông (thư viện-kho) mà mỗi nhóm bạn đọc có quyền mượn, trả.

Các tính năng ghi mượn:

- Thay đổi hạn trả: Thủ thư có thể thay đổi hạn trả khác.

- Mượn tại chỗ và mượn về: Tư liệu mượn sẽ được tính vào những hạn ngạch khác nhau và áp dụng các hạn trả khác nhau.

- Kiểm soát được số lượng và thông tin của bạn đọc vào đọc sách trong kho mở.

- Thông tin trạng thái bạn đọc: thông tin về tình trạng mượn của bạn đọc (tư liệu đang mượn, quá hạn, tiền phạt, thẻ hết hạn...) sẽ tự động được hiển thị. Thủ thư cũng có thể truy cập nhanh vào lịch sử mượn của bạn đọc.

- Hủy yêu cầu mượn: Có thể rút lại yêu cầu mượn của tư liệu đang ghi mượn.

- Tích hợp mã vạch: Có thể nhập thông tin mượn bằng phím hoặc trực tiếp thông qua thiết bị quét mã vạch.

- Gia hạn cho bạn đọc: Một hoặc toàn bộ các tư liệu bạn đọc đang giữ có thể được gia hạn.

Các tính năng ghi trả:

- Thông báo tiền phạt: Chương trình phải có khả năng tự động tính toán và thông báo số tiền phạt mà người dùng phải trả.
- Bỏ qua tiền phạt: Có thể ra lệnh cho chương trình bỏ qua tiền phạt mà bạn đọc phải nộp.
- In biên nhận phạt: Biên nhận thu tiền phạt (nếu có) phải tự động được in ra khi phiên ghi trả kết thúc.
- Thay đổi ngày ghi trả: Thủ thư có thể thay đổi thời điểm ghi trả (giá trị ngầm định là ngày tháng hiện thời).

Xem và thông báo tư liệu giữ quá hạn:

- Gửi thư thông báo qua email: Với những bạn đọc có đăng ký địa chỉ email với thư viện, thư thông báo có thể được gửi qua email.

Truy cập dữ liệu lịch sử mượn:

- Cho phép tra cứu dữ liệu về lịch sử mượn.

Báo cáo và thống kê:

- Phần mềm cho phép báo cáo và thống kê theo các tiêu chí sau:
 - + Đầu mục quá hạn.
 - + Bạn đọc quá hạn.
 - + Danh sách bạn đọc nợ tiền.
 - + Đầu mục đang được mượn.
 - + Danh sách bạn đọc có hiệu lực.
 - + Tổng hợp lưu thông tài liệu.
 - + Báo cáo tần suất lưu thông.
 - + Tình hình phí và thu phí.
 - + Tình hình bổ sung bạn đọc.

e) Module Ấn phẩm định kỳ

Hỗ trợ tiêu chuẩn MARC cho định dạng vốn tư liệu (MARC format for Holding and Location), cung cấp giao diện đồ họa Winform cho phép dễ dàng tích hợp với các module khác trong Phân hệ tự động hóa. Các tính năng cơ bản:

- Tạo các biểu ghi thư mục cho ấn phẩm định kỳ
- Tạo các biểu ghi vốn tư liệu, biểu ghi nhận, biểu ghi chờ nhắc nhở
- Tạo các biểu ghi đầu mục từ các phát hành của ấn phẩm định kỳ
- Dự đoán các phát hành tương lai qua trường 853
- Kiểm tra nhập các biểu ghi tự sinh hoặc nhập tay
- Hiện thị thông tin lưu trữ thực trên OPAC.

Phân hệ Thư viện số

a) Module Quản lý kho tư liệu số

- Kho tư liệu số là một tập hợp các không gian lưu trữ web, ở đó các tệp tin tài liệu số được lưu trữ và có một địa chỉ web duy nhất cho mỗi tệp tin và thư mục. Phần mềm cung cấp nhiều phương thức và giao diện cho phép thư viện quản lý và cập nhật các tệp tin tài liệu một cách dễ dàng nhất.

- Cho phép tải lên tải xuống các thư mục tệp tin không giới hạn số lượng, độ sâu thư mục và các tệp tin con.

- Cho phép duyệt xem, sửa, xóa, sao chép và dán các thư mục tệp tin trong kho tư liệu số trên giao diện thân thiện tương tự duyệt thư mục tệp tin trên Windows Explorer.

- Cho phép tải lên tải xuống các tệp tin có kích thước hàng chục đến hàng trăm MB.

- Tích hợp trình biên tập HTML với việc chọn lựa các tệp tin nhúng từ kho tư liệu số trợ giúp cho thư viện tạo lập các trang web dễ dàng.

- Cung cấp khả năng cập nhật nội dung các tệp tin thông dụng một cách dễ dàng thông qua việc tự động tải về và mở bằng chương trình soạn thảo tương ứng, tự động tải lên và ghi lại những thay đổi.

- Cho tạo các ảnh đại diện theo 2 kích thước chỉ bằng một click chuột.

- Ảnh xạ địa chỉ vật lý và địa chỉ web (URL) của các tệp tin, cho phép mở xem tệp tin bằng trình duyệt.

- Cho phép thư viện quản lý không chỉ các tệp tin tài liệu số mà bao gồm mọi tệp tin chia sẻ như: ảnh đại diện, ảnh độc giả, các tệp dữ liệu và cấu hình khác.

- Tích hợp các phương thức bảo mật, sao lưu dự phòng của hệ điều hành máy chủ, đảm bảo an toàn cho các tệp tin tài liệu số, dễ dàng khôi phục khi có sự cố xảy ra.

b) Module Biên tập tài liệu số

- Tuân thủ hoàn toàn tiêu chuẩn MARC cho siêu dữ liệu mô tả, sử dụng chung dữ liệu thư mục với Phân hệ tự động hóa thư viện. Cho phép tối ưu về mặt lưu trữ và quy trình xử lý dữ liệu của nhân viên thư viện.

- Phải tuân thủ hoàn toàn tiêu chuẩn METS cho việc xây dựng và bảo trì các tài liệu số, dễ dàng đóng gói chia sẻ siêu dữ liệu theo tiêu chuẩn này với các phần mềm khác.

- Giao diện windows thân thiện, dễ sử dụng với nhiều tiện ích hỗ trợ người dùng như: kéo thả, trình đơn cảm ngữ cảnh... cho phép người biên tập dễ dàng hơn trong việc xây dựng các đối tượng số phức tạp.

- Dễ dàng chọn lựa các tệp tin nguồn từ kho tư liệu số, hỗ trợ liên kết nhiều thư mục tệp tin cùng lúc.

- Tự động tạo các ảnh đại diện cho các tệp tin nguồn có định dạng khả thi: html, jpg, png, gif, tif, bmp... Mỗi tệp tin chỉ thực hiện tạo ảnh đại diện một lần, tối ưu việc xử lý dữ liệu.

- Đáp ứng việc xây dựng mọi dạng tài liệu số phổ biến hiện nay: sách, báo tạp chí theo từng bài hoặc số hoặc năm xuất bản, các tệp ảnh, bản đồ, âm thanh, phim...

- Dễ dàng liên kết các biểu ghi thư mục từ cơ sở dữ liệu thư mục thông qua việc tìm kiếm, sao chép và dán liên kết. Một biểu ghi METS có thể liên kết tới nhiều biểu ghi thư mục đóng vai trò là siêu dữ liệu mô tả.

- Các xử lý tự động luôn được đưa vào để hỗ trợ người biên tập tạo

lập và đánh số nhanh nhất: danh sách các tệp tin nguồn; danh mục chương, bài, trang theo nhiều qui tắc số thập phân, số la mã, alphabet...

- Hỗ trợ biên tập nhúng siêu dữ liệu mô tả hoặc siêu dữ liệu nội dung tệp tin dạng html, điều đó cho phép phần mềm linh hoạt đáp ứng hoàn chỉnh tiêu chuẩn đề ra.

c) Module Tra cứu tài liệu số

- Cung cấp hai loại giao diện tra cứu: trên Windows cho nhân viên tác nghiệp và trên web cho độc giả.

- Cung cấp nhiều tính năng tìm kiếm hiệu quả: tìm lướt, tìm theo từ khóa và tìm chuyên gia.

- Cho phép tìm kiếm theo nhiều dấu hiệu tìm kiếm khác nhau đáp ứng chuẩn quốc tế về tìm kiếm như sử dụng các toán tử logic, toán tử lân cận, toán tử chặt cụt, toán tử so sánh.

- Hỗ trợ các ký tự thay thế như * thay cho một chuỗi, ? thay cho một ký tự

- Sử dụng nháy kép để yêu cầu tìm kiếm cả cụm từ chính xác.

- Hỗ trợ việc xây dựng các biểu thức tìm kiếm với số lượng điều kiện kết hợp không giới hạn, cho phép sử dụng ký tự thay thế và các dấu ngoặc, thỏa mãn mọi yêu cầu tìm kiếm thông tin của độc giả.

- Tốc độ tìm kiếm nhanh, chính xác thỏa mãn mọi yêu cầu tìm tin chuyên nghiệp với CSDL lớn hàng trăm nghìn đến hàng triệu biểu ghi.

- Cung cấp đầy đủ thông tin và các chức năng tương tác giữa độc giả và thư viện: thông tin độc giả, tình trạng tài liệu, đặt mượn...

- Cung cấp nhiều phương thức trình diễn những thông tin đặc biệt như danh mục giới thiệu khuyến đọc, các tài liệu mới cập nhật, các tài liệu được truy cập nhiều nhất, trình diễn các ảnh đại diện sinh động...

d) Module Trình diễn tài liệu số

- Cung cấp một giao diện tiện lợi, đáp ứng đầy đủ yêu cầu trình diễn tài liệu theo tiêu chuẩn METS ở các khía cạnh: siêu dữ liệu mô tả,

cấu trúc...

- Sơ đồ cấu trúc vật lý của tài liệu được trình diễn như một bộ lật trang tuần tự. Sơ đồ cấu trúc logic của tài liệu được trình diễn dạng cây mục lục hoặc ở chế độ ảnh đại diện nếu có. Các phương thức di chuyển phong phú đem lại cho độc giả sự tự nhiên, tiện lợi trong khi khai thác tài liệu.

- Diện tích trình bày sơ đồ cấu trúc tài liệu và diện tích trình bày nội dung dễ dàng được thu hẹp hay mở rộng tùy thuộc vào sự tập trung cần thiết của độc giả.

- Tối đa hóa việc trình diễn bằng các chương trình nhúng trên giao diện web, giảm thiểu việc tải về và mở đối với các tài liệu thông thường: nội dung web, pdf, ảnh, âm thanh, phim..., hạn chế vi phạm bản quyền tài liệu.

- Cung cấp các tiện ích phóng to, thu nhỏ, xoay chiều với các tệp tin ảnh thông thường.

- Cung cấp nhiều tiện ích trình bày đặc biệt với ảnh có độ phân giải cao phục vụ các mục đích bảo tàng nghệ thuật, bản đồ, chẩn đoán hình ảnh...

- Trình diễn siêu dữ liệu mô tả kết hợp trong giao diện trình bày nội dung tệp tin.

- Nhiều tiện ích cần thiết được bổ sung nếu cần đối với các bộ sưu tập đặc biệt hoặc yêu cầu trình bày đặc biệt.

e) Module Quản lý lưu thông tài liệu số

- Quản lý hồ sơ độc giả với đầy đủ các thông tin cần thiết, cho phép lập tức thiết lập cảnh báo và các tình trạng hạn chế khác có thể ngăn chặn truy cập của độc giả.

- Mỗi độc giả được cung cấp một tài khoản truy cập, người quản trị có thể thiết lập các chế độ kích hoạt hoặc khóa tài khoản truy cập của độc giả.

- Cho phép thiết lập chính sách theo bộ sưu tập + nhóm độc giả. Cho phép thư viện thiết lập cách tính phí theo thời gian truy cập hoặc lượt truy

cập của độc giả đến tài liệu.

- Cho phép độc giả tự quyết định thời lượng truy cập và mức phí sẽ phải trả theo chính sách được thiết lập của thư viện.

f) Module Tìm kiếm toàn văn (fulltext search)

- Sử dụng công nghệ tìm kiếm Google Desktop Search Engine, cho ra kết quả tìm kiếm tốt nhất với cách thức sử dụng tiện lợi nhất.

- Triển khai dễ dàng và khôi phục dễ dàng trong trường hợp cần cài đặt lại Khai thác kết quả tìm kiếm nhưng vẫn kiểm soát truy cập tới toàn văn thông qua chức năng của phân hệ Quản lý lưu thông tài liệu số.

2.2.2.4. Cấu trúc dữ liệu của phần mềm thư viện số

Để hiểu hệ thống thư viện số vận hành như thế nào trước hết cần tìm hiểu mô hình khái quát cấu trúc dữ liệu của một hệ thống phần mềm thư viện số hiện đại. Các quan niệm về dữ liệu – đối tượng quản lý của phần mềm sẽ quyết định kiến trúc của phần mềm và các vấn đề mà hệ thống đó giải quyết. Một hệ thống phần mềm thư viện số thường có phải có cấu trúc dữ liệu cơ bản như sau:

- a. Các tệp tin tài liệu: Đây là đối tượng quản lý chính của một hệ thống phần mềm thư viện số. Mỗi một tài liệu có thể là một hoặc nhiều tệp tin nằm trên không gian lưu trữ web của hệ thống máy tính của thư viện và như vậy mỗi tệp tin có địa chỉ web riêng của nó, địa chỉ này chính là đầu mối để liên kết tệp tin trong các siêu dữ liệu đối tượng số. Sử dụng web làm phương tiện xuất bản tài liệu số là cách thức hiệu quả nhất vì tính phổ biến và năng lực web đã hoàn toàn áp đảo các phương thức khác.
- b. Các biểu ghi thư mục: Tương tự như đối với các tài liệu truyền thống, mỗi tài liệu số cần có một biểu ghi thư mục mô tả làm cơ sở cho việc tìm kiếm qua OPAC và quản lý tài liệu đó. Đối với các hệ thống thư viện số giản đơn biểu ghi thư mục chứa thông tin liên kết trực tiếp tới địa chỉ tệp tin tài liệu số. Ví dụ thông tin địa chỉ tài liệu

điện tử được tham chiếu trong trường 856 với tiêu chuẩn MARC21 và DC.Identifier với Dublin Core. Với tham chiếu đơn hệ thống thư viện số không cho phép thể hiện tường minh cấu trúc các tài liệu phức tạp kết hợp nhiều tệp tin và các thông tin mô tả như một tạp chí nhiều bài, một bài giảng nhiều thể loại tài liệu kết hợp... ngoài ra còn nhiều những giới hạn về kỹ thuật khác.

- c. Các biểu ghi siêu dữ liệu đối tượng số: Lịch sử phát triển thư viện số trên thế giới đã trải qua nhiều dự án với nhiều đề xuất các tiêu chuẩn thư viện số khác nhau mà không được phổ biến rộng rãi cho đến METS(Metadata Encoding and Transmission Standard) – tiêu chuẩn mã hóa và trao đổi siêu dữ liệu. METS là tiêu chuẩn để mã hóa siêu dữ liệu mô tả, quản trị, cấu trúc, bản quyền và các dữ liệu cần thiết cho thu thập, bảo trì và cung ứng các nguồn tài nguyên số. METS là một tiêu chuẩn lớn và phức tạp về mặt kỹ thuật, không tĩnh như tiêu chuẩn MARC hay Dublin Core, việc vận dụng nó cần phải linh hoạt trong thực tế rất phụ thuộc vào nhà cung cấp giải pháp cũng như yêu cầu cụ thể của thư viện.
- d. Các biểu ghi bạn đọc: Đối tượng phục vụ của thư viện là bạn đọc, mỗi bạn đọc cần được phân loại và lưu trữ các thông tin cần thiết liên quan đến họ để thư viện có thể thiết lập lên các chính sách truy cập tới tài liệu số, quản lý truy cập và thu phí.

Hiện trạng và xu hướng áp dụng các phần mềm quản trị thư viện tích hợp tại Việt Nam

a. Phần mềm thương mại

Lịch sử phát triển và ứng dụng phần mềm quản trị thư viện tích hợp gắn liền với lịch sử phát triển của ngành thư viện tại Việt Nam. Sau khi ngành thư viện Việt Nam đưa vào áp dụng các tiêu chuẩn nghiệp vụ thư viện dựa trên tiêu chuẩn quốc tế của thư viện như khổ mẫu Marc21, qui tắc biên mục AACR2, khung phân loại DDC cách đây hơn 10 năm,

các công ty phần mềm Việt Nam đã bắt đầu xây dựng các phần mềm quản trị thư viện tích hợp dựa trên những tiêu chuẩn, nghiệp vụ thư viện trên thế giới, dựa trên qui trình, nghiệp vụ quản lý thư viện và từ tư vấn hỗ trợ của các nhà nghiên cứu, quản lý thư viện tại Việt Nam. Tiêu biểu trong số này là các phần mềm iLib của công ty CMC, Libol của công ty Tinh Vân và Vebrary của công ty Lạc Việt.

Có thể nói, các phần mềm thương mại quản trị thư viện tích hợp mang thương hiệu Việt Nam, tuy có rất nhiều hạn chế tại thời điểm đó, nhưng đã làm thay đổi bộ mặt của hệ thống thư viện Việt Nam. Từ việc không có phần mềm sử dụng hoặc sử dụng hệ thống phần mềm lạc hậu (CDS/ISIS), các thư viện đã được áp dụng một phần mềm quản lý theo phương pháp tiên tiến của thế giới. Vì vậy, các phần mềm này đã được các thư viện đón nhận một cách tích cực. Tính đến thời điểm năm 2014 phần lớn các thư viện đại học, thư viện công cộng tỉnh/thành phố tại Việt Nam sử dụng phần mềm thương mại mang thương hiệu Việt Nam. [30]

Bên cạnh phần mềm mang thương hiệu Việt, hệ thống thư viện còn đón nhận một số sản phẩm phần mềm thương mại của nước ngoài như Virtua của hãng VTLS, Aleph của hãng Ex Libris và Millennium của hãng Innovative Interfaces. Đây là các phần mềm được sản xuất bởi các công ty nước ngoài có trên 30 năm kinh nghiệm trong lĩnh vực phần mềm thư viện và cũng là các phần mềm được sử dụng rộng rãi trên thế giới nên đảm bảo 100% tiêu chuẩn nghiệp vụ thư viện cũng như các yêu cầu của thư viện Việt Nam. Tuy nhiên, do giá thành quá cao nên cho đến nay chỉ khoảng 10 thư viện lớn tại Việt Nam đang sử dụng các phần mềm nước ngoài này.

Dựa trên các ý kiến đánh giá, nhận xét của các thư viện trực tiếp sử dụng và các chuyên gia trong lĩnh vực thư viện, các phần mềm quản

trị thư viện tích hợp thương mại đang được sử dụng tại Việt Nam đang gặp một số bất cập cơ bản sau:

- *Phần mềm quản trị thư viện tích hợp thương mại nước ngoài:*
 - Chi phí mua bản quyền và bảo trì quá cao, vượt quá khả năng của đại đa số thư viện Việt Nam
 - Rất khó tùy biến
- *Phần mềm quản trị thư viện tích hợp thương mại Việt Nam*
 - Rất ít được cập nhật và không theo kịp xu thế phát triển của thư viện thế giới (một số công ty đã ngừng phát triển phần mềm)
 - Tiêu chuẩn thư viện không đồng nhất và không đáp ứng 100% chuẩn quốc tế về thư viện
 - Không có khả năng chia sẻ liên kết với nhau và với hệ thống thư viện trên thế giới
 - Rất khó để kết nối tới các ứng dụng phần mềm và thiết bị khác liên quan đến thư viện (phần mềm quản lý tài nguyên số, phần mềm tìm kiếm tập trung, thiết bị RFID vvv...)
 - Phần mềm được tùy biến tùy tiện cho mỗi thư viện
 - Các vấn đề về kỹ thuật, phát triển phụ thuộc hoàn toàn vào các nhà cung cấp
 - Đội ngũ phát triển phần mềm và hỗ trợ không ổn định (đa số đã chuyển công tác khác, đội ngũ mới trình độ thấp)
 - Dịch vụ hỗ trợ kém
 - Chi phí phần mềm chỉ phù hợp với các thư viện đại học, thư viện tỉnh, thành phố. Các hệ thống thư viện huyện, xã, cao đẳng, thư viện các trường học không có khả năng để sử dụng phần mềm.

Việc áp dụng các ứng dụng CNTT, phần mềm quản lý thư viện vào công tác quản lý, khai thác thư viện đã đem lại một số thành tựu nhất định:

- Tin học hóa, tự động hóa công tác nghiệp vụ thư viện
- Đã áp dụng các tiêu chuẩn quốc tế vào các sản phẩm thư viện
- Đã áp dụng các quy trình vận hành thư viện theo tiêu chuẩn quốc tế
- Trình độ nghiệp vụ của cán bộ quản lý, cán bộ thư viện ngày càng nâng cao
- Các sản phẩm và dịch vụ thư viện đã tiếp cận bạn đọc/người sử dụng dễ dàng hơn

Tuy nhiên cho tới thời điểm này, các giải pháp phần mềm quản lý thư viện đang được sử dụng tại Việt Nam đã bộc lộ rất nhiều hạn chế, ảnh hưởng nghiêm trọng đến sự phát triển bền vững, lâu dài, có tính hội nhập quốc tế của ngành thư viện Việt Nam.

Về phần mềm có xuất xứ nước ngoài, tuy đáp ứng hoàn toàn các tiêu chí về kỹ thuật, tính cập nhật và hội nhập quốc tế, song với chi phí lớn (Vào thời điểm năm 2017 là khoảng 3 tỷ đồng cộng với hàng trăm triệu phí bảo hành bảo trì hàng năm), thì việc áp dụng rộng rãi cho toàn bộ hệ thống thư viện Việt Nam là bất khả thi. Số thư viện có khả năng tài chính để mua và duy trì phần mềm nước ngoài là con số quá nhỏ trong tổng số hàng ngàn thư viện tại Việt Nam. Vì vậy, sự ảnh hưởng của nó đến bức tranh toàn cảnh của hệ thống thư viện Việt Nam là rất ít.

Trong khi đó các phần mềm thương mại Việt Nam lại gặp nhiều khó khăn, trở ngại trong quản lý hoạt động thư viện.

b. Phần mềm mã nguồn mở

Phần mềm mã nguồn mở là phần mềm với mã nguồn được công bố và sử dụng một giấy phép nguồn mở. Giấy phép này cho phép bất cứ ai

cũng có thể nghiên cứu, thay đổi, cải tiến và phân phối phần mềm ở dạng chưa thay đổi hoặc đã thay đổi

Các phần mềm quản lý thư viện mã nguồn mở đã và đang được sử dụng rộng rãi trên thế giới gồm:

- Phần mềm quản lý thư viện tích hợp: Emilda, MicroLCS, Evergreen, phpmymlibrary, Koha, OpenBiblio, Avanti ...
- Phần mềm quản lý tài nguyên số: Invenio, Greenstone, EPrints, Dspace ...
- Phần mềm tìm kiếm tập trung: Vufind

Các phần mềm quản lý thư viện mã nguồn mở đã và đang được sử dụng tại Việt Nam gồm:

- Phần mềm quản lý thư viện tích hợp: Koha
- Phần mềm quản lý tài nguyên số: Dspace, Greenstone
- Phần mềm tìm kiếm tập trung: Vufind

Việc áp dụng phần mềm quản lý thư viện mã nguồn mở có nhiều ưu điểm như sau

- Các phần mềm mã nguồn mở nói chung và phần mềm quản lý thư viện mã nguồn mở nói riêng được phát triển bởi cộng đồng các nhà chuyên môn, phát triển, lập trình và sử dụng trên toàn thế giới
- Sự đóng góp của cả cộng đồng thư viện gồm các chuyên gia thư viện, các nhà lập trình hay các công ty cung cấp dịch vụ thư viện đảm bảo phần mềm luôn được cập nhật các tiêu chuẩn, tính năng mới nhất trong lĩnh vực thư viện.
- Việc tham gia vào chuỗi phát triển và ứng dụng phần mềm mã nguồn mở cho thư viện đảm bảo thư viện Việt Nam:
 - o Luôn luôn được song hành với sự phát triển của thư viện thế giới.

- Phát triển bền vững và nâng cao khả năng hội nhập, trao đổi thông tin với các thư viện quốc tế
- Có điều kiện tiếp xúc, trao đổi, học hỏi kinh nghiệm chuyên môn với đồng nghiệp trên thế giới
- Tiết kiệm chi phí tối đa:
 - Vì là mã nguồn mở nên thư viện Việt Nam sẽ không phải trả chi phí bản quyền phần mềm. Do vậy, các thư viện dù nhỏ nhất cũng có điều kiện áp dụng.
 - Các công ty cung cấp dịch vụ sẽ chịu trách nhiệm hỗ trợ cộng đồng và sẽ chỉ thu phí dịch vụ khi thư viện có yêu cầu.
- Làm chủ Công nghệ: chỉ có sử dụng phần mềm mã nguồn mở chúng ta mới có thể thoát khỏi được sự ảnh hưởng, phụ thuộc vào các hãng cung cấp phần mềm độc quyền trên thế giới.

Kết luận

Qua những phân tích ở trên, có thể khẳng định ứng dụng phần mềm nguồn mở là lựa chọn tối ưu cho quản lý các thư viện quy mô vừa và nhỏ ở Việt Nam. Đặc biệt trong điều kiện kinh phí đầu tư cho thư viện còn rất hạn chế.

2.3 CÁN BỘ THƯ VIỆN SỐ

Như chúng ta nhận thấy, ngoài sách và tạp chí ra, một thư viện cần bao gồm con người, thông tin, tài chính, công nghệ và nhiều tài nguyên khác. Vì tài nguyên con người (human resource) là một động lực tích cực nhất, nó đóng vai trò là một tài nguyên có ảnh hưởng lớn trong việc kiểm soát và sử dụng các tài nguyên khác. Thiếu tài nguyên con người; thì không tài nguyên nào hữu dụng cả. Bởi vậy, một trong những nhiệm vụ có tính sứ mệnh khi dịch chuyển một thư viện có tính di sản trước kia vào một thư viện số đó là nhằm chuyển đổi thành công tài nguyên con

người của một thư viện trở thành các cán bộ thư viện số để đáp ứng đòi hỏi của quá trình dịch chuyển này.

Trên thế giới, việc ứng dụng công nghệ mới, bao gồm việc sử dụng các loại phương tiện truyền thông, đang được triển khai rất mạnh mẽ. Tài nguyên số, các sản phẩm và dịch vụ thông tin – thư viện hiện đại đã trở nên phổ biến trong mọi hệ thống thư viện ở tất cả các quốc gia. Xây dựng và phát triển thư viện số với sự hỗ trợ đắc lực của công nghệ thông tin đang là nhu cầu thực sự cấp thiết. Cán bộ thư viện phải vươn lên để nắm giữ những công nghệ này. Ngoài những hiểu biết cần thiết về ứng dụng công nghệ thông tin như khả năng sử dụng phần mềm quản trị thư viện, phần mềm tạo lập và quản trị tài nguyên số... , cán bộ thư viện cũng cần được huấn luyện về kỹ năng duy trì hoạt động của máy tính trong thư viện. Mặc dù khi sự cố máy tính xảy ra, việc xử lý hầu hết đều do nhà cung cấp máy tính hay các chuyên gia máy tính đảm nhận. Tuy nhiên, trong một số trường hợp không quá phức tạp thì người cán bộ thư viện phải biết cách tự giải quyết

2.3.1. Nội dung công việc chủ yếu

Trong kỷ nguyên số, vai trò của cán bộ thư viện thay đổi nhanh chóng. Họ sẽ phát triển để trở thành những tổ chức và chuyên gia thông tin số trong xã hội. Vì vậy, so với cán bộ thư viện truyền thống, nội dung công việc của họ rất khác biệt.

Công việc chủ yếu của cán bộ thư viện số là:

- Lựa chọn, bổ sung, bảo quản, tổ chức và quản lý các bộ sưu tập số;
- Thiết kế kết cấu kỹ thuật cho thư viện số;
- Mô tả nội dung và thuộc tính của đầu mục hoặc đối tượng (siêu dữ liệu);
- Lập kế hoạch, thực hiện và hỗ trợ các dịch vụ số như định hướng thông tin, tư vấn và chuyển giao;
- Tạo lập giao diện thân thiện người dùng trên toàn bộ hệ thống

mạng;

- Xây dựng các chính sách và tiêu chuẩn liên quan đến thư viện số;
- Thiết kế, duy trì và chuyển giao các sản phẩm thông tin với giá trị gia tăng;
- Bảo vệ quyền sở hữu trí tuệ đối với thông tin số trong môi trường mạng; và
- Đảm bảo an ninh thông tin.

Để có thể thực hiện tốt vai trò, chức năng trong giai đoạn mới, ngoài những kiến thức và kỹ năng truyền thống, cán bộ thư viện số cần được được chú trọng đào tạo, bồi dưỡng về:

- Kỹ năng sống và làm việc khoa học, làm việc nhóm và giao tiếp cộng đồng, đặc biệt giao tiếp với bạn đọc.
- Kiến thức công nghệ thông tin và truyền thông đủ để thực hiện việc cung cấp các dịch vụ ngày càng dựa nhiều vào công nghệ thông tin, máy tính và hệ thống mạng.
- Kiến thức về các tiêu chuẩn để giúp họ có thể đóng góp vào việc xây dựng thư viện đạt chuẩn, thực hiện tốt hợp tác, chia sẻ, hòa nhập trong nước và quốc tế.
- Kiến thức quản lý nói chung và quản lý tài chính, nhân sự nói riêng
- Trang bị kỹ năng học tập liên tục để không ngừng tự nâng cao trình độ chuyên môn và bắt kịp với sự tiến bộ của khoa học, công nghệ

2.3.2. Cách thức phục vụ

Cho dù thư viện có phát triển theo hướng nào, thì mục tiêu của nó là đáp ứng nhu cầu thông tin và mong muốn hiểu biết của nhân loại sẽ không bao giờ thay đổi. Trong những thư viện số, cán bộ thư viện số sẽ cung cấp cho bạn đọc những dịch vụ đa dạng, tiên tiến, năng động và linh hoạt theo cách thức đầy sáng tạo, bao gồm:

- Phân tích và xử lý nhiều loại tài nguyên thông tin khác nhau;
- Thúc đẩy và tổ chức các giá trị tiềm ẩn trong mọi thông tin;

- Cung cấp những sản phẩm và dịch vụ thông tin có giá trị gia tăng cao đúng lúc và đúng đối tượng;
- Chuyển giao thông tin đúng đến người dùng và cung cấp các dịch vụ chuyên biệt và định hướng người dùng.

Hiện nay tại Việt Nam, một số thư viện lớn như Trung tâm Thông tin – Thư viện, Đại học Quốc gia Hà Nội; Thư viện Trung tâm, Đại học Quốc gia TP Hồ Chí Minh; các Trung tâm Học liệu Thái Nguyên, Huế, Đà Lạt, Cần Thơ; Cục Thông tin – Tư liệu... đã triển khai hàng loạt phương thức phục vụ hiện đại như dịch vụ tra cứu số (digital reference services), cung cấp thông tin chọn lọc trực tuyến; truy cập, khai thác CSDL thư mục, CSDL tóm tắt, CSDL toàn văn; giải đáp thông tin qua điện thoại; chat online; truyền tệp,...

2.4. DỊCH VỤ THƯ VIỆN SỐ VÀ NGƯỜI DÙNG THƯ VIỆN SỐ

2.4.1. Dịch vụ thư viện số

Khi thư viện chuyển đổi từ mô hình thư viện truyền thống sang mô hình thư viện số thì các dịch vụ thư viện cũng có những thay đổi nhất định. Trong mô hình thư viện truyền thống thì các dịch vụ cơ bản của thư viện là: phục vụ tại chỗ (tra cứu, tìm tin, đọc tại chỗ); mượn về nhà; in ấn, photocopy tài liệu... Trong mô hình thư viện số với mục đích phục vụ 24/24, không phụ thuộc vào khoảng cách địa lý thì xuất hiện thêm các dịch vụ thư viện số như: tra cứu, tìm tin, đọc tài liệu online; in ấn, mượn liên thư viện, bao gói thông tin, gia hạn tài liệu qua mạng...

Dịch vụ tìm tin trực tuyến (OPAC) được cung cấp cho bạn đọc ngay tại thư viện và từ xa. Ngoài CSDL thư mục của thư viện mình, nhiều thư viện còn hỗ trợ bạn đọc tìm kiếm thông tin của các cơ quan thông tin - thư viện khác cả trong và ngoài nước bằng cách kết nối trực tuyến hoặc thu thập và cài đặt vào máy chủ. Với việc tạo lập cổng thông tin (portal), một vài thư viện lớn đã tạo điều kiện cho bạn đọc sử dụng

nhiều dịch vụ mới như giao tiếp trực tuyến (Chat reference), giới thiệu Bộ sưu tập số hóa...

Các dịch vụ thư viện số đã mang lại cho bạn đọc các lợi ích rất to lớn. Ngay trong khi tra cứu tìm kiếm tài liệu, bạn đọc đã biết được tài liệu họ muốn tìm hiện đang có tại các thư viện nào, kho nào... Bạn đọc có thể đặt các yêu cầu như: mượn, sao chụp, nhận bản copy điện tử ngay trong hệ thống này mà không phải đến các thư viện đó. Dịch vụ mượn liên thư viện sẽ thực hiện một cách tự động chuyển yêu cầu này đến thư viện có cuốn sách mà độc giả có yêu cầu. Thư viện có sách sẽ chuyển cuốn sách này theo đường bưu điện đến thư viện của bạn đọc. Bạn đọc sẽ làm thủ tục mượn tại thư viện của họ. Với yêu cầu sao chụp tài liệu, bạn đọc sẽ nhận được các bản copy do thư viện gửi tới nhà qua đường bưu điện. Hiện nay, một số thư viện hiện đại trên thế giới đang thực hiện một dịch vụ mới: dịch vụ cung cấp bản sao điện tử (các file dữ liệu) tới địa chỉ E-mail của bạn đọc có yêu cầu.

Một dịch vụ khác của thư viện số cung cấp cho bạn đọc là Dịch vụ thông báo thường xuyên các tài liệu mới theo yêu cầu riêng của độc giả (Current Awareness Service - SDI: Dịch vụ thông tin có chọn lọc). Dịch vụ này được thực hiện một cách tự động ngay khi các tài liệu này được cập nhật vào các CSDL hoặc thông tin sẽ được cung cấp dựa trên yêu cầu của bạn đọc.

Một trong số các dịch vụ nổi bật khác của thư viện số là chia sẻ các kết quả nghiên cứu thông qua việc tạo ra các xuất bản phẩm điện tử, làm cho việc xuất bản các kết quả nghiên cứu không bị lệ thuộc vào các nhà xuất bản.

Theo báo cáo của Dự án “2014 Library Edition, NMC Horizon Library Project”, do sự tăng nhanh về mức độ truy cập tài liệu số trên thiết bị di động, các thư viện đại học và nghiên cứu cần có giải pháp phù hợp để chuyển giao những nội dung sẵn có dưới dạng truyền thống (in

ấn) trước kia tới máy tính bảng và điện thoại thông minh, đồng thời phân phối nhiều loại hình sản phẩm và dịch vụ mới hơn như sách điện tử, tạp chí khoa học điện tử.... Thói quen đọc đang thay đổi với mức tăng nhanh việc truy cập, khai thác thông tin trên thiết bị di động. Xu hướng này đòi hỏi nội dung thông tin/tài liệu cần được định dạng theo hình thức mới, có thể giúp người dùng tin đọc được trên mọi thiết bị đầu cuối. Trong khi tệp tin PDF đã và đang trở thành thói quen chung để truy cập nội dung số kể từ những năm 1990, thì chúng cũng trở nên phù hợp hơn cho máy tính cá nhân, máy tính xách tay và các thiết bị lớn khác. Những giải pháp chuyển giao nội dung trên định dạng EPUB 3 được xem như là một lựa chọn có thể phát triển tốt cho việc đọc trên các thiết bị di động nhỏ. Bạn đọc thư viện mong muốn truy cập nhiều hơn tới tài liệu số, bao gồm sách điện tử, video, và những hình ảnh hoá dữ liệu trên thiết bị di động. Để bắt kịp xu hướng này, thư viện đang làm việc với các nhà cung cấp và nhà xuất bản để phân phối sách điện tử, chủ yếu thông qua mô hình truy cập vĩnh viễn và thuê bao hàng năm

Nhiều thư viện ở Mỹ hiện đang bắt đầu phát triển những chính sách cho việc sử dụng thiết bị di động. Nhận ra tầm quan trọng ngày càng tăng của điện thoại trong đời sống hàng ngày, Hệ thống Thư viện Đại học Penn (Penn State University Libraries) đã dành riêng một khu vực sử dụng điện thoại để tránh những cuộc điện thoại mất sự yên tĩnh trong môi trường thư viện.

Ủy ban Hệ thống thông tin (JISC - Joint Information Systems Committee) của Vương quốc Anh đề xuất một chương trình xây dựng Hạ tầng Di động cho Thư viện (JISC Mobile Infrastructure for Libraries Programme) đã thiết lập lên một dự án “Mobilising Academic Content Online” để phát triển một bộ công cụ thực hành nhằm giải quyết những thách thức liên quan đến chuyển giao nội dung tài liệu học tập trên thiết bị di động. Tương tự như vậy, một dự án được cấp quỹ gần đây tại Đại

học California Davis (University of California Davis) sẽ tạo ra một lộ trình cho các kế hoạch đầu tư có tính chiến lược vào phần mềm, tiêu chuẩn, và nghiệp vụ mới. Những giải pháp công nghệ và nhiều tài nguyên thông tin trực tuyến hiện cũng đã sẵn sàng để giúp thủ thư hiểu biết và phát triển các ứng dụng, cũng như website thư viện phù hợp với thiết bị di động hơn. Tạp chí Techsource của Hiệp hội Thư viện Mỹ (ALA) hiện cung cấp nhiều thông tin và hướng dẫn cách thức cải tiến website di động của thư viện. “23 Mobile Things” là một khoá học trực tuyến giúp tìm hiểu tiềm năng của các công cụ di động nhằm đổi mới dịch thư viện.

Hệ thống Thư viện Đại học Duke (Duke University Libraries) đang triển khai ứng dụng đọc tạp chí điện tử trên máy tính bảng nhằm làm cho tài nguyên thông tin của thư viện trở lên thân thiện với thiết bị di động hơn, đồng thời cho phép bạn đọc của thư viện tìm duyệt, đọc, và theo dõi tạp chí khoa học hiện hành dễ dàng. Thư viện Đại học Exeter (University of Exeter) cũng triển khai dịch vụ nội dung di động bằng việc cung cấp một danh mục tài nguyên thông tin có thể truy cập dễ dàng qua website di động và ứng dụng di động. Khi các site di động và ứng dụng di động phát triển phổ biến, thì các thư viện đại học và nghiên cứu cũng cần cung cấp những giải pháp phù hợp cho mọi thiết bị di động, như thiết kế hỗ trợ hay tương thích mọi màn hình di động. Các website tại thư viện Đại học Toronto (University of Toronto), Đại học Princeton (Princeton University), Đại học ETH (ETH Bibliothek), Đại học Arizona (University of Arizona) đã đảm bảo rằng việc xem thông tin có thể được tối ưu hoá cho mọi kích cỡ màn hình, thậm chí cả những màn hình vẫn chưa được sáng tạo [37].

2.4.2. Người dùng thư viện số

Khác với môi trường thư viện truyền thống, người dùng tin trong môi trường thông tin điện tử không bị giới hạn bởi thời gian và không gian, họ có thể truy cập tới nguồn tin thông qua mạng máy tính. Tuy nhiên, yêu

cầu đối với người dùng tin điện tử phải cao hơn, họ cần có các kỹ năng cần thiết về máy tính, mạng máy tính, tìm kiếm thông tin cũng như kiến thức chuyên môn.

Nếu như đối với thư viện truyền thống, việc xác định mối quan hệ giữa các yếu tố cấu thành của thư viện rất dễ dàng thì với thư viện số lại khó xác định người dùng tin. Thế giới điện tử cung cấp cho người dùng tin khối lượng thông tin khổng lồ. Tuy nhiên, chất lượng thông tin lại phụ thuộc rất nhiều vào khả năng của người sử dụng như: khả năng sử dụng máy tính, khả năng xác định nguồn tin, khả năng tìm và định vị nguồn tài nguyên thông tin điện tử. Người dùng tin và nguồn thông tin không bị giới hạn bởi thời gian và không gian, thông qua môi trường điện tử người dùng tin có thể truy cập trực tuyến đến các nguồn tin bất cứ khi nào họ cần, ở bất cứ đâu.

Người dùng tin trong môi trường thông tin điện tử rất đa dạng, khó xác định bởi tính phi không gian của nó. Với tính chất là kho thông tin điện tử thì ở bất cứ đâu cũng có thể truy cập được thông tin mình cần. Chính vì vậy, có thể nói người dùng tin điện tử phong phú hơn, đa dạng hơn, “mềm hơn” người dùng tin truyền thống.

Nếu như với thế giới người dùng tin là khó xác định thì cộng đồng người dùng lại rất dễ xác định bởi nó được hình thành tại một địa điểm cụ thể.

Người dùng tin có thể tương tác với nguồn tin để tìm được thông tin cần thiết cho nhu cầu của mình. Trong môi trường thư viện số, giao diện ở đây là số như trang web, cho phép người dùng tin sử dụng thông tin thông qua chiến lược tìm kiếm phù hợp, bên cạnh đó còn cung cấp dịch vụ tư vấn và những thông tin phản hồi từ phía người dùng tin.

Trong môi trường thư viện số cần có giao diện chuyển đổi, một mặt đảm bảo thực hiện chức năng phiên dịch yêu cầu tin của người dùng tin thông qua các toán tử, mặt khác thỏa thuận các điều khoản và trình bày thông tin cho người sử dụng. Với giao diện này, người dùng tin có thể

lựa chọn bất cứ thông tin nào cần. Đồng thời có thể đưa ra ý kiến phản hồi về thông tin đó.

Để có được giao diện với nguồn tin cần có tập hợp thông tin xác định và đó chính là tập hợp các đối tượng số có trong kho dữ liệu số của thư viện số.

Ngoài ra, để có thể tìm kiếm được thông tin trong môi trường thư viện số còn cần có các siêu dữ liệu nguồn, tức là những dữ liệu về các đối tượng số. Siêu dữ liệu có tác dụng giúp cán bộ thư viện quản lý, tổ chức kho thông tin, giúp người dùng tin trong quá trình thực hiện chiến lược tìm tin của họ.

Khác với thế giới người dùng tin, thế giới thông tin là bất kỳ thông tin nào trên thế giới, dưới bất kỳ hình thức và khổ mẫu nào.

Từ năm 2003, nhiều nhà nghiên cứu đã tham gia vào nghiên cứu việc sử dụng hiệu quả các website thư viện đại học và nghiên cứu nhằm đưa ra cách tốt nhất để tổ chức nội dung cho người dùng tin trong môi trường số. Các nhu cầu về tìm kiếm tài liệu cho học tập, nghiên cứu; thông tin chung về thư viện; xác định tài khoản người dùng... luôn là những tính năng cần thiết cho bạn đọc. Việc truy cập thông tin của các đối tượng người dùng tin hiện nay đã thay đổi do sự phổ biến nhanh chóng của công nghệ di động. Thư viện đại học và nghiên cứu đã triển khai nhiều giải pháp để đáp ứng người dùng tin trên di động, bao gồm tạo ra phiên bản website di động, ứng dụng di động, và thiết kế website tự động thích nghi với mọi loại kích cỡ màn hình di động. Những người lãnh đạo và cán bộ thư viện đang phối hợp chặt chẽ với các chuyên gia công nghệ để tìm ra cách thức giúp bạn đọc truy cập nguồn lực thông tin thư viện hiệu quả nhất trên nhiều loại hình thiết bị di động khác nhau.

Sự nổi lên của công nghệ di động cũng sẽ làm thay đổi luồng công việc học tập và nghiên cứu. Ngày nay, cán bộ giảng dạy, nghiên cứu và sinh viên đang ngày càng hứng thú hơn với việc tìm kiếm trong cơ sở dữ

liệu thư mục, đọc tóm tắt và toàn văn dưới dạng số, sau đó tìm và lưu tài liệu thông qua website trên thiết bị di động hoặc ứng dụng di động của thư viện thay vì phải tới trực tiếp thư viện để mượn tài liệu. Điều quan trọng thư viện không chỉ xây dựng các tính năng cho các thông tin cơ bản về thư viện, mà còn ưu tiên những nhiệm vụ giúp cho việc tiến hành nghiên cứu có năng suất cao thông qua thiết bị di động. Tối ưu năng lực tìm kiếm tài nguyên số cho thiết bị di động là một trong những tính năng hữu ích nhất, cùng với việc tích hợp các cơ sở dữ liệu trực tuyến sẵn có của thư viện vào cùng một ứng dụng web sẽ tạo điều kiện cho bạn đọc tìm kiếm sách, bài báo ... từ một vị trí duy nhất trên thiết bị di động của họ [37].

CÂU HỎI ÔN TẬP CHƯƠNG 2

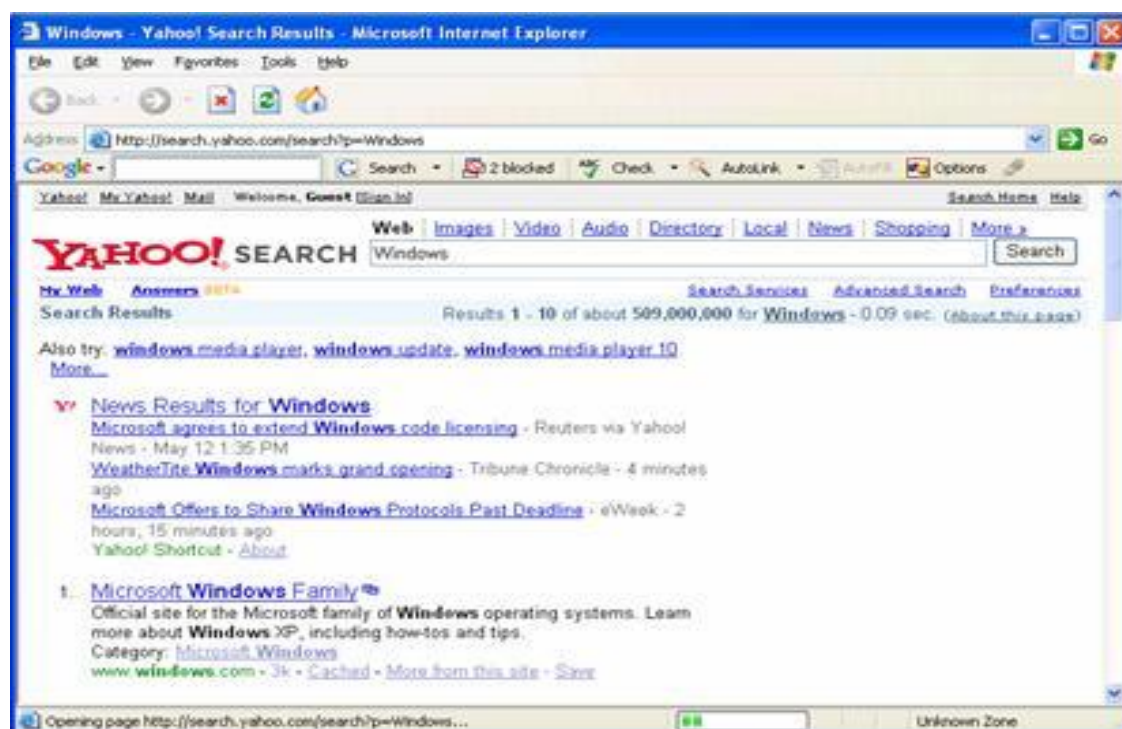
1. Thư viện số gồm những thành phần nào?
2. Trình bày các các đối tượng số và các siêu dữ liệu!
3. Hạ tầng phần cứng gồm những thành phần nào?
4. Nêu các yêu cầu của phần mềm ứng dụng

CHƯƠNG 3: SƯU TẦM VÀ KHAI THÁC THÔNG TIN SỐ

Mục tiêu đào tạo: Hướng dẫn người học sử dụng công cụ và chiến lược tìm kiếm để truy cập khai thác các tài nguyên thông tin điện tử (bao gồm các CSDL thương mại). Đồng thời trang bị kiến thức chung về số hóa nguồn tin nội sinh và hiểu rõ hệ thống phát hiện và tìm kiếm tập trung.

3.1 THÔNG TIN TRÊN INTERNET

Internet là nơi chứa đựng thông tin rất lớn của cả thế giới, về mọi lĩnh vực khoa học. Xét về mặt thương mại thì thông tin trên Internet có 2 dạng miễn phí và không miễn phí. Ở Việt Nam, người sử dụng Internet chủ yếu là sử dụng các phần thông tin miễn phí trên mạng.



Hình 3.1 Thông tin trên Internet

Nguồn thông tin quan trọng nhất trên Internet là thông tin trong các CSDL của các cơ quan thông tin, các viện nghiên cứu, các nhà xuất bản trên thế giới. Đây là nguồn thông tin lớn và quan trọng phục vụ tốt cho

công tác học tập và nghiên cứu nhưng chúng ta ít được tiếp cận vì phần lớn những CSDL này khi sử dụng phải trả tiền nên chúng ta thường sử dụng những CSDL miễn phí hay nói đúng hơn là chúng ta sử dụng những phần miễn phí của các CSDL trên mạng.

Một số đặc điểm của thông tin trên Internet:

- Nội dung trên Internet luôn được cập nhật và bổ sung, không có bất kỳ một số liệu thống kê chính xác nào về lượng thông tin có thể truy cập được trên Internet

- Tài liệu trên Internet không được xử lý bằng một hệ thống hợp chuẩn nào. Nếu danh mục tài liệu trong các thư viện được xử lý bao gồm những từ khóa chuẩn có kiểm soát thì nguồn tin trên Internet hoàn toàn không sử dụng bất cứ công cụ nào tương tự như vậy.

- Không có sự đảm bảo nào cho những thông tin mà bạn tìm được trên Internet về tính chính xác và sự cập nhật

- Một nhà cung cấp dịch vụ có thể thay đổi địa chỉ web, một cá nhân hoặc một nhóm tài trợ có thể không còn thời gian hoặc tiền bạc để duy trì một trang web, đó là một trong rất nhiều lý do khiến trang web trên Internet có thể biến mất mà không cần báo trước.

- Internet là một kho tài nguyên thông tin khổng lồ, nó có phạm vi toàn cầu, do đó để tìm một mẫu thông tin nhỏ thôi đôi khi là điều không thể thực hiện được hoặc bạn cần phải chọn lọc trong rất nhiều thông tin rác.

Nếu chúng ta không nắm vững các nguyên tắc cơ bản cũng như có một định hướng rõ ràng về hướng tìm kiếm của mình thì người sử dụng rất dễ lạc trong khối lượng thông tin đồ sộ của Internet

hoặc tìm ra những thông tin không cần thiết lắm. Công tác sưu tầm và khai thác thông tin số là một nhiệm vụ quan trọng của người cán bộ thư



viện trong hiện tại và tương lai. Để làm tốt công tác này, người cán bộ thư viện cần hiểu rõ về các công cụ tìm kiếm thông tin số, các CSDL điện tử miễn phí và có phí ...

3.1.1 Công cụ tìm kiếm (Search Engine)

Công cụ tìm kiếm hay còn gọi là bộ máy tìm kiếm (Search Engine) là một công cụ cơ bản dùng để tìm kiếm các thông tin theo những chủ đề xác định mà người sử dụng quan tâm, nó được coi là một công cụ thiết yếu và quan trọng nhất.

Các công cụ tìm kiếm làm việc theo nguyên tắc tìm kiếm trong CSDL được tự động xây dựng bởi một robot, không phải do con người xây dựng. Công cụ tìm kiếm sẽ so sánh các từ bạn đánh vào cửa sổ tìm kiếm với các từ được viết ở các trang web mà nó lưu trữ.

Lượng thông tin mà các công cụ tìm kiếm có thể bao quát thường dao động từ một số nhỏ và trong một phạm vi hẹp về nội dung cho đến số lượng lớn các trang web có thể xử lý được. Tuy nhiên không có một công cụ tìm kiếm nào có thể bao quát được toàn bộ thông tin trên Internet về một chủ đề.

Kết quả tìm kiếm có phù hợp hay không là phụ thuộc vào khả năng sử dụng các tính năng và cú pháp của công cụ tìm kiếm và diện bao quát của công cụ tìm kiếm mà bạn sử dụng.

Điểm mạnh của công cụ tìm kiếm:

- Mức độ cập nhật cao hơn danh bạ chủ đề.
- Mức độ bao quát cao hơn và kết quả tìm kiếm thường đầy đủ hơn so với danh bạ web.
- Có thể tìm thông tin rất đặc trưng
- Thân thiện với người sử dụng.

Nhược điểm của công cụ tìm kiếm:

- Kết quả tìm tin có độ chính xác không cao
- Kết quả tìm tin nhiều khi quá lớn, gây khó khăn cho việc lựa chọn

kết quả thích hợp. Để giải quyết vấn đề này các công cụ tìm kiếm thường cung cấp khả năng tìm theo các yếu tố giới hạn như subject, dạng tài liệu, ngày tháng..., hoặc tìm theo toán tử.

Công cụ tìm kiếm được sử dụng khi:

- Tìm kiếm thông tin chi tiết/đặc trưng (các chủ đề được mô tả bằng các thuật ngữ đặc trưng)

- Cần có nhiều liên kết nhưng không quan tâm lắm về chất lượng

- Tìm kiếm một cách toàn diện

Một số công cụ tìm kiếm phổ biến:

- Yahoo: www.yahoo.com

- Google: www.google.com

- Alta Vista: www.altavista.com

- Info seek: www.infoseek.com

- Excite: www.excite.com

- WebCrawler: www.webcrawler.com

- Hotbot: www.hotbot.com

- Vina seek: www.vinaseek.com

- PanVietnam: www.panvn.com

Một số kỹ năng tìm tin với Search Engines

- Sử dụng các từ khóa đặc trưng, tránh sử dụng những từ chung chung như: “tin tức”, “sự kiện” ...

- Có thể sử dụng ký tự * để đại diện cho ký tự bất kỳ nào đó. VD: run* sẽ cho kết quả là runs, running ...

Để mở rộng các chức năng tìm kiếm, cũng như tạo thêm nhiều tiện dụng cho người dùng, các công cụ tìm kiếm cũng đã hỗ trợ thêm nhiều phép toán lên từ khóa. Dĩ nhiên mỗi công cụ tìm kiếm có thể sẽ hỗ trợ những phép toán khác nhau. Ở đây chỉ nêu ra các phép toán được hỗ trợ bởi hầu hết các công cụ tìm kiếm:

- Dùng phép "+": Để tìm các trang có mặt tất cả các chữ của từ

khóa mà không theo thứ tự nào hết thì viết nối các chữ này với nhau bằng dấu +. Thí dụ: Tìm trang nói về cách thức viết Linux scripts có thể dùng bộ từ khóa: +Linux +script +tutor

- Dùng phép "-": Trong số các trang Web tìm được do quy định của từ khóa thì máy truy tìm sẽ loại bỏ các trang mà nội dung của chúng có chứa chữ (hay cụm từ) trong ngoặc kép đứng ngay sau dấu trừ. Thí dụ: Khi tìm tin tức về các loại xe dùng kỹ thuật lai mới chưa có bán trên thị trường nhưng không muốn các trang bán xe hay các trang nói về hai kiểu xe Prius (của Toyota) và kiểu xe Insight (của Honda) lọt vào danh sách truy tìm thì có thể thử từ khóa: +car +hibrid -sale -Prius -Insight

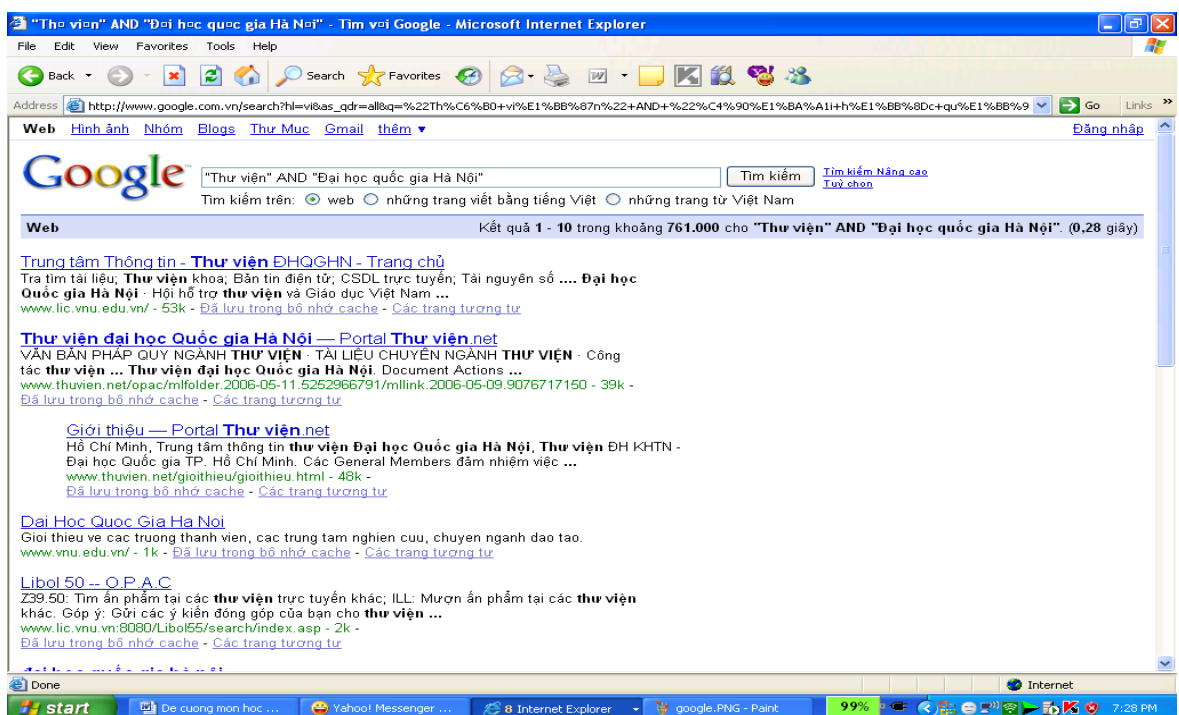
- Dùng dấu ngoặc kép " ": Khi muốn chỉ thị máy truy tìm nguyên văn của cụm từ, có thể dùng dấu ngoặc kép. Thí dụ: Để tìm lại nguyên tác và nội dung bài thơ có câu nước non nặng một lời thề thì có thể thử dùng từ khoá với ngoặc kép "Nước non nặng một lời thề"

- Dùng các phép toán của đại số Bool: Hiện tại, nhiều máy truy tìm hỗ trợ thêm các phép toán như là OR, AND và NOT. Khi dùng thì tên của các phép toán này bắt buộc phải viết chữ in hoa. Phép toán Bool đòi hỏi điền vào đúng vị trí qui định một từ (hay một cụm từ trong ngoặc kép) giữ vai trò của toán tử. Ngoài ra, đa số máy truy tìm chỉ hoạt động tốt trong một số lượng giới hạn các phép toán Bool cho một bộ từ khoá. Lời khuyên chung là không nên dùng quá 6 phép toán Bool cho cùng một bộ từ khoá và không phải máy truy tìm nào cũng hỗ trợ đầy đủ các phép toán AND, OR hay NOT

+ **OR**: Có cú pháp là (Toán tử 1) OR (Toán tử 2). Lệnh này cho phép tìm những trang Web nào có chứa một trong các toán tử của phép toán OR của bộ từ khoá. Thí dụ để tìm các bài viết về Nguyễn Trãi trong cả tiếng Việt và tiếng nước ngoài thì có thể dùng bộ từ khoá: "Nguyễn Trãi" OR "Nguyen Trai"

+ Các máy truy tìm có thể dùng OR là: AltaVista, AOL Search, Excite, Google, Inktomi (HotBot, MSN), Ask Jeeves, Lycos, Northern Light, HotBot và Gigablast.

+ **AND**: Có cú pháp (Toán tử 1) AND (Toán tử 2). Phép toán AND nhằm yêu cầu máy truy tìm kiếm các trang có sự hiện diện của tất cả các toán tử. Thí dụ “Thư viện” AND “Đại học quốc gia Hà Nội” sẽ giúp truy tìm các trang có mặt đồng thời hai chữ Thư viện và chữ Đại học quốc gia Hà Nội



Hình 3.2 Tìm tin với cú pháp “AND”

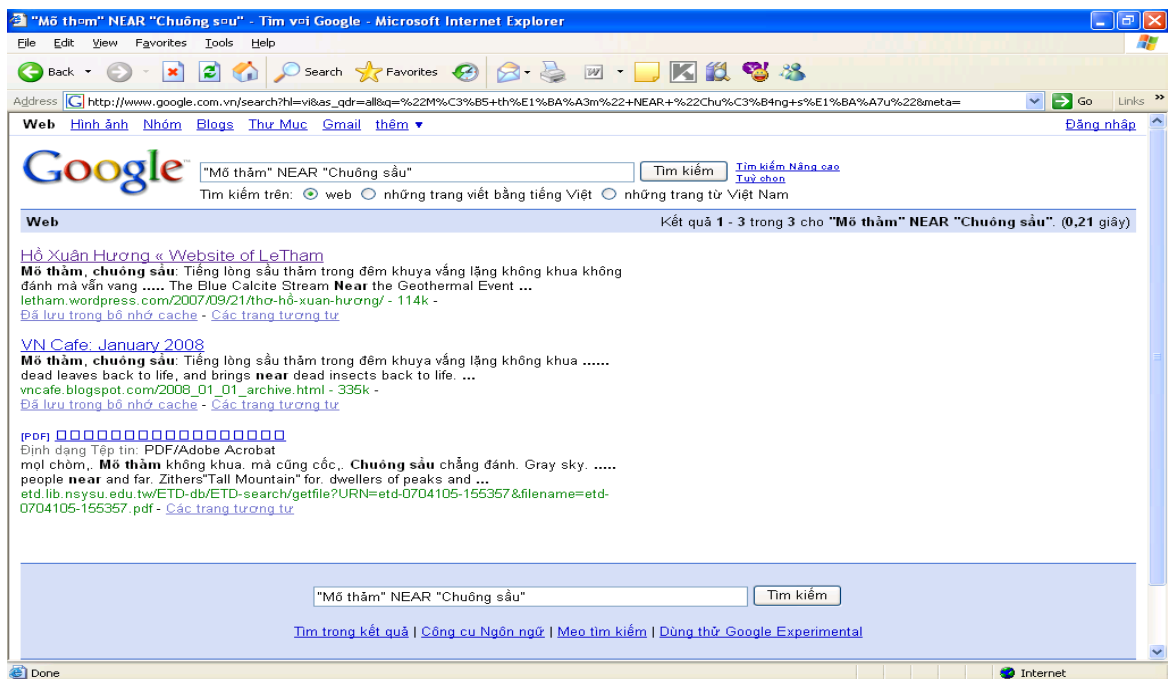
+ Một số trang truy tìm sẽ dùng AND như là mặc định (trong đó có Google). Bạn cũng có thể thay thế bằng cách dùng dấu + trong một số trường hợp nào máy truy tìm không có chức năng của đại số Bool.

+ Các trang hỗ trợ phép toán AND là: Google, AltaVista, AOL Search, Excite, Northern Light, Yahoo và Gigablast...

+ **NOT**: Phép này hoàn toàn tương tự như cách dùng dấu -. Nghĩa là, sự truy tìm sẽ loại bỏ những trang mà nội dung có chứa toán tử đứng ngay sau phép toán NOT. Tuy nhiên trong nhiều máy truy tìm có hỗ trợ

thì phép này cũng chỉ được dùng có một lần cho một bộ từ khoá. Thí dụ để tìm tài liệu hướng dẫn về ngôn ngữ lập trình Java có thể dùng trên Altavista "Java tutor" NOT book

- + Các trang có thể dùng NOT là Google, AOL Search, Excite, Inktomi (HotBot, MSN), Northern Light và Gigablast.
- Dùng phép **NEAR**: Dùng để truy tìm những trang Web mà nội dung của chúng có các thành tố của từ khoá nằm gần nhau.



Hình 3.3 Tìm tin với cú pháp "NEAR"

Phép toán này rất có lợi để tìm ra những trang có một cụm từ, một khái niệm, một định nghĩa hay một lời phát biểu mà bạn không nhớ hết được. Ví dụ: Tìm lại nguyên văn câu thơ và tác giả bằng bộ từ khoá: "Mồ thắm" NEAR "Chuông sầu"

- Dùng dấu ngoặc đơn () để chẻ nhánh. Dùng ngoặc đơn cho phép tìm nhiều kết hợp phức tạp. Thí dụ để truy tìm tài liệu về cách tạo ra CD ROM có khả năng tự khởi động có thể thử từ khoá: "tự khởi động" AND (CD OR CDROM OR CD-ROM) AND ("hướng dẫn")
- Hỗ trợ cho kiểu phân nhánh bằng ngoặc đơn là Google, AltaVista, AOL Search, Excite, Inktomi (MSN), Northern Light

Lưu ý:

- Trong mọi trường hợp thì từ khoá sai chính tả sẽ không thể có hiệu quả.
- Cách tốt nhất là dùng chữ in hoa cho các phép toán.
- Các trang tìm kiếm của Hotbot hay MSN thì bạn phải chuyển sang chọn chức năng "Boolean phrase" khi dùng các phép toán Boolean.

3.1.2. Tài nguyên điện tử

Đây là tài nguyên phong phú nhất bao gồm những thông tin về chính quyền, kinh tế, thương mại, giáo dục, học thuật, giải trí,...do các cơ sở có thẩm quyền xuất bản trên web như chính phủ, trường đại học, viện nghiên cứu, hội đoàn, công ty,...và luôn cả cá nhân độc lập

Hình thức bao gồm: Bản tin, tạp chí, sách điện tử và những hình thức đa phương tiện khác

Thường thì truy cập tự do, miễn phí; vấn đề là chúng ta phải biết chọn lọc và đánh giá nguồn tin

Một số kho tài nguyên điện tử miễn phí trên mạng:

- Directory of Open Access Journals (<http://www.doaj.org/>): Gồm 132.990 bài trích toàn văn từ 2669 tạp chí về toàn bộ các lĩnh vực khoa học: Khoa học xã hội; Nghệ thuật; Nông nghiệp; Công nghệ sinh học; Hóa học; Khoa học về môi trường; Khảo cổ học; Lịch sử; Triết học; Tôn giáo; Khoa học kỹ thuật; Chính trị và luật pháp; Ngôn ngữ và văn học; Sức khỏe
- FFT (free full text): <http://www.freefulltext.com/>
- CSDL Tạp chí khoa học Việt Nam trực tuyến (Việt journals online - VJOL) (<http://www.vjol.info.>): VJOL là một cơ sở dữ liệu các tạp chí khoa học Việt Nam trên tất cả các lĩnh vực khoa học, có 16 tạp chí với 34 mục lục, liệt kê 418 bài viết, trong đó có 294 bài viết được cung cấp toàn văn dưới dạng PDF. Mục tiêu của VJOL là quảng bá

các tạp chí khoa học tham gia VJOL cũng như các công trình nghiên cứu mà các tạp chí chuyển tải tới đông đảo bạn đọc.



Hình 3.4 CSDL Tạp chí khoa học Việt Nam trực tuyến (VJOL)

3.1.3. Chiến lược tìm kiếm thông tin

Chiến lược tìm kiếm thông tin gồm 7 bước

Bước 1: Phân tích yêu cầu tìm tin

- Tự đặt câu hỏi để làm rõ yêu cầu của mình. Ví dụ:
- + Tôi muốn biết thông tin về thư viện số
- + Tôi cần các thông tin về phần mềm
- Biến yêu cầu của mình thành 1 câu hoàn chỉnh. Ví dụ:
- + Phần mềm thư viện số
- Phân chia yêu cầu thành những khái niệm nhỏ
- + Khái niệm 1: “Phần mềm”
- + Khái niệm 2: “thư viện số”

Bước 2: Diễn đạt lệnh tìm kiếm

Cú pháp của lệnh tìm là cách thức chúng ta sử dụng để liên kết các

khái niệm một cách phù hợp cho lệnh tìm. Các công cụ tìm kiếm khác nhau trong việc liên kết các thuật ngữ tìm kiếm.

Sử dụng các phép toán, các toán tử logic. VD: “Phần mềm” AND “thư viện số”

Một số cách thu hẹp phạm vi tìm kiếm khác:

- Giới hạn theo định dạng file (.pdf, .doc, ...)
- Giới hạn theo ngôn ngữ (Tiếng Anh, tiếng Pháp ...)
- Giới hạn theo từng loại địa chỉ web (Ví dụ chỉ tìm các tài liệu từ các trang web có đuôi .edu)
- Giới hạn theo địa điểm xuất hiện của từ tìm kiếm (ở tên tài liệu hoặc trong nội dung)
- Tìm tranh ảnh hoặc bản đồ

Bước 3: Phân nhóm yêu cầu thông tin

- Phân loại yêu cầu tìm tin: tìm chính xác hay tìm tương đối .
- So sánh nhu cầu tìm tin với các tính năng của máy tìm kiếm: Ví dụ một hay một vài thuật ngữ có nhiều nghĩa trong nhiều ngữ cảnh khác nhau (“phần mềm” OR “software”) AND (“thư viện số” OR “Digital library”)

Bước 4: Chọn công cụ tìm kiếm phù hợp

- Chọn công cụ tìm kiếm phù hợp với thông tin mà bạn cần.
- Xem xét cách thức làm việc của từng công cụ tìm và diễn đạt lại lệnh tìm để có thể khai thác tối đa các chức năng của công cụ tìm đó.
- Cố gắng thực hiện việc tìm kiếm trên nhiều công cụ tìm kiếm khác nhau. Nếu các kết quả tìm từ công cụ tìm kiếm chưa thỏa mãn nhu cầu tin của bạn, hãy sử dụng nhiều công cụ khác nhau để tìm kiếm vì không một công cụ tìm kiếm nào có thể bao quát toàn bộ các trang web đang hiện hữu trên Internet.
- Xem các kết quả tìm và sử dụng các thuật ngữ được sử dụng trong các văn bản tìm được để tìm kiếm lại.

Bước 5: Tìm lời khuyên từ mọi người

- Nếu bạn không thể tìm được các thông tin bằng các công cụ tìm kiếm, hãy tìm sự giúp đỡ từ những người có kinh nghiệm: hỏi tác giả của tài liệu, gửi câu hỏi qua email và trợ giúp trực tiếp từ Thư viện Quốc hội Mỹ ...

5 bước tìm kiếm trên đây là một cách hữu hiệu để tìm kiếm trên Internet và trong nhiều trường hợp bạn có thể tìm ra rất nhiều tài liệu. Tuy nhiên độ tin cậy của tài liệu tìm được trên mạng cần được xem xét cẩn thận, vì vậy trong quá trình tìm tin bạn cần phải thực hiện những công việc tiếp theo sau đây:

Bước 6: Nếu bước đầu chưa thành công - hãy thử lại

- Để trở thành một người tìm tin có kỹ năng, bạn sẽ phải luôn xem xét lại các bước mình đã tiến hành trong quá trình tìm kiếm và tìm những cách khác nhau, diễn đạt lại lệnh tìm kiếm, sử dụng các toán tử tìm kiếm khác, hoặc thậm chí xem xét lại nhu cầu thông tin của mình. Bạn sẽ trở nên thành thạo với việc sử dụng các công cụ tìm kiếm.

Bước 7: Đánh giá kết quả tìm

- Internet là một kho thông tin phong phú nhưng không được kiểm soát, do đó cần phải đánh giá chất lượng và độ chính xác của bất cứ thông tin nào tìm được trên Internet

- Một số tiêu chí đánh giá:

+ Nguồn tác giả (tác giả có nổi tiếng trong lĩnh vực này không? tác giả có được những tác giả khác hay những người, cơ quan đáng tin cậy đề cập đến hay không? tài liệu có đề cập đến thông tin của tác giả hay không?...)

+ Nơi phát hành (có tên bất kỳ tổ chức nào trong văn bản bạn đọc không? liệu tổ chức đó có tiếng trong lĩnh vực bạn nghiên cứu không? bạn có thể xác định được mối quan hệ giữa tác giả với nơi phát hành/máy chủ không? trang web đó là của cá nhân hay tổ chức? ...)

- + Quan điểm của tác giả
- + Trích dẫn và nội dung của tài liệu
- + Độ chính xác của thông tin (phương pháp được trình bày trong tài liệu có phù hợp với chủ đề không? dữ liệu được sử dụng có thể xác định được độ chính xác không? ...)
- + Tính thời sự của thông tin (ngày cập nhật cuối cùng ...)

3.2 CSDL TRỰC TUYẾN THƯƠNG MẠI

Bao gồm những CSDL về bài tạp chí, tài liệu hội nghị, báo cáo khoa học, luận án tiến sĩ, sách điện tử,... và cả bằng sáng chế

Hiện nay có nhiều CSDL thương mại được nhiều cơ sở lớn và tập đoàn liên quốc gia trên thế giới cung cấp với giá thành tương đối cao

Có một số tổ chức vận động nhiều thư viện trên thế giới cùng mua để giá thành được hạ xuống, chẳng hạn như PERIT

Đối với các thư viện ở Việt Nam thường sử dụng hình thức mua quyền sử dụng trong một thời gian giới hạn

Một số CSDL trực tuyến thương mại:

- Springer Ebooks: Là bộ sưu tập sách điện tử chuyên ngành với rất nhiều đầu sách đã đoạt giải Nobel với những TÍNH NĂNG VƯỢT TRỘI.

Tính năng cho thư viện:

- + Hơn 3000 đầu sách điện tử và các tài liệu tham khảo hàng năm
- + Đặt mua một lần sở hữu dài hạn theo hình thức Ownership Business Model.

- + Không giới hạn người sử dụng đồng thời.
- + Được nhóm thành 12 bộ sưu tập chủ đề thuận tiện cho tìm kiếm
- + Tích hợp hoàn toàn với mục lục thư viện
- + Sẵn có bản ghi MARC 21
- + Dữ liệu thống kê hữu dụng

Tính năng cho các nhà nghiên cứu

+ Kết nối trực tiếp với hơn 10.000 đầu tài liệu và 3000 đầu sách nghiên cứu mới bổ sung hàng năm

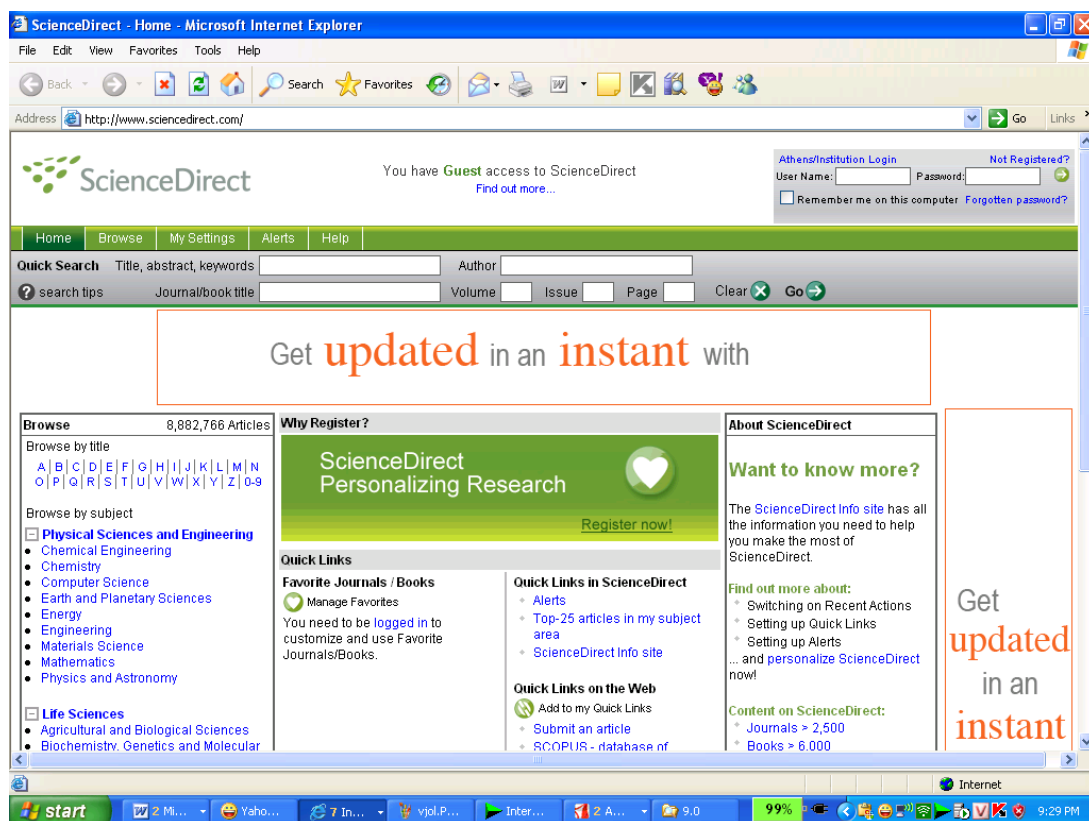
+ Dữ liệu có liên kết: kết nối liền mạch giữa sách điện tử với Tạp chí điện tử cũng như tài liệu tham chiếu trên cùng một giao diện tìm kiếm SpringerLINK

+ Các chức năng tìm kiếm và tìm lướt đơn giản và thân thiện. Có thể tìm kiếm theo từng chương, dễ dàng xác định nội dung cần thiết.

+ Các tính năng hỗ trợ định hướng tìm kiếm bao gồm cả từ điển và thesauri.

- Tạp chí điện tử toàn văn ScienceDirect Subject Collection: ScienceDirect là một dịch vụ chuyển giao hơn 1800 tạp chí điện tử toàn văn có chỉ số ảnh hưởng khoa học cao với gần 6 triệu bài báo về các lĩnh vực và chủ đề khoa học, công nghệ khác nhau, được công bố bởi các nhà khoa học, nhà nghiên cứu hàng đầu trên thế giới. ScienceDirect có các phương án lựa chọn cấp phép tối ưu, từ phương thức cơ bản là truy cập trực tiếp vào nguồn dữ

liệu tới việc hỗ trợ khách hàng cùng chia sẻ một nguồn dữ liệu (Consortium hoặc Share Programs). Hình thức cấp phép căn bản của ScienceDirect là thu phí hàng năm đối với các truy cập điện tử dựa trên việc xác định phần trăm giá trị của phí tiếp cận bản in.

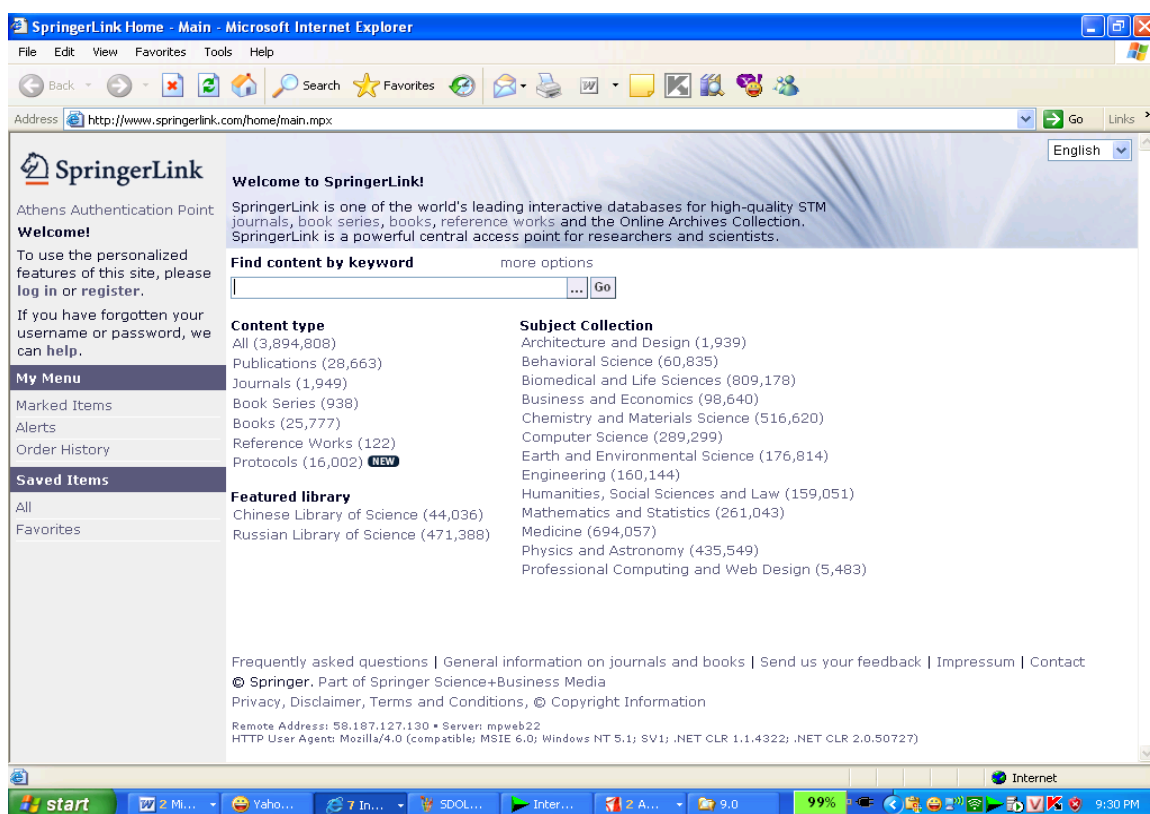


Hình 3.5 CSDL Tạp chí điện tử ScienceDirect

- Tạp chí điện tử SpringerLINK: SpringerLINK là nguồn dữ liệu điện tử hàng đầu của NXB Springer dành cho các nhà nghiên cứu trong hầu hết các lĩnh vực khoa học tự nhiên và công nghệ. SpringerLINK hiện có trên 1,200 tạp chí toàn văn có giá trị khoa học cao của các chuyên ngành:

- + Y tế và sức khỏe cộng đồng (Medicine & Public Health)
- + Khoa học cuộc sống (Life science)
- + Hóa học (Chemistry)
- + Toán (Mathematics)
- + Vật lý (Physics)
- + Kinh tế và khoa học quản lý (Economics & Management Science)
- + Khoa học máy tính (Computer science)
- + Cơ khí (Engineering)
- + Tâm lý học (Psychology)

+ Khoa học thư viện Nga (Russian Library of Science)



Hình 3.6 CSDL Tap chỉ điện tử SpringerLINK

Lưu ý: Trên đây chỉ là một số CSDL điện tử ví dụ, ngoài ra còn rất nhiều các CSDL điện tử thương mại khác như: Ebrary, EBSCO, CRC, IEEE, ACM ... Việc đặt mua các CSDL điện tử phụ thuộc vào kinh phí cũng như nhu cầu khai thác khác nhau của các thư viện

3.3 SỐ HÓA NGUỒN TIN NỘI SINH

3.3.1. Khái niệm và vai trò của số hóa

Đây là khái niệm vẫn còn mới đối với cộng đồng những người làm công tác thông tin, thư viện. Có nhiều định nghĩa khác nhau nhưng có nội dung chung đều cho rằng: *Số hóa tài liệu là quá trình chuyển các dạng dữ liệu/tài liệu truyền thống như các bản viết tay, bản in trên giấy, hình ảnh, âm thanh, phim, dữ liệu toàn văn với nhiều định dạng khác nhau sang dữ liệu trên máy tính và được máy tính nhận*



biết được như tài liệu ban đầu. Hay nói cách khác số hoá tài liệu là hình thức chuyển đổi các tài liệu truyền thống thành dạng tài liệu số mà máy tính có thể hiểu được. Sản phẩm sau khi số hóa tài liệu chính là nguồn Tài nguyên số/dữ liệu số - các dữ liệu dạng chữ, hình ảnh, âm thanh... được máy tính nhận biết đúng định dạng và được sử dụng trên máy tính.

Như vậy, thuật ngữ số hóa (Digitization) được sử dụng để chỉ quá trình chuyển đổi thông tin trong các đối tượng thực sang dạng điện tử. Trong xã hội, đối tượng thực phổ biến chứa thông tin bao gồm các dạng tài liệu, văn bản, tranh vẽ, bản đồ, băng hình, băng ghi âm ... Kết quả của việc số hóa các đối tượng nguồn tin thực sự được chuyển sang dạng điện tử. Như vậy, số hóa được coi là một phương thức tạo lập tài nguyên thông tin điện tử

Tài nguyên thông tin điện tử có thể được định nghĩa khái quát là tập hợp có tổ chức những bộ sưu tập thông tin kiến thức của các đối tượng số (digitized objects) hoặc đã được số hóa, được lưu trữ theo các công nghệ đặc biệt mà có thể truy cập, chia sẻ, khai thác theo các giao thức và thủ tục tiêu chuẩn xác định trong môi trường điện tử. Với các ưu điểm vốn có, tài nguyên thông tin số đóng vai trò rất lớn trong hoạt động thông tin, cụ thể trong việc:

- Kiểm soát tài nguyên thông tin
- Bảo vệ an toàn và lâu dài các tài liệu gốc (điều này đặc biệt có ý nghĩa khi số hóa các tài liệu có giá trị quý, hiếm như các chứng cứ của lịch sử, là di sản văn hóa ...)
- Nâng cao năng lực khai thác thông tin của người dùng tin
- Thúc đẩy mở rộng việc chia sẻ thông tin trong Hệ thống thông tin Quốc gia

Tuy nhiên trong bất cứ hệ thống thông tin nào, điều quan trọng không phải hệ thống đó có bao nhiêu thông tin mà giá trị của hệ thống thể hiện ở chỗ chúng quản trị được những loại thông tin gì và tổ chức khai thác các

thông tin đó như thế nào mới là yếu tố quan trọng.

Một khi nguồn tài nguyên thông tin số là mạch huyết, là linh hồn của thư viện số thì công tác số hóa tài liệu có vai trò vô cùng quan trọng. Trước hết việc số hóa tài liệu sẽ giúp giảm thiểu đáng kể diện tích, không gian lưu giữ; Thứ hai là giúp việc bảo quản, duy trì tuổi thọ của tài liệu truyền thống được lâu hơn; Thứ ba là dễ dàng mở rộng phạm vi cộng đồng người sử dụng nguồn tài nguyên thông tin của cơ quan thông tin, thư viện; Thứ tư là tiện ích trong việc truy xuất tìm kiếm thông tin ở bất kỳ đâu vào bất cứ thời điểm nào một cách nhanh chóng, dễ dàng; Thứ năm là thuận lợi trong việc chia sẻ nguồn tài nguyên thông tin của thư viện với các thư viện khác. Thứ sáu là giảm thiểu tối đa sức người, sức của cho việc quản lý nguồn tài nguyên thông tin truyền thống; Thứ bảy là góp phần nhanh chóng tái tạo thông tin mới có giá trị gia tăng cao hơn.

3.3.2. Chính sách, kế hoạch và quy trình số hóa

Có sáu nguyên tắc được xác định nhằm chọn tài liệu để số hóa hướng đến việc phát triển sưu tập thư viện số:

- Tính hữu dụng: Hữu dụng là lý do cơ bản trước tất cả mọi quyết định phát triển sưu tập. Tài liệu có tần suất sử dụng cao (như giáo trình, tài liệu tham khảo mà các giáo viên thường yêu cầu tất cả sinh viên tìm đọc);
- Nhu cầu nội bộ: Sưu tập nội bộ được xây dựng để phục vụ nhu cầu nội bộ và chi phí cho tài nguyên nội bộ phải được thuyết minh vì lợi ích nội bộ - chẳng hạn như đối với thư viện đại học, yêu cầu học tập, giảng dạy, và nghiên cứu là ưu tiên;
- Tài liệu mới: Mặc dù sưu tập cũ mang tính lịch sử là cần thiết cho nghiên cứu, nhưng tài liệu mới vẫn ưu tiên hơn;
- Tài liệu liên quan đến bản gốc: Những tài liệu mà người muốn tìm

hiều không thể tiếp cận được bản gốc (ví dụ các văn bản viết tay - "manuscript" của các nhà thơ, nhà văn, các nhà chính trị, hoặc các bản tuyên ngôn có chữ ký của các lãnh tụ như bản tuyên ngôn độc lập của Hoa Kỳ hiện có tại Thư viện Quốc hội Hoa Kỳ, vv...). Trên thực tế, còn có rất nhiều thể loại viết tay trên những chất liệu khác nhau. Việc số hoá các bản viết tay đó tạo điều kiện tiếp cận thuận lợi hơn cho các nhà nghiên cứu;

- Tài liệu quý hiếm: Tài liệu quý hiếm, lâu năm, độc giả không thể trực tiếp sử dụng, dễ hư hỏng - chẳng hạn như tài liệu chữ Nôm trên giấy bồi;

- Chuyển đổi nhận thức: Ngày càng có nhiều thông tin chuyển sang dạng số. Tài liệu giúp người sử dụng chuyển đổi nhận thức để làm quen việc sử dụng dạng thông tin này là ưu tiên.

Chúng ta cần phải cân nhắc mức độ ưu tiên đối với những nguyên tắc trên trong việc chọn tài liệu để số hóa.

Trong điều kiện hiện tại, việc phát triển tài nguyên thông tin số có thể nhìn nhận theo 3 mức như kịch bản sau:

- Số hóa toàn phần (fully digital resource)
- Song song tồn tại tài liệu và nguồn lực số hóa (parallel resources)
- Số hóa hồi cố (Retrospective digitization)

Hiện nay phần lớn các cơ quan Thông tin - thư viện chuyên nghiệp đều xây dựng các nguồn tài nguyên số từ các nguồn tài liệu và các ấn phẩm. Như vậy trên thực tế nguồn tin số hóa được tạo lập vẫn song song tồn tại cùng với các tài nguyên thông tin hiện hữu trên giấy

Việc số hóa các tài liệu là công việc tốn kém, đòi hỏi nhiều kinh phí, lao động và trang thiết bị và phí chuyển đổi về tổ chức. Do vậy, việc xây dựng và phát triển kho tài nguyên số không thể làm tràn lan mà phải có chọn lọc, trọng tâm, trọng điểm

Tùy theo các điều kiện cụ thể, các đơn vị thông tin phải có các nghiên cứu cần thiết như: loại tài liệu nào cần số hóa, khối lượng, quy mô,

phương thức lựa chọn ... để lập kế hoạch sao cho phù hợp

Khi xây dựng tài nguyên số, từ quan điểm lợi ích của người dùng tin và từ quan điểm pháp luật tránh rơi vào vi phạm lỗi bản quyền, cần đặc biệt chú ý tới việc xây dựng các bộ sưu tập. Một bộ sưu tập thường bao gồm nhiều tài liệu dưới nhiều dạng thức khác nhau: văn bản, hình ảnh, âm thanh. Ví dụ: một bộ sưu tập về đề tài “1000 năm Thăng Long - Đông Đô - Hà Nội” sẽ bao gồm những tài liệu dạng văn bản về lịch sử, văn hóa, phong tục...; tài liệu dạng hình ảnh về các điểm di tích, các mẫu trang phục, các lễ hội ...; tài liệu âm thanh về những điệu nhạc, bài hát, làn dân ca ... Một sưu tập thông tin số như vậy phải qua một quá trình hình hành để tạo nên những cấu trúc hỗ trợ cho việc truy tìm và có thể xuất bản, đưa ra trên các phương tiện khác nhau trên mạng Internet, trên CD-ROM, trong các CSDL...

Xây dựng tài nguyên số bằng phương thức trên có nghĩa là tổ chức lại thông tin, biến chúng trở thành nguồn lực, làm cho các thông tin đó trở lên phổ biến hơn đối với đông đảo người dùng tin mà chỉ với các ấn phẩm truyền thống rất khó, nếu như không muốn nói là không thể thực hiện. Tuy nhiên, vấn đề phổ biến thông tin trong môi trường số hóa có mặt pháp lý, mà những người tham gia cần phải am hiểu các điều khoản của luật xuất bản, sở hữu trí tuệ, quy định bảo mật, pháp lệnh lưu trữ ... để hành động đúng, để không rơi vào lỗi vi phạm pháp luật.

Sử dụng các công nghệ xử lý ảnh và nhận dạng thực hiện chuyển đổi các tập tin ảnh sang văn bản điện tử MS Word. Nếu muốn chuyển đổi văn bản trên giấy thành văn bản điện tử, chỉ cần quét hoặc chụp văn bản bằng máy quét hoặc máy ảnh số, sau đó chuyển đổi tập tin ảnh sang Word.

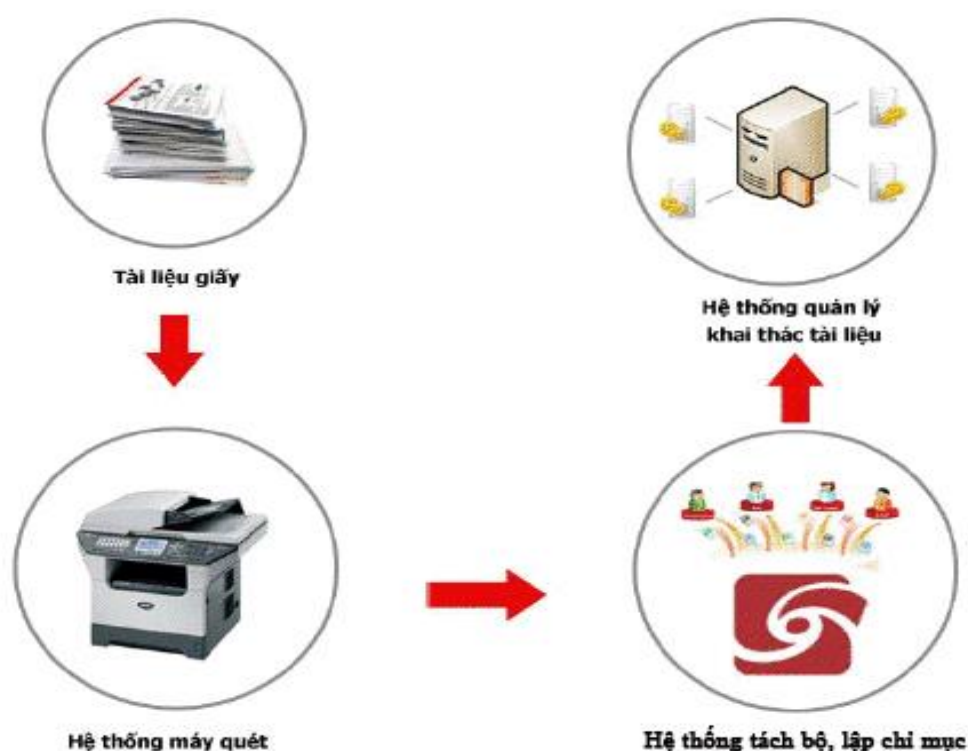
Có thể chuyển đổi mọi định dạng ảnh phổ biến như PDF, JPEG, TIFF... sang văn bản điện tử song vẫn giữ lại chính xác cấu trúc logic của tài liệu. Sử dụng công nghệ nhận dạng thông minh, các dàn trang phức tạp trong tài liệu như đầu trang, chân trang, mục lục, chú thích,

bảng biểu, đồ thị, ... được phát hiện và tái cấu trúc Tiếp đó, có thể tạo kho tài liệu điện tử phục vụ cho việc lưu trữ và khai thác.

Tổng quan hệ thống số hóa dữ liệu

Hệ thống số hóa tốt cần có khả năng chuyển đổi nhanh chóng một khối lượng lớn tài liệu giấy sang các tài liệu số trong một thời gian ngắn mà vẫn đảm bảo giữ nguyên cấu trúc của tài liệu.

Mô hình tổng quan của hệ thống số hóa tài liệu được mô tả ở hình vẽ dưới đây:



MÔ HÌNH TỔNG QUAN HỆ THỐNG SỐ HÓA TÀI LIỆU

Hình 3.7 Mô hình tổng quan hệ thống số hóa tài liệu

Quy trình hoạt động của hệ thống số hóa tài liệu được mô tả như sau:

Ở bước đầu tiên, các tài liệu giấy được đưa vào quét tại các máy quét chuyên dụng. Hệ thống máy quét này cho kết quả quét là các ảnh số, các ảnh này cần có độ phân giải đủ tốt để hỗ trợ tiến trình nhận dạng sau đó. Phần mềm nhận dạng chữ ký tự (OCR) có nhiệm vụ nhận dạng và kết

xuất kết quả nhận dạng thành các tệp văn bản có thể soạn thảo được, ta gọi chúng là tài liệu số. Tùy vào mục đích sử dụng kết quả nhận dạng, các tài liệu số cần được kết xuất ở định dạng thích hợp. Một số định dạng phổ biến có thể kể ra ở đây bao gồm: text, MS Word, MS Excel, MS PowerPoint, hay PDF. Bước cuối cùng của dây chuyền số hóa là đưa các tài liệu số này vào hệ thống thông tin để phục vụ nhu cầu khai thác của người dùng cuối. Các hệ thống thông tin đó có thể là hệ quản lý văn bản, thư viện điện tử, hệ thống tra cứu điện tử.

Là khâu không thể thiếu được của dây chuyền số hóa, phần mềm nhận dạng có nhiệm vụ chuyển ảnh quét tài liệu gốc sang tài liệu số. Mục đích cơ bản của việc số hóa là để khai thác tài liệu số bằng các cách thức như biên tập lại, trích dẫn và tìm kiếm. Bởi vậy phần mềm số hóa công nghiệp phải đáp ứng một số tiêu chí tối thiểu sau:

- Nhận dạng tiếng Việt với độ chính xác 97% trở lên
- Có cơ chế hoạt động tự động và liên tục 24/7
- Có cơ chế kiểm soát lỗi và giao diện soát lỗi thân thiện
- Dàn trang của tài liệu số sau nhận dạng giống như tài liệu giấy
- Hỗ trợ kết xuất tài liệu số ra các định dạng phổ biến như

Microsoft Office files,

- PDF, PDF/A.

Trên thị trường Việt Nam hiện nay có các phần mềm nhận dạng tài liệu tiếng Việt sau: VnDOCR của Viện công nghệ thông tin, phần mềm mã nguồn mở Tesseract và giải pháp số hóa của ABBYY.

3.3.3. Thiết bị số hóa

Một trong số các thiết bị số hóa thông dụng là máy quét. Các máy quét rất đa dạng về giá cả, hình dạng và kích thước. Chúng có giá từ 100USD cho các máy quét hình phẳng cho đến 50.000USD cho các máy quét công nghiệp cỡ lớn của các nhà sản xuất như Bell & Howell. Rất

nhiều website cung cấp đa dạng máy quét. Để tìm những website này, bạn chỉ cần dùng từ khóa “scanners” vào Google, Altavista hoặc Yahoo.

Kết quả của một trang tài liệu được quét là một tập tin máy tính mà thông thường ở định dạng TIFF hoặc Bitmap. Định dạng nén TIFF phiên bản 4 là dạng tốt nhất. Trung bình một trang được nén và được chuyển thành định dạng này chỉ chiếm khoảng 50Kb, trong khi ở định dạng Bitmap không nén sẽ là 2Mb.

Các máy quét hình phẳng giá thấp

Các loại máy quét hình phẳng là rẻ nhất và được sử dụng nhiều nhất. Thuộc nhiều hãng khác nhau: HP, Agfa, Acer v.v..., giá cả từ 100USD đến 300USD. Chúng đều có thể quét hình trắng đen hay màu. Do chi phí thấp nên có thể trang bị cho mỗi máy tính một máy quét riêng.



Điểm bất lợi của những máy in này là cho ra những hình ảnh của trang tài liệu ở mức trung bình, tỉ lệ quét thấp, không bền trong những môi trường ẩm thấp và khá dễ hư. Chúng ta phải quét từng trang một. Mỗi trang phải được định vị cẩn thận theo lề bảng quét. Hiệu suất của những máy in này kém.

Mặc dù các nhà sản xuất khẳng định rằng mỗi trang tài liệu có thể được quét trong vòng chưa tới một phút nhưng thực tế cho thấy rằng khó có thể đạt tới mức 12 trang mỗi tiếng. Tiến trình quét thường làm ì ạch máy tính gắn kết. Do vậy những máy in này chỉ hữu dụng cho các công việc nhỏ (số lượng trang cần quét ít- từ 200 đến 400 trang một tháng một cách thường xuyên) hoặc các công việc xảy ra một lần từ 1.000 đến 2.000 trang

Máy quét cấp thấp có ngăn để giấy

Các máy quét này thường có giá từ 500USD cho đến 1.200USD. Có thể quét từ 10 đến 50 trang tài liệu một lần. Vì vậy người điều khiển không cần có mặt liên tục tại máy quét. Điều này sẽ làm gia tăng số lượng

trang đến 150- 200 trang/ngày. Những loại máy in này có tuổi thọ cao hơn, thường thì khoảng từ 30.000 đến 50.000 trang.

Điểm bất lợi của chúng là tại một thời điểm chỉ quét một mặt tài liệu - ngăn để các trang tài liệu phải được đảo lại để quét mặt sau của tài liệu. Và điều này có thể gây ra vấn đề bởi vì ngăn để giấy rất thường gặp trục trặc và đôi lúc làm kẹt giấy.

Những loại này hữu ích cho công việc quét từ 1.500 đến 3.000 trang/tháng. Các máy quét màu, để quét hình màu thì nhất thiết ta phải có máy quét màu. Nhưng nói chung, chưa đến 5% các ấn phẩm chứa màu cộng với bìa tài liệu. Vì vậy một máy quét hình phẳng giá thành thấp như kể trên là thường là đáp ứng được nhu cầu. Chúng ta nên chọn máy quét có độ phân giải lên đến 600dpi.



Các máy quét 2 mặt chuyên nghiệp

Các máy quét chuyên nghiệp là các máy tốt và đáng tin cậy, có khả năng xử lý một số lượng lớn trang tài liệu- từ 2.000 đến 10.000 trang/ngày. Chúng có hệ thống khay để giấy tự động, xử lý các nhóm gồm từ 50 đến 200 trang. Các máy quét tốt nhất và nhanh nhất thuộc dạng này có thể quét cả 2 mặt tài liệu cùng lúc.

Các máy quét này yêu cầu máy tính kết nối với nó phải mạnh và có dung lượng ổ cứng ít nhất là 10 -20Gb, giá từ 5.000 - 50.000USD. Chẳng hạn như:

- Máy quét Cannon DR-6020 giá khoảng 5.000USD, có thể quét 2 mặt tài liệu, .2000 trang/ngày và tuổi thọ từ 600.000 - 800.000 trang. Các máy quét nhãn hiệu Bell&Howell và Fujitsu, giá từ 10.000 - 500.00USD, có tuổi thọ đến hàng triệu trang.

- Các máy quét phích nhỏ có giá từ 15.000USD đối với loại bán tự động cho đến 80.000USD đối với loại tự động hoàn toàn.

Máy số hóa sách đóng tập

Hiện nay trên thị trường đã xuất hiện những máy số hóa sách chuyên dụng. Tuy nhiên giá thành rất đắt lên đến hàng trăm nghìn USD, chỉ phù hợp với những đơn vị có khối lượng tài liệu cần số hóa nhiều, hoặc những đơn vị cung cấp dịch vụ số hóa chuyên nghiệp. Một số ưu điểm của hệ thống thiết bị này là:

- Tư thế gáy sách không bị ép thẳng: cho phép việc sao chụp sách tự động không phá huỷ và giảm tối thiểu việc dẫn căng gáy sách và căng trang giấy khi lật giở
- Sức căng khi lật giở trang được tối thiểu hóa.
- Tư thế trang đặt phẳng: Hai chiếc kẹp đặt ở mép trang gàn gáy sách tự động nhẹ nhàng trợ giúp làm phẳng trang giấy ngay sau khi trang được lật giở. Những chiếc kẹp này tạo ra một lực ép nhẹ vào trang giấy mà không tạo ra sức căng trên gáy sách. Những chiếc kẹp dễ điều chỉnh theo các loại kích cỡ sách khác nhau và chỉ chạm vào mép giấy với sự tiếp xúc hướng xuống.
- Khả năng lật giở trang tự động: Kỹ thuật hút chân không hoặc kỹ thuật cánh tay robot nằm bên dưới phần điều khiển và sẽ thực hiện lặp đi, lặp lại chức năng của mình cùng một kiểu và ở cùng một mức độ. Đầu hút nhấc trang sách trên diện rộng để giở trang một cách nhẹ nhàng, chứ không nhấc giở phía góc trang. Cánh tay robot lật giở trang rất đều đặn ở cùng một lực đẩy. Việc lật giở trang bằng tay trong thời gian dài sẽ dẫn đến tình trạng là người vận hành lật giở không đều tay, có thể gây rách giấy hoặc các hư hỏng khác.
- Sao chụp tự động năng suất cao: Có khả năng sao chụp tự động



lên đến hàng nghìn trang/ giờ.

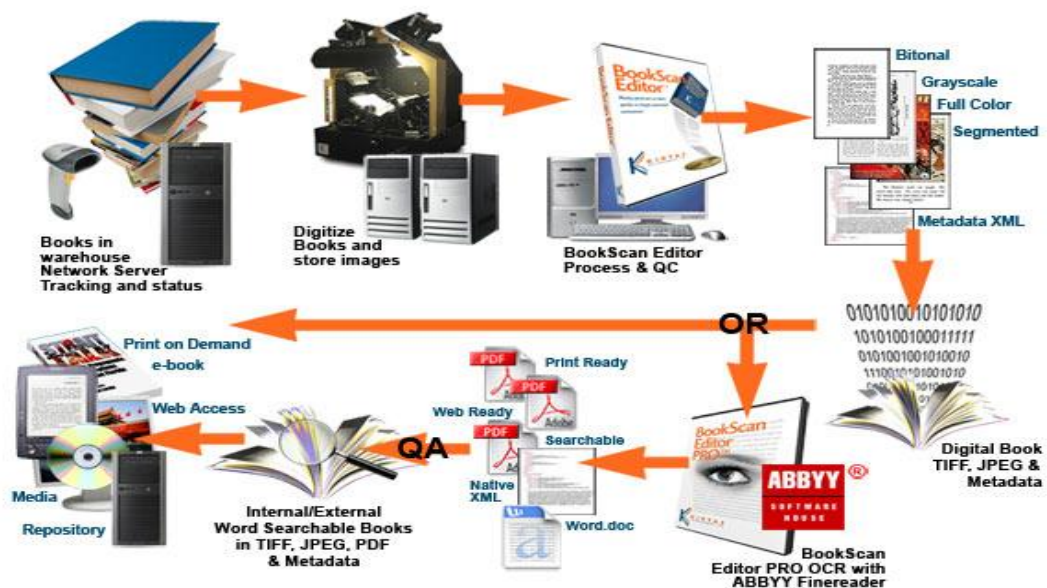
- Kỹ thuật chụp lấy trang sách: Hệ thống sử dụng máy ảnh kỹ thuật số hoặc quét quang

Mỗi máy quét đều có phần mềm riêng được cài đặt trên máy tính để điều khiển máy quét. Một số máy quét có card được cài đặt vào máy tính để tăng tốc độ quét.

3.3.4. Nhận dạng ký tự quang học: OCR

Nhận dạng kí tự hay còn gọi là hệ thống OCR làm công việc chuyển thể các hình ảnh được quét thành văn bản. Đầu vào là một hình ảnh kỹ thuật số ở định dạng TIFF hoặc Bitmap, tốt nhất là ảnh có chất lượng cao. Đầu ra là văn bản hoặc trang web, cơ bản là các định dạng RTF, PDF, Word hoặc HTML.

Sau đây là các bước cơ bản để chuyển thể tài liệu giấy tờ thành dạng kỹ thuật số:



Hình 3.8 Các bước chuyển đổi tài liệu giấy thành dạng điện tử

- Chụp, căn chỉnh trang, nâng cao chất lượng hình ảnh, làm trắng tự động theo khối dữ liệu (định dạng ảnh TIFF hoặc JPG)

- Chuyển sang định dạng PDF ảnh tĩnh và thêm bookmark cho tài liệu.

- Xử lý nhận dạng ký tự quang học (OCR): Chuyển tài liệu từ dạng ảnh sang PDF hoặc Word có thể tìm kiếm toàn văn.

- Biên mục siêu dữ liệu: Siêu dữ liệu mô tả (theo chuẩn Dublin Core), siêu dữ liệu cấu trúc, siêu dữ liệu kỹ thuật.

Tuân theo những bước này, bạn kiểm tra chất lượng các tập tin kết quả và lưu chúng ở định dạng thích hợp.

Trên thị trường có rất nhiều chương trình nhận dạng ký tự tốt với giá cả từ 100USD đến 400USD, chẳng hạn như:

- Read-Iris (<http://www.readiris.com>)
- Omnipage (<http://www.omnipage.com>)
- Fine-Reader (<http://www.finereader.com>)
- VNDOCR (<http://www.vndocr.com>)

Tất cả thông tin bao gồm cả nhà phân phối địa phương đều có thể được tìm thấy trên các website của các nhà sản xuất.

Để tự thực hiện công việc quét tài liệu cần có máy quét, phần mềm nhận dạng ký tự và sự đảm bảo về chất lượng. Do cần phải quyết định sự lựa chọn giữa tự quét hay bàn giao cho đối tác thực hiện việc quét.

3.4. HỆ THỐNG PHÁT HIỆN VÀ TÌM KIẾM TẬP TRUNG

Hệ thống phát hiện và tìm kiếm tập trung là một ứng dụng phát hiện và tìm kiếm các nguồn tin: thương mại, nội sinh và các nguồn truy cập mở khác của thư viện trên một giao diện tìm kiếm duy nhất. Tài liệu có thể được tìm trong:

- Mục lục thư viện (Cơ sở dữ liệu thư mục)
- Kho tài nguyên số nội sinh (Digital Repository)
- Các nguồn tài nguyên bên ngoài khác: cơ sở dữ liệu thư mục của các thư viện khác, các kho dữ liệu số bên ngoài, các truy cập mở (Open access: sách điện tử, bài giảng điện tử, tạp chí điện tử mở...), các cơ sở dữ liệu trực tuyến của các nhà xuất bản hàng đầu thế giới

Giải pháp phần mềm tìm kiếm tài nguyên tập trung là giải pháp mới kể cả trên bình diện thế giới: chỉ xuất hiện cách đây khoảng 5 đến 7 năm và chỉ có khoảng 8 giải pháp trên toàn thế giới.

Tại Việt Nam, không có phần mềm tìm kiếm tài nguyên tập trung được sản xuất bởi các công ty trong nước và chỉ có một vài thư viện đang sử dụng phần mềm tìm kiếm tài nguyên tập trung của nước ngoài.

Bên cạnh rào cản là chi phí triển khai và duy trì quá lớn, phần mềm tìm kiếm tập trung của nước ngoài khó áp dụng được tại Việt Nam là do nó chỉ có thể tìm kiếm tài nguyên được mô tả theo đúng tiêu chuẩn quốc tế về thư viện. Các tài nguyên tại Việt Nam đa phần không đáp ứng được tiêu chí này (do được mô tả và quản lý bởi phần mềm của Việt Nam không đáp ứng 100% tiêu chuẩn quốc tế về thư viện).

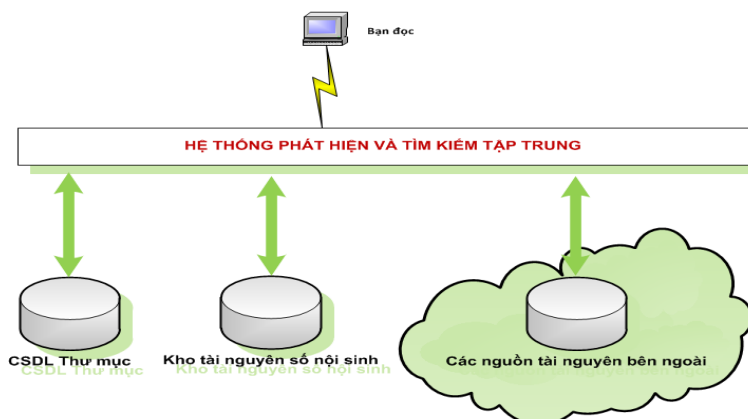
Các hệ thống phát hiện và tìm kiếm tập trung phổ biến trên thị trường hiện nay có: WorldCat Local, Primo, Summon, Encore, EBSCO Discovery Service, SirsiDynix Enterprise, VuFind (giải pháp phần mềm mã nguồn mở)...

Một số điểm chung của các hệ thống phát hiện và tìm kiếm tập trung:

- Đối tượng tìm kiếm: Tất cả các nguồn tài nguyên bên trong và bên ngoài thư viện mà hệ thống được phép kết nối tới.
- Giao diện tìm kiếm: Mặc định một ô tìm kiếm duy nhất (giống Google), có giao diện tìm kiếm nâng cao (nếu cần)
- Hiện thị kết quả tìm kiếm:
 - + Kết quả tìm kiếm được hiển thị theo mức độ phù hợp
 - + Kết quả tìm kiếm được lọc theo các nhóm, các tiêu chí khác nhau: theo định dạng, theo bộ sưu tập, theo địa điểm, theo ngôn ngữ, theo năm xuất bản...
 - + Đưa ra gợi ý liên quan đến từ khóa tìm kiếm
- Quản lý truy cập bạn đọc: hỗ trợ tính năng SSO (Single sign-on) –

đăng nhập một lần duy nhất

- Khai thác tài nguyên:
 - + Tích hợp với hệ thống quản lý thư viện tích hợp để thực hiện công việc tìm kiếm tài liệu, xem trạng thái tài liệu và các giao dịch mượn/ trả tài liệu online
 - + Kết nối và xem toàn văn các tài liệu số (nếu được phép)



Hình 3.9 Mô hình hệ thống phát hiện và tìm kiếm tập trung

Phần mềm tìm kiếm tập trung Vufind

Vufind là phần mềm tìm kiếm mã nguồn mở được phát triển bởi trường đại học Villanova năm 2010 cho phép người dùng tìm kiếm các nguồn tài liệu trên một giao diện duy nhất và vượt trội hơn giao diện cổ điển OPAC. Vufind đã và đang được sử dụng và tùy biến bởi khoảng 64 trường, tổ chức trên thế giới.

Các tính năng nổi bật của Vufind chính là các khả năng tương tự như các phần mềm thương mại như:

- Tích hợp, tìm kiếm và truy cập tới Mục lục thư viện và kho tài nguyên số nội sinh
- Phương thức tìm kiếm
- Kết quả tìm kiếm có thể được lọc và giới hạn theo nhu cầu
- Đa ngôn ngữ

Bên cạnh đó, do Vufind là phần mềm mã nguồn mở, Vufind cho phép khả năng dễ dàng tùy biến cấu hình các phân hệ chức năng, tùy biến hiển

thị hoặc tùy biến, giới hạn kết quả tìm kiếm dựa trên nhu cầu và mục đích của từng thư viện. Điều này có thể thực hiện được do Vufind được module hóa trong thiết kế, giúp cho việc bật, tắt các tính năng trở nên dễ dàng.

CÂU HỎI ÔN TẬP CHƯƠNG 3

1. Hãy giới thiệu các công cụ tìm kiếm và phân tích chiến lược tìm kiếm thông tin!
2. Trình bày các thiết bị số hóa và phương pháp nhận dạng ký tự quang học (OCR)!
3. Phân tích đặc điểm của hệ thống phát hiện và tìm kiếm tập trung!

CHƯƠNG 4: BỘ SƯU TẬP VÀ BIÊN MỤC TÀI LIỆU SỐ

Mục tiêu đào tạo: Giúp người học nắm vững những vấn đề chung về bộ sưu tập số như khái niệm, ý nghĩa, công cụ xây dựng, cách khai thác bộ sưu tập và đặc điểm một số bộ sưu tập mẫu. Giới thiệu khái quát về biên mục analog, biên mục digital và phương thức chuyển đổi biểu ghi từ MARC sang Dublin Core

4.1 CÁC KHÁI NIỆM

Sưu tập - collection: tập trung vào đối tượng số hóa, bao gồm văn bản, hình ảnh và âm thanh cùng với phương thức truy cập, truy hồi, chọn lọc, tổ chức, bảo trì sưu tập đó. Sưu tập do chuyên gia thư viện tạo nên. Một sưu tập - collection thông tin bao gồm nhiều tài liệu dưới nhiều hình thức.

Tài liệu - document: là thông điệp mang thông tin dưới hình thức điện tử. Tài liệu là đơn vị cơ sở từ đó bộ sưu tập thông tin được xây dựng, mặc dù chúng có thể có những cơ sở hạ tầng và những tập tin kết hợp riêng. Nói chung tài liệu bao gồm văn bản, hình ảnh, âm thanh hay video.

Bộ sưu tập thông tin là một sưu tập bao gồm nhiều tài liệu dưới nhiều dạng thức khác nhau: văn bản, hình ảnh, âm thanh, hình ảnh động, 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 tùy thuộc vào phương tiện và dạng thức của tài liệu đó. Ví dụ muốn đọc một văn bản PDF thì phải tích hợp trình ứng dụng Adobe Acrobat hay muốn xem một đoạn video thì phải tích hợp trình ứng dụng Windows Media Player, vv...



Hình 4.1 Ví dụ minh họa về một bộ sưu tập bằng phần mềm
Greenstone

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

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, tùy theo hình thức tài liệu được sưu tầm và tùy theo chủ đề, đề tài được quan tâm; tuy nhiên hoàn toàn giống nhau về phương cách xây dựng và hiển thị.



Hình 4.2 Ví dụ minh họa về một bộ sưu tập bằng phần mềm Dspace

Những chuyên viên thư viện số trong một thư viện với sự phối hợp và tương tác với người sử dụng sẽ tạo nên những bộ sưu tập cần thiết và hữu ích cho thư viện, phục vụ tốt công tác nghiên cứu, giảng dạy, học tập trong những cơ sở đào tạo; đồng thời đáp ứng nhu cầu khảo cứu của học giả, nhà nghiên cứu và nhiều đối tượng khác. Đây là công việc thường xuyên của một thư viện đòi hỏi chuyên viên thư viện số phải phối hợp với bộ phận Dịch vụ thông tin để nắm bắt yêu cầu thông tin của người sử dụng nhằm phục vụ tốt cho việc sưu tầm thông tin; có kỹ năng số hóa tài liệu; am hiểu chuẩn Dublin Core trong công tác biên mục tài nguyên; và nhất là thành thạo trong việc sử dụng phần mềm nguồn mở Greenstone.

4.2 Ý NGHĨA VÀ NHU CẦU

Thế giới thư viện ngày nay đang trải qua một giai đoạn thay đổi lớn. Các thư viện đại học và thư viện nghiên cứu đang đối mặt với sự phát triển mới. Cùng với sự phát triển việc sử dụng các công nghệ mới, sự gia

tăng thông tin số dưới dạng xuất bản phẩm điện tử và nguồn tin trên mạng, và sự có mặt Internet ở khắp mọi nơi.

Thực tế, việc phát triển CNTT và truyền thông 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 lên 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, sinh viên và cán bộ 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, phục vụ công tác giảng dạy, học tập và nghiên cứu khoa học. Một nhu cầu về lưu trữ, tổ chức và truy cập thông tin đã và đang đặt các thư viện trước ngưỡng cửa của thư viện số. Thông tin cần được chọn lọc, tổ chức và phục vụ dưới dạng kỹ thuật số.

Để làm được điều đó việc xây dựng, phát triển các bộ sưu tập số cần phải có một kế hoạch, một chiến lược phát triển lâu dài. Sau đây là một số bước tạo lập bộ sưu tập thông tin:

- Xác định các loại tài liệu cần thiết trong một lĩnh vực cụ thể hay trong toàn bộ sưu tập
- Đối chiếu tính xác đáng, mức độ yêu cầu của tài liệu nhằm đảm bảo sự chọn lựa tốt nhất đã được thực hiện mà không vượt quá kinh phí cho phép
- Số hóa các tài liệu đã chọn với giá hợp lý theo phương cách hiệu quả nhất
- Đóng gói các tài liệu số thành các bộ sưu tập và phân phối các bộ sưu tập tới đông đảo bạn đọc

Toàn bộ quá trình phải tập trung vào việc cung cấp tài liệu đáp ứng nhu cầu thông tin đã được xác định và phù hợp, gắn liền với hoạt động của thư viện trong quá trình xác định vai trò, nhiệm vụ, mục tiêu hoạt động của mình. Chính vì vậy mà phát triển bộ sưu tập không thể là một yếu tố hay quá trình đơn lẻ; mà ngược lại nó cần được đặt trong mối quan hệ chung

mà ở đó nhu cầu sử dụng của các nhóm đối tượng người dùng, cũng như mối tương quan với các quan hệ bên ngoài cần phải được xác định và thỏa mãn.

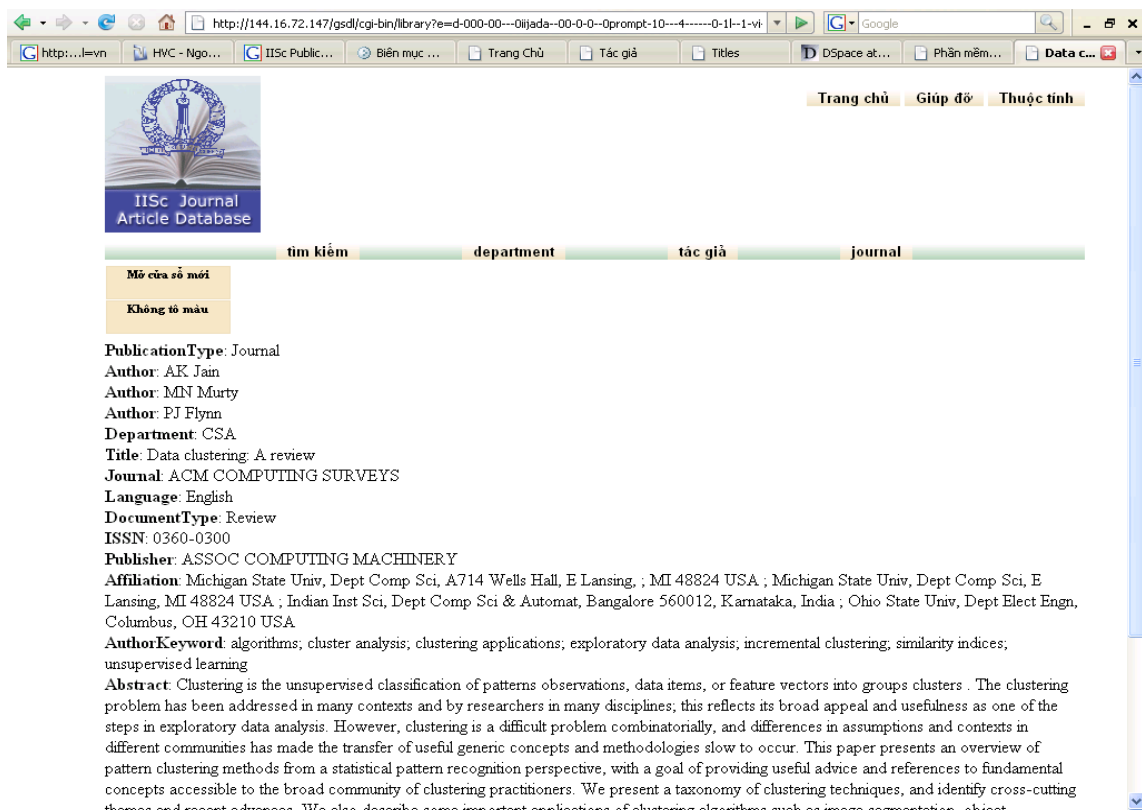
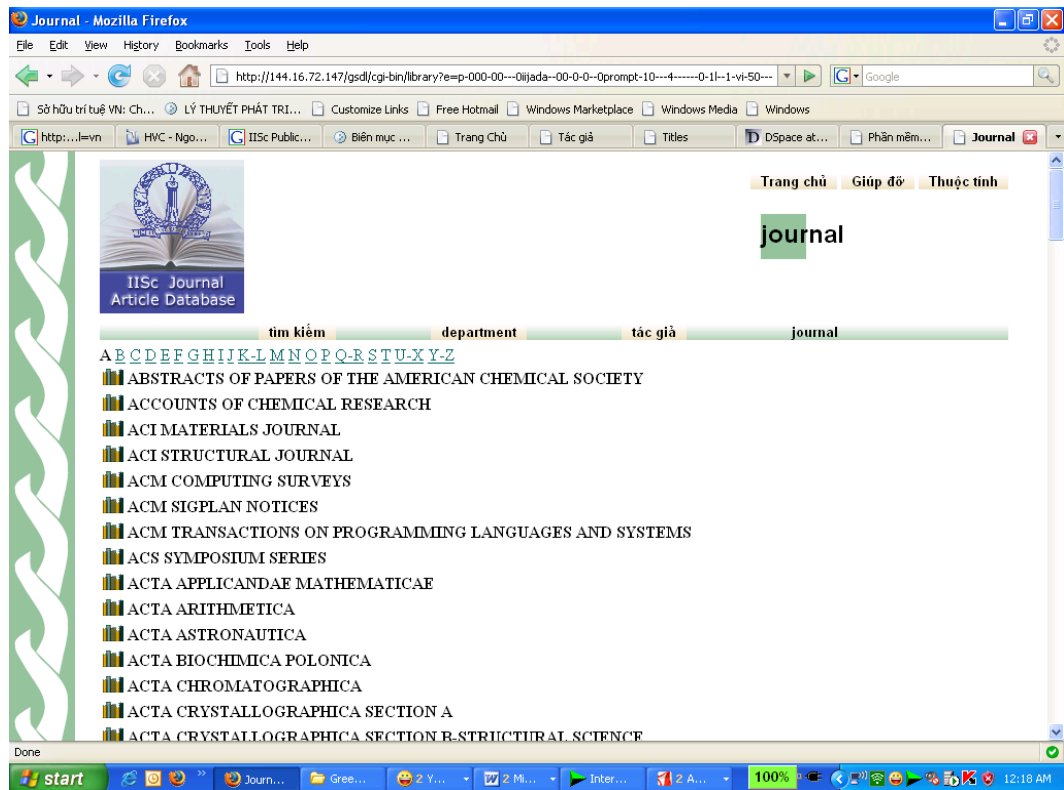
4.3 MỘT SỐ BỘ SUU TẬP MẪU

Như khái niệm đã đề cập một bộ sưu tập thông tin là một sưu tập bao gồm nhiều tài liệu dưới nhiều dạng thức khác nhau: văn bản, hình ảnh, âm thanh, hình ảnh động ...; dưới đây là một số bộ sưu tập mẫu được thể hiện theo các dạng thức khác nhau; phần lớn các bộ sưu tập mẫu này được xây dựng bằng phần mềm nguồn mở Greenstone.

4.3.1. Các bộ sưu tập theo loại hình xuất bản

Bộ sưu tập tạp chí điện tử:

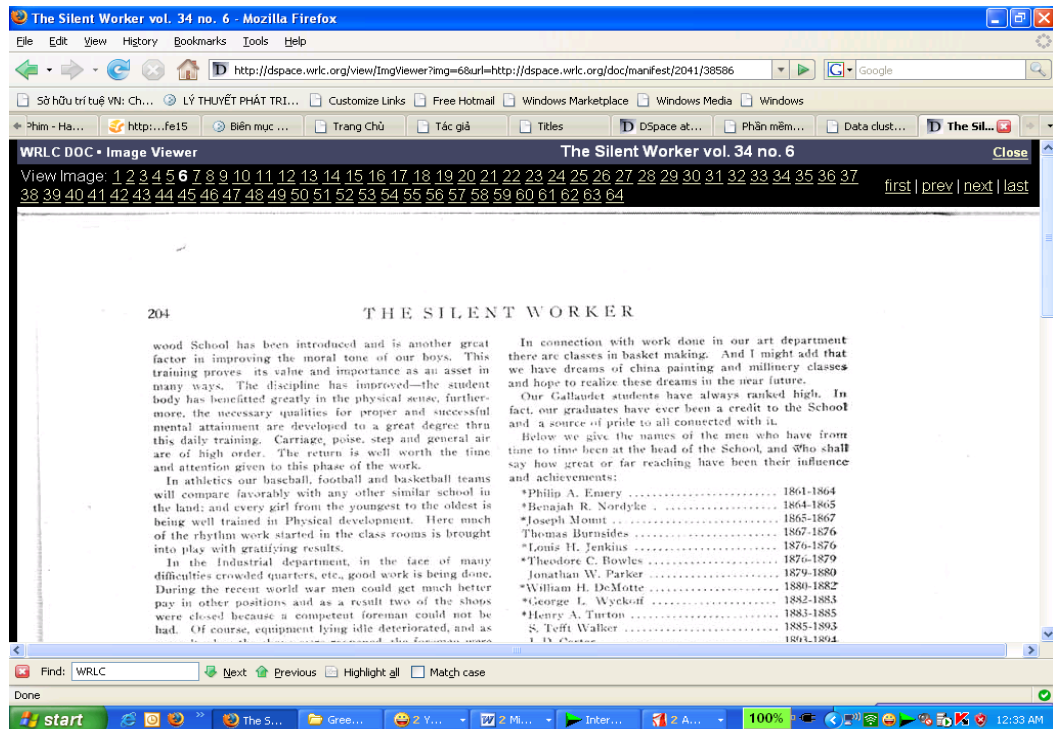
Với Greenstone ta có thể tạo ra các sưu tập về tạp chí điện tử theo từng số xuất bản mà không phá vỡ cấu trúc mục lục nguyên gốc của tạp chí điện tử đó, làm cho người sử dụng không thấy có sự khác biệt của tạp chí khi đưa vào quản lý bởi phần mềm Greenstone và tạp chí điện tử nguyên gốc xuất bản trên mạng của nhà xuất bản như: TidBITS, Journal of Artificial Intelligence Research (The New Zealand Digital Library). Ngoài ra cũng có thể sưu tập hoàn toàn là danh mục tạp chí (chỉ biên mục quản lý nguồn mà không quản lý dữ liệu toàn văn của từng bài tạp chí) mà người sử dụng khi kết nối nội dung sẽ kết nối trực tiếp đến tạp chí gốc của nhà xuất bản : IISc Publication (Indian Institute of Science, India)



Hình 4.3 và 4.4 Danh mục tạp chí của IISc Publications (Indian Institute of Science, India)

Bộ sưu tập báo: The Silent Worker (Washington Research Library)

Consortium (WRLC), USA. Bộ sưu tập này sử dụng dạng tài liệu toàn văn File format PDF của nguyên dạng hình ảnh scan báo không dùng File chuyển đổi sang HTML (tham khảo thêm tại địa chỉ: <http://www.aladin.wrlc.org/gsd/cgi-bin/library?c=gasw&a=d&c1=CL4>)



Hình 4.5 và 4.6 Bộ sưu tập Argus (Illinois Wesleyan University, USA)

Bộ sưu tập eBooks: Books From the Past (Culturenet Cymru and the Welsh Books Council) sưu tập sách toàn văn xuất bản từ đầu thế kỷ 19 đến giữa thế kỷ 20, của 2 ngôn ngữ English và Welsh. Người sử dụng có thể tra cứu nhan đề, tác giả và thời gian xuất bản. (Tham khảo thêm tại địa chỉ <http://www.booksfromthepast.org/>)



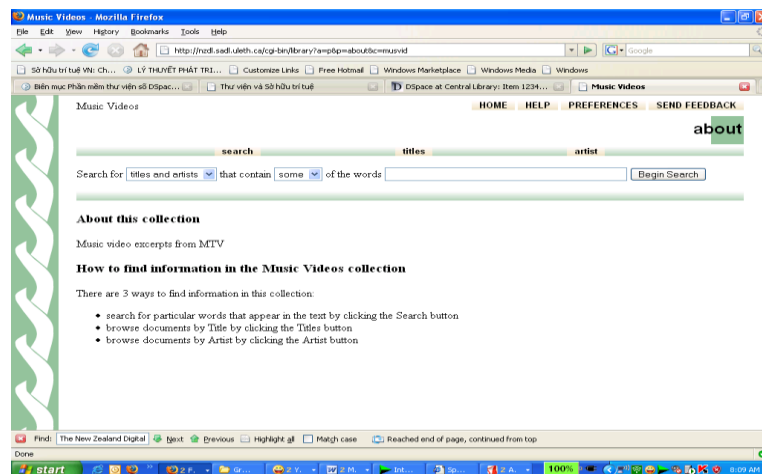
Hình 4.7 và 4.8 Ví dụ minh họa bộ sưu tập Books From the Past

4.3.2. Các bộ sưu tập theo dạng lưu trữ tài liệu

Tài liệu sưu tập là âm thanh, video:

- Sưu tập Music Library (The New Zealand Digital Library) tập hợp gần 10.000 giai điệu nhạc người sử dụng có thể nghe nhạc trực tiếp từ sưu tập này với sự hỗ trợ của các phần mềm Media: Media Player, Windows MPEC... kích hoạt tự động từ máy tính hay có thể sao lưu vào máy tính của mình.

- Sưu tập Music Video (The New Zealand Digital Library) là sưu tập Video ca nhạc từ MTV mà người sử dụng tìm kiếm theo tên bài, thể loại nhạc hay tên các ca sĩ trình diễn : Eric Clapton, George Michael và Elton John...



Hình 4.9 và 4.10 Ví dụ minh họa bộ sưu tập Music Video (Tham khảo thêm tại địa chỉ <http://nzdl.sadl.uleth.ca/cgi-bin/library>)

Tài liệu sưu tập là hình ảnh

Việc sử dụng Greenstone để xây dựng các sưu tập hình ảnh của các thư viện được giới thiệu rất nhiều trên Internet rất đa dạng cách thể hiện cấu trúc và hiển thị, có 2 sưu tập được đánh giá là sưu tập lớn và chất lượng. Trong hai sưu tập này thể hiện sự kết hợp chặt chẽ của biên mục truyền thống với dữ liệu Metadata và hiển thị biểu ghi này cùng với chế độ Thumbnail hình ảnh, và khi người sử dụng Click chọn vào đó mới hiển thị chế độ View hình ảnh ở cỡ lớn.

- The Virginia Civil War Archive của Washington Research Library Consortium (WRLC), USA. Là sưu tập có hơn 400 hình minh họa từ tạp chí Harper's Weekly từ 1861-1865

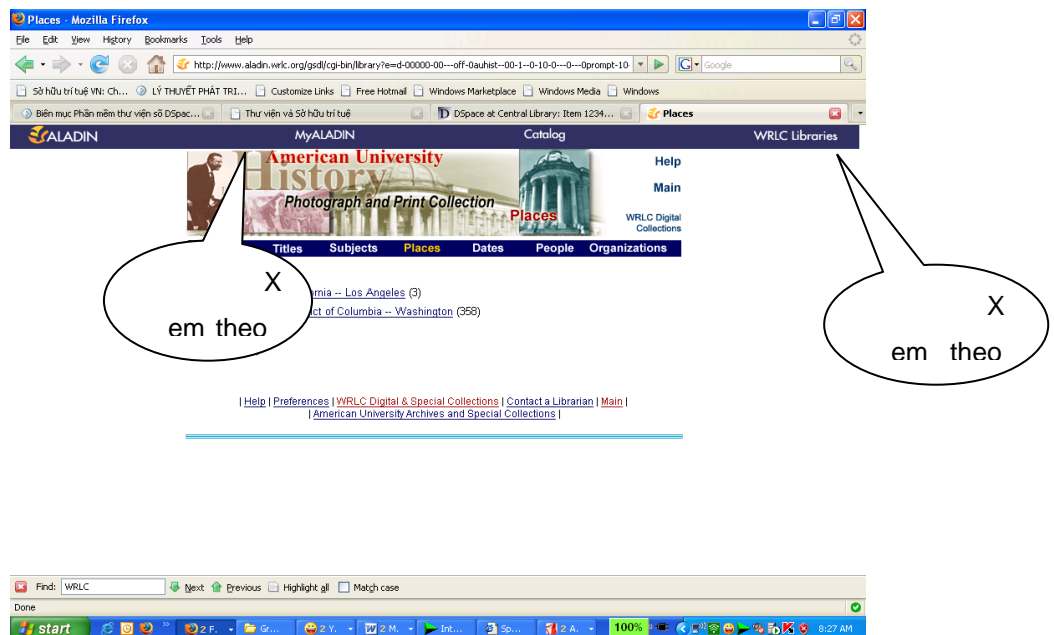
- American University History Photograph and Print Collection của Washington Research Library Consortium (WRLC), USA. Sưu tập này bao gồm các ảnh chụp về lịch sử, sinh hoạt, những sự kiện liên quan trong lịch sử nước Mỹ. Từ tòa nhà đầu tiên của trường vào năm 1896 đến 1960



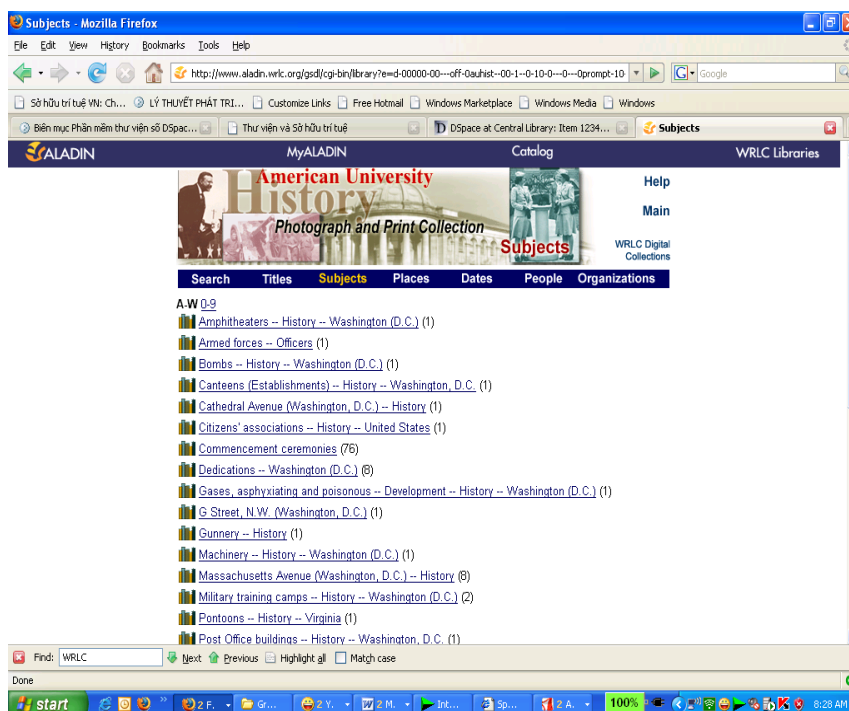
Hình 4.11 Ví dụ minh họa bộ sưu tập The Virginia Civil War Archive



Hình 4.12 Ví dụ minh họa bộ sưu tập *The Virginia Civil War Archive* (Tham khảo thêm tại địa chỉ <http://www.aladin.wrlc.org/gsd/catalog/vacw/vacw.shtml>)



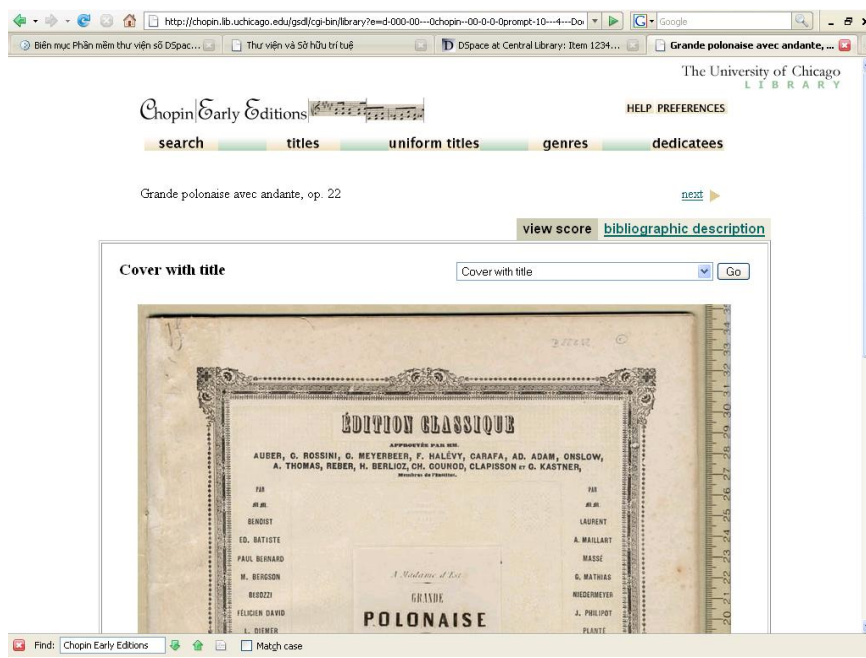
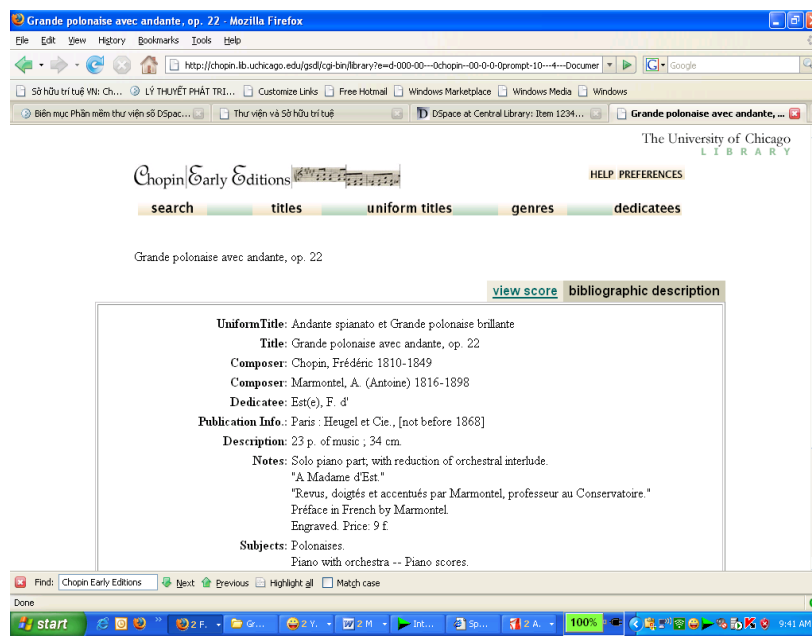
Hình 4.13 Ví dụ minh họa bộ sưu tập *American University History Photograph and Print Collection* (<http://www.aladin.wrlc.org/gsd/catalog/auhlist/auhlist.shtml>)



Hình 4.14 Ví dụ minh họa bộ sưu tập American University History Photograph and Print Collection

(<http://www.aladin.wrlc.org/gsd/collect/auhlist/auhlist.shtml>)

- Sưu tập Chopin Early Editions: Là sưu tập hơn 400 tác phẩm âm nhạc của Frédéric Chopin của The University of Chicago Library, USA. Sưu tập giữ nguyên gốc hình ảnh scan từ các bản nhạc để bảo đảm tính nguyên gốc, chúng ta có thể sao lưu và in ấn để sử dụng. Sưu tập sử dụng dạng hiển thị tài liệu nguyên bản của những trang nhạc được in hay chép tay theo từng trang, từng chương theo cấu trúc mục lục của tài liệu gốc. Các tác phẩm của Frédéric Chopin thường được xuất bản trùng lặp nhau trên thế giới với các kiểu bản in khác nhau, Greenstone cho phép người sử dụng có thể kết nối đến các tài liệu giống nhau để so sánh và đối chiếu. Trong Chopin Early Editions có thể tìm kiếm hay liệt kê theo nhan đề, thể loại và số bản nhạc.

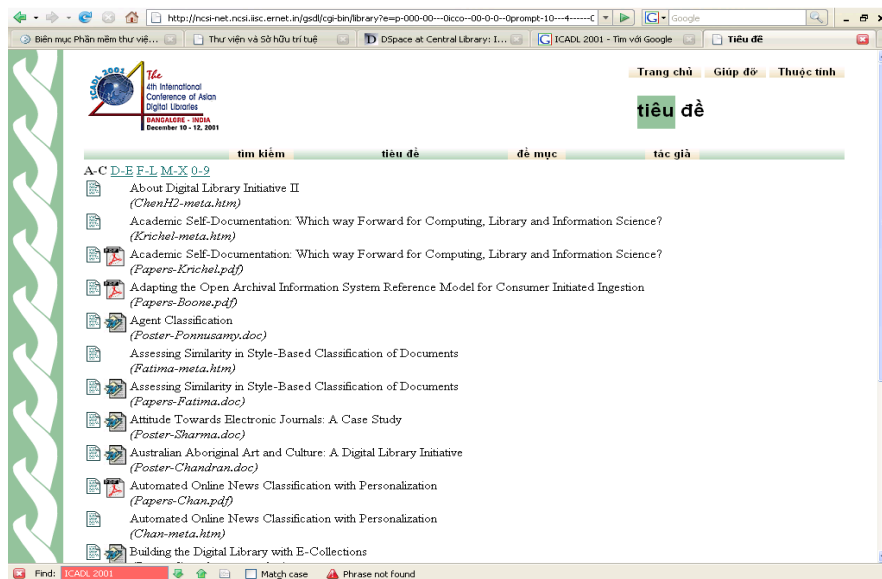
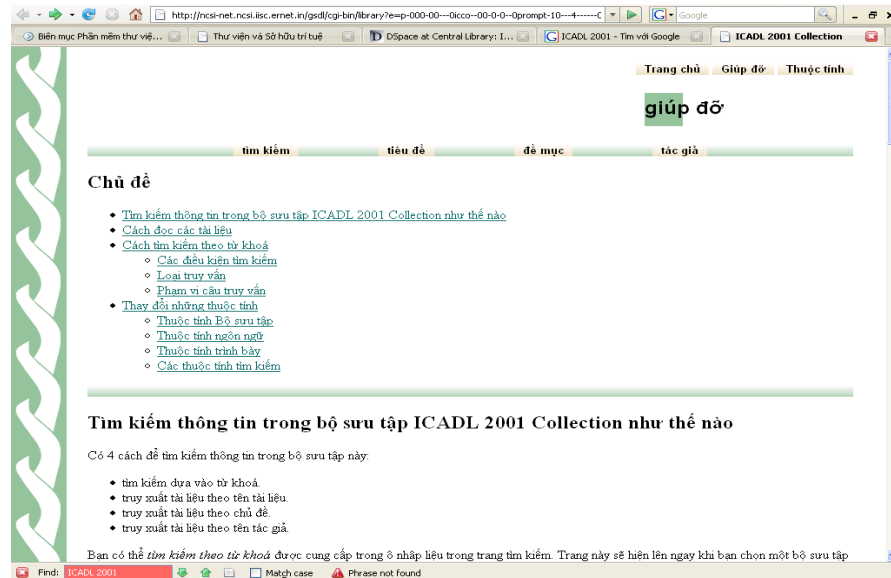


Hình 4.15 và 4.16 Ví dụ minh họa bộ sưu tập Chopin Early Editions
(Tham khảo thêm tại địa chỉ <http://chopin.lib.uchicago.edu/>)

Tài liệu sưu tập là văn bản, hình ảnh, âm thanh

ICADL 2001 Sample Collection : Là sưu tập các báo cáo khoa học và các tài liệu liên quan của Hội nghị Thư viện số Châu Á lần thứ 4 từ ngày 10-12/12/2001 tại Bangalore, Ấn Độ. Tập hợp các tài liệu theo nhiều dạng toàn văn HTML, PDF, Doc và các Slide báo cáo bằng Microsoft PowerPoint kèm theo thuyết minh MP3. Với sưu tập này người sử dụng

có thể tìm hiểu và tiếp cận hội nghị một cách chi tiết đến từng báo cáo. Trong sưu tập tìm kiếm theo nhan đề, chủ đề và tác giả báo cáo. Đây là một sưu tập rất đa dạng về loại hình tài liệu toàn văn với sự kết hợp tài liệu đa phương tiện (Tham khảo thêm tại địa chỉ: <http://ncsi-net.ncsi.iisc.ernet.in/gsd/cgi-bin/library?a=p&p=home&l=vi&w=utf-8>).



Hình 4.17 và 4.18 Ví dụ minh họa bộ sưu tập ICADL 2001

Trên đây, là một số mẫu mô hình các sưu tập được tổ chức bởi các thư viện trên thế giới cho chúng ta cái nhìn tổng quát về việc tổ chức bộ sưu tập thông tin, với nhiều cách thể hiện, thay đổi giao diện bên ngoài và giao diện tìm kiếm cũng như tổ chức dữ liệu khác nhau. Nhưng đều cùng một

mục đích chính là làm sao đưa được sưu tập đến người sử dụng với các khả năng tìm kiếm đa dạng và tiếp cận tài liệu toàn văn một cách hiệu quả nhất.

4.4 SỬ DỤNG CÁC BỘ SƯU TẬP

Dịch vụ tham khảo: Dịch vụ tham khảo cung cấp thông tin cho người sử dụng từ mọi nguồn có trong và ngoài thư viện. Sử dụng phần mềm Greenstone để tổ chức thông tin thành những bộ sưu tập chuyên ngành theo những đề tài được yêu cầu. Những bộ sưu tập này được tái đóng gói lên CD-ROM để cung cấp cho người sử dụng. Người sử dụng có thể truy tìm và lướt tìm thông tin theo từ khóa, tác giả, nhan đề, đề mục và những điểm truy cập khác trên chính CD-ROM của mình.

Công tác địa chỉ: Những tài liệu địa chỉ của địa phương bao gồm đủ mọi hình thức từ những di chỉ, mẫu vật đến công trình kiến trúc, tài liệu văn bản cổ, vv... đều có thể số hóa; rồi biên mục trên từng tài liệu và tổ chức trong một hay nhiều bộ sưu tập theo đề tài. Thật dễ dàng trong việc lưu trữ phục vụ công tác nghiên cứu cũng như quãng bá rộng rãi phục vụ công tác du lịch.

Kho tài nguyên học tập trong trường đại học: Từng giảng viên hay nhà nghiên cứu, sưu tầm tài liệu theo chuyên ngành của mình, tổ chức thành bộ sưu tập rồi đóng góp cho thư viện một bản

Lưu trữ và quản lý công văn: Trong việc lưu trữ và quản lý công văn, mỗi văn bản được xem như một tài liệu. Sử dụng chuẩn Dublin Core linh hoạt để biên mục từng tài liệu và quản lý theo cấp, đề tài, ngày, cơ quan chủ quản, người ban hành và ký công văn, vv... Người sử dụng dễ dàng truy tìm và lướt tìm Bộ sưu tập công văn cho dù lớn bao nhiêu.

4.5 CÔNG CỤ XÂY DỰNG BỘ SƯU TẬP

Hiện nay trên thế giới, xu thế phát triển thư viện số đã trở thành một phần không thể thiếu trong toàn cảnh hoạt động thông tin thư viện. Các thư viện tại Việt Nam đang áp dụng nhiều giải pháp khác nhau để quản lý tài

nguyên số:

- Sử dụng module quản lý tài nguyên số, tích hợp trong phần mềm quản trị thư viện tích hợp
- Sử dụng các phần mềm quản trị văn bản
- Sử dụng phần mềm quản lý tài nguyên số của nước ngoài (Content Pro của hãng Innovative Interface)
- Một số ít thư viện sử dụng các phần mềm tạo lập và quản trị tài nguyên số mã nguồn mở (Dspace, Greenstone,...)

Các phần mềm thương mại Việt Nam đều xây dựng phân hệ quản lý tài nguyên số trong phần mềm quản trị thư viện tích hợp. Do vậy, phần lớn các thư viện sử dụng phần mềm quản trị thư viện tích hợp thương hiệu Việt cũng có hệ thống quản lý tài nguyên số.

Tuy việc tích hợp hai phần mềm thư viện vào làm một ban đầu sẽ mang đến một số thuận lợi cho thư viện nhưng về lâu dài sẽ ảnh hưởng nghiêm trọng đến sự phát triển mở rộng và liên kết của thư viện

- Tài nguyên số chỉ có thể sử dụng nội tại tại thư viện chủ sở hữu, không có khả năng chia sẻ, trao đổi, liên kết với các hệ thống thư viện tại Việt Nam và trên thế giới
- Tài nguyên số không có tính mở cho cộng đồng (không có khả năng tìm kiếm, thống kê được bằng công cụ Google hoặc Goolge Scholar), một trong các tiêu chí để đánh giá xếp hạng đại học trên www.webometrics.info

Đối với phần mềm quản lý tài nguyên số của nước ngoài, do chi phí quá cao nên hiện nay chỉ có số ít thư viện đang sử dụng tại Việt Nam và trong tương lai cũng khó có khả năng được áp dụng rộng rãi.

Các công ty phần mềm cũng đã bắt đầu xây dựng, chuyển giao các sản phẩm hỗ trợ cho quá trình tạo lập bộ sưu tập số để xây dựng kho tài nguyên

số hóa của các thư viện: Greenstone, Dspace, Eprints, Digitools, Content Pro, Kipos, Libol 6.0 Nổi lên trong phân khúc phần mềm mã nguồn mở là phần mềm mã nguồn mở Greenstone và phần mềm mã nguồn mở Dspace đang được ứng dụng rất rộng rãi và bước đầu đã thành công tại một số thư viện như: Thư viện Đại học Khoa học Tự nhiên thành phố Hồ Chí Minh, Thư viện Hải Phú, Trung tâm Thông tin -Thư viện, Đại học Quốc gia Hà Nội, Thư viện Đại học Đà Lạt...

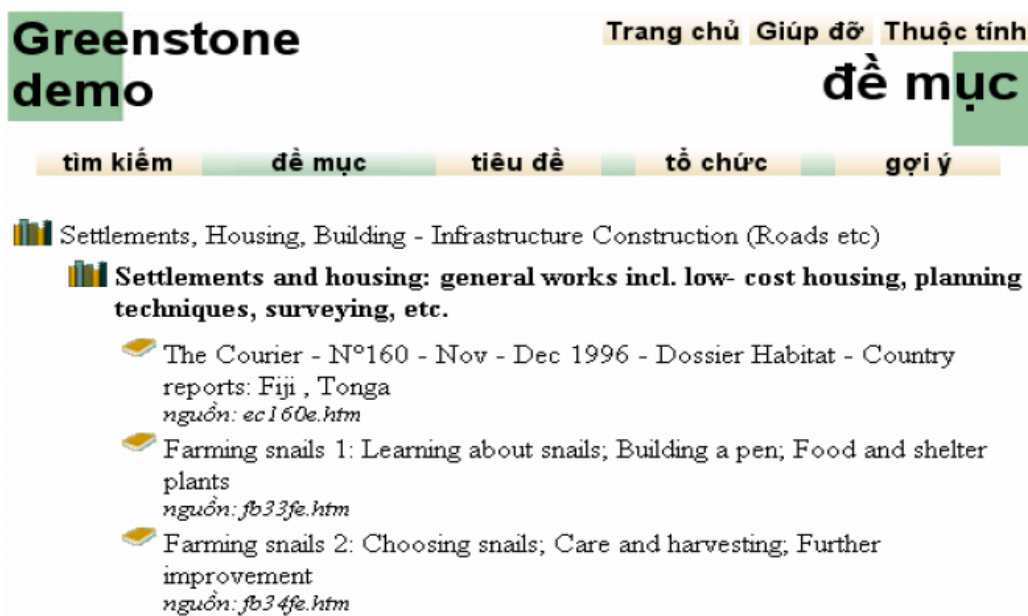
Greenstone là một bộ phần mềm dùng để xây dựng và phân phối các bộ sưu tập thư viện số, một giải pháp mới để tổ chức và xuất bản thông tin trên Internet và CD-ROM, cung cấp phương tiện dễ dàng cho người sử dụng truy tìm toàn văn và lướt tìm dựa vào metadata

Greenstone là sản phẩm của dự án New Zealand Digital Library của Đại học Waikato (<http://wichbaby.info-science.uiowa.edu/cgi-bin/library>), được phát triển và phân phối với sự tham gia của UNESCO và Human Info NGO. Đây là phần mềm nguồn mở được cung cấp theo thỏa thuận của GNU General Public License. Greenstone hiện đang được sử dụng rộng rãi trên khắp thế giới, đặc biệt trong các cơ quan của UNESCO và thư viện trong các trường đại học.

Đặc trưng của Greenstone

- Truy cập qua trình duyệt web, cả ở chế độ cục bộ (local) và từ xa (remote)
- Chạy được trên nhiều hệ điều hành: Windows, Unix, Macintosh
- Tìm kiếm toàn văn và tìm kiếm theo từng trường riêng biệt
- Khả năng trình duyệt linh động, đa dạng
- Cấu trúc duyệt tài liệu được xây dựng hoàn toàn tự động
- Tận dụng các metadata sẵn có trong tài liệu, giúp người tạo lập bộ sưu tập không phải làm bằng tay
- Khả năng linh động, dễ mở rộng hệ thống nhờ các thành phần như plugin, classifier.

- Hỗ trợ xử lý tài liệu với nhiều loại ngôn ngữ
- Cung cấp giao diện đa ngôn ngữ
- Ngoài các bộ sưu tập văn bản, hình ảnh thông thường, Greenstone còn cho phép tạo các bộ sưu tập hình ảnh, âm thanh đa phương tiện (multimedia)
 - Khả năng lưu trữ rất lớn tới hàng Gigabyte dữ liệu
 - Thêm mới bộ sưu tập đơn giản, có hiệu quả tức thì
 - Khả năng xuất bản các bộ sưu tập ra CD, với đầy đủ tính năng để có thể tự cài đặt và chạy độc lập
 - Các bộ sưu tập dễ dàng được mang, chuyển, phân phối, chia sẻ
 - Theo phong cách WYSIWYG (What you see is what you get), dễ dùng, thuận tiện
- 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



Hình 4.19 Màn hình duyệt tài liệu theo đề mục

Ngoài phần mềm GreenStone, phần mềm DSpace cũng được sử dụng phổ biến ở Việt Nam, thậm chí trong thời gian gần đây, số lượng người sử dụng DSpace có chiều hướng tăng lên.

DSpace là một bộ phần mềm hỗ trợ giải pháp xây dựng và phân phối các bộ sưu tập số hóa trên Internet. Nó cung cấp một phương thức mới trong việc tổ chức và xuất bản thông tin trên Internet. DSpace do HP và The MIT Libraries phát triển vào năm 2002, hiện nay có hơn 600 trường đại học và các tổ chức văn hoá ở trên 70 quốc gia sử dụng phần mềm số DSpace để quản lý và chia sẻ nguồn tài nguyên: Sách, Tạp chí, Luận văn và các sưu tập Hình ảnh, Âm thanh và Phim. Đây là một phần mềm mã nguồn mở cho phép các thư viện, các cơ quan nghiên cứu phát triển và mở rộng, tại địa chỉ <http://www.dspace.org>.

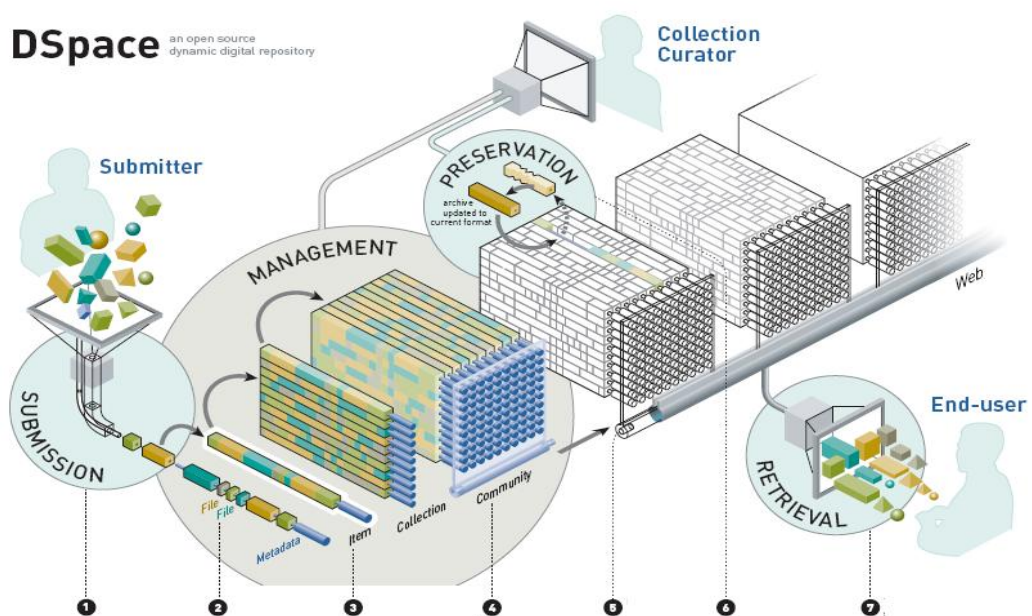
DSpace được sử dụng cơ bản như một phần mềm lưu trữ và phân phối tài liệu số. DSpace có ba vai trò chính:

- Giúp cho việc thu nhận và quản lý tài liệu được dễ dàng, bao gồm siêu dữ liệu của tài liệu
- Giúp cho việc truy cập tài liệu được dễ dàng, bằng cả việc liệt kê và tìm kiếm
- Giúp cho việc bảo quản tài liệu lâu dài.

Cấu trúc DSpace

- Dựa trên giao diện web giúp bạn dễ dàng để tạo ra một tài liệu bằng cách gửi tập tin vào DSpace. DSpace đã được thiết kế để xử lý bất kỳ định dạng tài liệu văn bản từ đơn giản đến dữ liệu và video kỹ thuật số.
- Các tập tin dữ liệu, cũng được gọi là bitstreams, được tổ chức gắn vào với nhau. Mỗi bitstream có một định dạng kỹ thuật và các thông tin kỹ thuật.
- Một tài liệu là một "lưu trữ nguyên tử" bao gồm nhóm có liên quan đến nội dung và liên kết đến các mô tả siêu dữ liệu. Siêu dữ liệu được lập chỉ mục cho việc duyệt và tìm kiếm. Các tài liệu được tổ chức vào các bộ sưu tập.

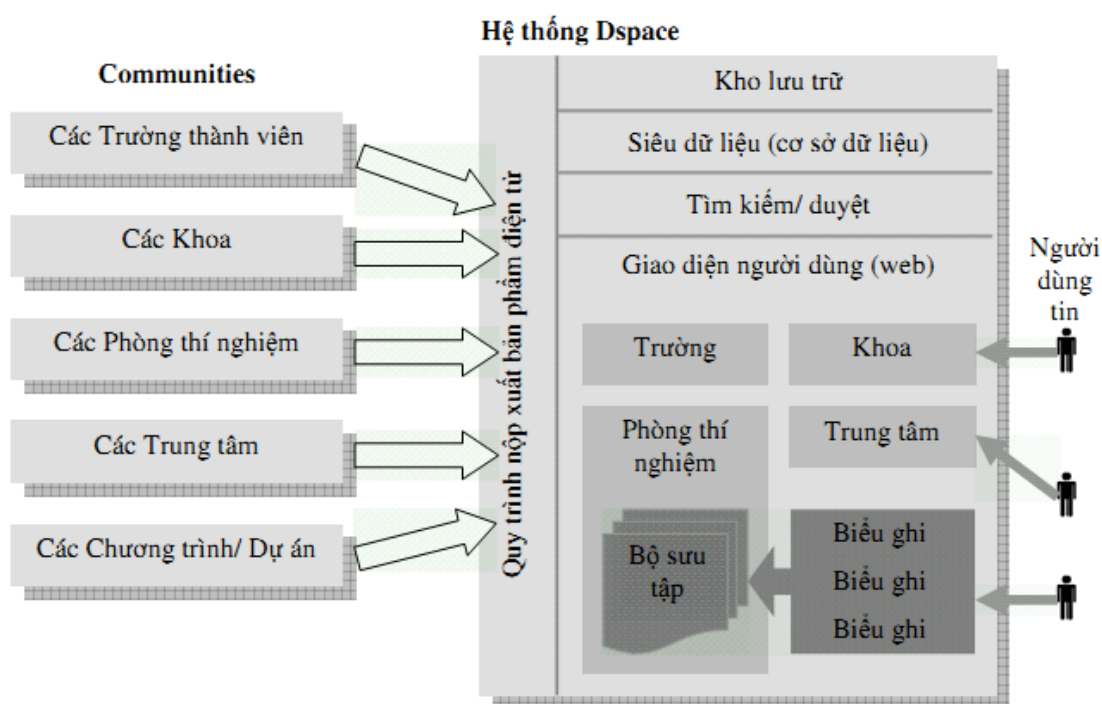
- Một đơn vị là cấp cao nhất về nội dung trong DSpace. Chúng tương ứng với các thành phần của tổ chức như các phòng, ban, khoa, ngành, phòng thí nghiệm hoặc các trung tâm ...
- Kiến trúc modul của DSpace modular cho phép mở rộng ranh giới giữa các đơn vị và bộ sưu tập.
- DSpace duy trì tập tin một cách đáng tin cậy để cung cấp chức năng lưu trữ, nơi có thể truy cập các tập tin được lưu giữ như với các định dạng kỹ thuật, công nghệ, truyền thông phát triển theo thời gian cho nhiều loại tài liệu.
- Giao diện người dùng cuối ở dạng web hỗ trợ việc duyệt và tìm kiếm các tài liệu được lưu trữ. Sau khi một tài liệu được chọn, các tập tin có thể được hiển thị trong một trình duyệt Web, trong khi các định dạng khác có thể được tải về và mở bằng một chương trình ứng dụng phù hợp.



Hình 4.20 Cấu trúc của Dspace

DSpace được thiết kế để tạo thuận lợi cho các cá nhân có thể đóng góp các tài nguyên số vào hệ thống một cách dễ dàng. Mô hình thông tin

của hệ thống được xây dựng xung quanh ý tưởng “Communities” tổ chức các đơn vị trực thuộc của một tổ chức nghiên cứu, một trường đại học có nhu cầu quản lý thông tin đặc biệt. Trong trường hợp của MIT (một trường đại học nghiên cứu lớn), “Communities” được định nghĩa là các trường thành viên, các khoa, phòng thí nghiệm, và các trung tâm của MIT. Mỗi “communities” có thể thích ứng với hệ thống để đáp ứng các nhu cầu cụ thể của đơn vị và quản lý quá trình nộp các xuất bản phẩm điện tử.



Hình 4.21 Mô hình tổ chức thông tin trong DSpace

Ngoài các phần mềm mã nguồn mở như đã đề cập ở trên còn có các sản phẩm thương mại như: Content Pro IRX của hãng Innovative, CONTENTdm của OCLC, Digital Commons của Bepress, DigiTool của ExLibris, Open Repository của BioMed, VITAL của VTLS....

Đây là những giải pháp hoàn chỉnh cho việc quản lý các tài nguyên lưu trữ, các bộ sưu tập đặc biệt và các dạng tài liệu số khác. Với khả năng tìm kiếm mạnh và thiết kế cho việc thu thập các siêu dữ liệu, giải pháp cung cấp khả năng truy cập, nghe, nhìn toàn diện cho bạn đọc.

Một số tính năng nổi bật của các sản phẩm thương mại so với các hệ thống mã nguồn mở:

- Tìm kiếm
 - + Hỗ trợ tìm kiếm theo từ khóa và đánh giá kết quả tìm kiếm theo mức độ chính xác
 - + Có thể tìm kiếm theo từng bộ sưu tập
 - + Nhiều lựa chọn sắp xếp trong từng bộ sưu tập và trong kết quả tìm kiếm
 - + Đánh chỉ mục toàn văn cho các dạng file PDF, Word và các loại văn bản
 - + Có thể tích hợp với nền tảng tìm kiếm tập trung như: Encore, Primo, EBSCO Discovery services ...
 - Kiểm soát
 - + Có thể xác thực thông qua LDAP ở tầng các bộ sưu tập
 - + Giới hạn truy cập theo địa chỉ IP
 - + Từng tài liệu có thể được gắn mật khẩu
 - + Các tài liệu hoặc các bộ sưu tập có thể được giới hạn đối với người dùng công cộng
 - + Có thể tùy chỉnh “watermarking” cho các file ảnh
 - Dễ dàng cho nhân viên thư viện
 - + Quy trình đưa tài liệu số lên hệ thống dựa trên nền tảng Web
 - + Tự động và có kiểm soát hệ thống từ vựng cho trường tiêu đề
 - + Có thể cập nhật nhiều tài liệu vào bộ sưu tập cùng một lúc
 - + Có thể chuyển đổi dữ liệu dạng MARC và CSV
 - + Công cụ chuyển đổi cho Metadata
 - + Thống kê sử dụng: có thể theo dõi số lần được nhấp vào ở tất cả các tầng của kho dữ liệu, tên miền, các từ khóa được tìm kiếm
 - Công nghệ và tiêu chuẩn
 - + OAI-PMH (Open archives Initiative Protocol for Metadata

Harvesting)

- + Hỗ trợ hoàn toàn Unicode
- + Chuẩn METS, DublinCore cho siêu dữ liệu
- + Các dữ liệu video và âm thanh được truyền tải theo cơ chế streaming
- + Giao diện quản lý trên nền Java có thể chạy trên các hệ điều hành khác nhau như Windows, Mac, Linux
- Các định dạng hỗ trợ, đa dạng như: định dạng văn bản, định dạng hình ảnh, định dạng âm thanh, định dạng video ...

4.6 BIÊN MỤC TÀI LIỆU SỐ

4.6.1. Biên mục Analog

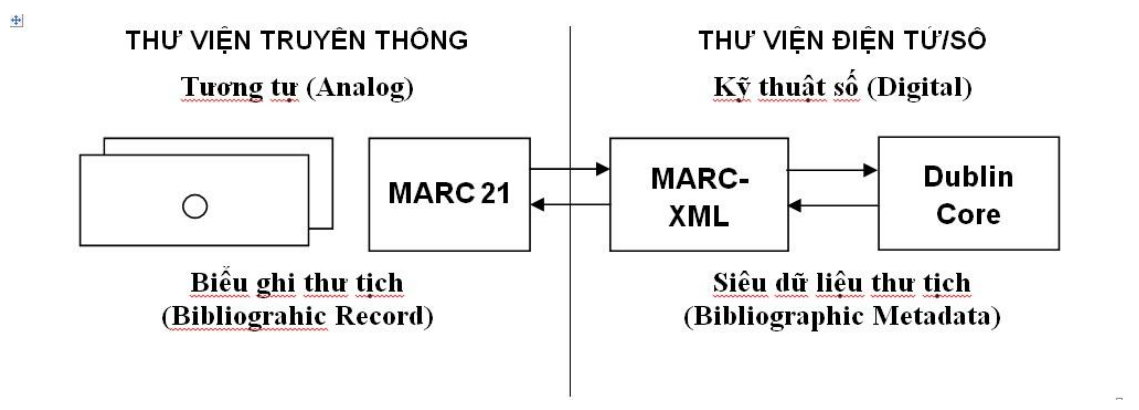
- Trong thư viện truyền thống, người ta biên mục để tạo nên những biểu ghi thư tịch nhằm xây dựng hệ thống tra cứu qua mục lục phiếu. Biểu ghi thư tịch hay mục lục phiếu miêu tả lý lịch của tài liệu: nhan đề, tác giả, đề mục, xuất bản, vv... Khi sử dụng máy tính, biểu ghi thư tịch này được biểu thị bằng một dạng thức máy đọc được (MARC). Cách biên mục này chỉ thể hiện được dạng thư tịch tức lý lịch chứ không có toàn văn và đa phương tiện, được gọi là biên mục theo dạng tương tự (analog).
- Phiếu mục lục và kho sách tách rời nhau
- Biểu ghi thư tịch MARC nhằm trao đổi dữ liệu qua máy tính hay trên mạng máy tính

4.6.2. Biên mục Digital

- Trong môi trường số, dữ liệu được đóng gói bằng ngôn ngữ XML. Cách biên mục phải thay đổi qua môi trường Web, nghĩa là các dữ liệu thư tịch phải được đóng gói, người ta gọi là biên mục theo dạng kỹ thuật số (digital). Các biểu ghi thư tịch trở thành siêu dữ liệu thư tịch - metadata. Vậy Metadata chính là phiếu mục lục miêu tả lý lịch tài liệu

được phát sinh tự động trong môi trường số.

- Metadata gắn liền với nội dung văn bản
- MARC chuyển đổi thành MARC-XML
- Chuẩn Dublin Core đáp ứng yêu cầu tổ chức tài nguyên số



Hình 4.22 Biên mục Analog và Digital

Khổ mẫu tiêu chuẩn siêu dữ liệu Dublin core là một tập hợp các yếu tố đơn giản nhưng hữu hiệu trong việc mô tả một loạt nguồn tin trên mạng. DC gồm 15 yếu tố, mà ngữ nghĩa được xác lập của nhiều chuyên gia. Các yếu tố mô tả trong DC đều không bắt buộc và có thể lặp, ngoài ra còn có một số lượng hạn chế các từ hạn định và định ngữ có thể sử dụng để tiếp tục tinh chỉnh ý nghĩa của các yếu tố.

4.6.3. Chuyển đổi MARC sang Dublin Core

Vì dữ liệu của MARC phong phú hơn Dublin Core nhiều nên trong việc chuyển đổi MARC sang Dublin Core, một trường riêng lẻ của MARC được ghép với một thành phần của Dublin Core có thể được lặp lại. Bảng dưới đây là ví dụ đối với những trường liên quan đến chủ đề.

Thành phần DC	Từ hạn định DC	Trường MARC
Chủ đề (Subject)	LCSH	600, 610, 611, 630, 650
Chủ đề	MeSH	600, 610, 611, 630, 650

(Subject)			
Chủ đề (Subject)	LCC		050
Chủ đề (Subject)	DDC		082
Chủ đề (Subject)	UDC		080

Bảng 4.1 Chuyển đổi MARC sang Dublin Core

Ngày nay sự đối nghịch giữa hai chuẩn MARC và Dublin Core không còn tồn tại mà ngày càng nhiều chuyên gia tìm những giải pháp để chuyển đổi nhau: MARC sang Dublin Core và Dublin Core sang MARC, chẳng hạn như tại website của Văn phòng Chuẩn MARC và phát triển mạng lưới của Thư viện Quốc hội Hoa Kỳ có hướng dẫn các phương thức chuyển đổi. Thậm chí còn có nhiều phần mềm chuyển đổi được phổ biến rộng rãi, chẳng hạn như những phần mềm MarcEdit của Đại học Bang Oregon, Hoa Kỳ mà ta có thể download để dùng miễn phí tại website: <http://oregonstate.edu/~reaset/marcedit/html/downloads.html>.

CÂU HỎI ÔN TẬP CHƯƠNG 4

1. Phân tích khái niệm và ý nghĩa của bộ sưu tập số
2. Giới thiệu các công cụ xây dựng bộ sưu tập số!
3. Phân tích các dạng biên mục tài liệu số và quy trình chuyển đổi dữ liệu

CHƯƠNG 5: CÁC VẤN ĐỀ VỀ SỞ HỮU TRÍ TUỆ VÀ BẢN QUYỀN

Mục tiêu đào tạo: Trang bị cho người học kiến thức chung về sở hữu trí tuệ và bản quyền trong lĩnh vực tài nguyên thông tin số.

5.1 MỘT SỐ KHÁI NIỆM VỀ SỞ HỮU TRÍ TUỆ

Trong thời đại của nền kinh tế tri thức, sở hữu trí tuệ là một trong những vấn đề cần được quan tâm. Đặc biệt là bảo hộ sở hữu trí tuệ trên mạng đã và đang đặt cho những người làm công tác quản lý thông tin nhiều thách thức, đòi hỏi phải nắm vững các vấn đề về sở hữu trí tuệ và bản quyền để có thể vận dụng một cách đúng đắn trong công tác thông tin thư viện.



Sở hữu trí tuệ (SHTT): là thuật ngữ mô tả những ý tưởng, sáng chế, công nghệ, tác phẩm nghệ thuật, âm nhạc và văn học, những cái vô hình khi mới được tạo ra nhưng trở nên đáng giá dưới dạng sản phẩm hữu hình.

Quyền sở hữu trí tuệ bao gồm các quyền sở hữu đối với sản phẩm của hoạt động trí tuệ và tinh thần như tác phẩm văn học, nghệ thuật, tác phẩm khoa học, sáng chế, kiểu dáng công nghiệp, thiết kế bố trí mạch tích hợp bán dẫn, nhãn hiệu, tên thương mại, bí mật kinh doanh, chỉ dẫn địa lý và giống cây trồng

Các đối tượng sở hữu trí tuệ được nhà nước bảo hộ bao gồm: Đối tượng quyền tác giả: Tác phẩm văn học, nghệ thuật và tác phẩm khoa học; đối tượng liên quan đến quyền tác giả như: cuộc biểu diễn, bản ghi âm, ghi hình, chương trình phát sóng tín hiệu vệ tinh mang chương trình được mã hoá. Đối tượng quyền sở hữu công nghiệp: Sáng chế, kiểu dáng công nghiệp, nhãn hiệu, chỉ dẫn địa lý, bí mật kinh doanh, tên thương mại, thiết kế bố trí mạch tích hợp bán dẫn; Đối tượng quyền đối với giống cây trồng:

Giống cây trồng và vật liệu nhân giống.

Sở hữu trí tuệ bao gồm ba nhóm: nhóm quyền tác giả (bản quyền tác giả), nhóm sở hữu công nghiệp (quyền sở hữu công nghiệp) và giống cây trồng (Điều 3 Luật SHTT).

Quyền tác giả là quyền sở hữu của cá nhân, pháp nhân đối với tác phẩm văn học, nghệ thuật và khoa học do mình sáng tạo ra hoặc sở hữu, thường gọi là bản quyền tác giả.

Quyền tác giả thường được xác lập đối với những người sáng tạo ra tác phẩm gốc về văn học, kịch, âm nhạc, nghệ thuật.

Quyền này cho phép người sáng tạo kiểm soát được việc khai thác, sao chép, cải biên, công bố tác phẩm của mình.

Đối tượng quyền liên quan đến quyền tác giả bao gồm các cuộc biểu diễn, bản ghi âm, ghi hình, chương trình phát sóng, tín hiệu vệ tinh mang chương trình mã hoá (Điều 4.3 Luật SHTT)

Tác phẩm là sản phẩm sáng tạo trong lĩnh vực văn học, nghệ thuật và khoa học thể hiện bằng bất kỳ phương tiện hay hình thức nào.

Tác phẩm còn có thể thể hiện bằng các ký hiệu thay cho chữ viết như chữ nổi, ký hiệu tốc ký, ký hiệu tương tự khác mà có thể sao chép bằng nhiều hình thức khác nhau (Điều 9 Nghị định 100/2000/NĐ-CP).

Tác phẩm văn học, nghệ thuật và khoa học được nhà nước bảo hộ là các loại tác phẩm:

- Tác phẩm văn học, khoa học, sách giáo khoa, giáo trình và tác phẩm khác được thể hiện dưới dạng chữ viết hoặc ký tự khác.

- Bài giảng, bài phát biểu và bài nói khác được thể hiện dưới dạng ngôn ngữ nói và được định hình dưới dạng vật chất nhất định (Điều 10 Nghị định 100/2006/NĐ-CP).

- Tác phẩm sân khấu bao gồm thuộc loại hình biểu diễn như kịch nói, nhạc vũ kịch, ca kịch, kịch câm, xiếc, múa, múa rối và các loại hình tác phẩm sân khấu khác (Điều 13 Nghị định 100/2006/NĐ-CP).

- Tác phẩm điện ảnh và tác phẩm được tạo ra theo phương pháp tương tự là tác phẩm được hợp thành bằng hàng loạt hình ảnh liên tiếp tạo nên hiệu ứng chuyển động kèm theo hoặc không kèm theo âm thanh, được thể hiện trên chất liệu nhất định có thể phân phối, truyền đạt tới công chúng bằng các thiết bị kỹ thuật, công nghệ. Tác phẩm điện ảnh gồm các loại phim truyện, phim tài liệu, phim khoa học, phim hoạt hình và các loại hình tương tự khác (Điều 16 Nghị định 100/2006/NĐ-CP).

- Tác phẩm báo chí gồm: phóng sự, ghi nhanh, tường thuật, phỏng vấn, phản ảnh, điều tra, bình luận, chuyên luận, ký báo chí và các thể loại khác nhằm đăng, phát trên báo in, báo nói, báo hình báo, điện tử hoặc các phương tiện khác (Điều 11 Nghị định 100/2006/NĐ-CP).

- Tác phẩm âm nhạc là tác phẩm được thể hiện dưới dạng nhạc nốt trong bản nhạc hoặc các ký tự âm nhạc khác có hoặc không có lời, không phụ thuộc vào việc trình diễn hay không (Điều 12 Nghị định 100/2006 NĐ-CP).

- Tác phẩm kiến trúc là bản vẽ thiết kế thể hiện ý tưởng sáng tạo về ngôi nhà, công trình xây dựng, quy hoạch không gian đã hoặc chưa xây dựng. Tác phẩm kiến trúc bao gồm các bản vẽ, thiết kế mặt bằng, mặt đứng, mặt cắt, phối cảnh, thể hiện ý tưởng sáng tạo về ngôi nhà, công trình, tổ hợp công trình kiến trúc, tổ chức không gian, kiến trúc cảnh quan một vùng, đô thị, hệ thống đô thị, khu chức năng, đô thị, khu dân cư (Điều 17 Nghị định 100/2006/NĐ-CP).

- Tác phẩm tạo hình là tác phẩm được thể hiện bằng đường nét, hình khối, màu sắc, bố cục như hội họa, đồ họa, điêu khắc, nghệ thuật sắp đặt và các hình thức thể hiện khác tồn tại dưới dạng độc bản. Đối với loại hình đồ họa có thể thể hiện đến phiên bản thứ 50, có số thứ tự và có chữ ký của tác giả (Điều 15.1 Nghị định 100/2006/Nđ-CP).

- Mỹ thuật ứng dụng là tác phẩm được thể hiện bằng đường nét, hình khối, màu sắc, bố cục với tính năng hữu ích, có thể gắn liền với đồ

vật hữu ích, được sản xuất hàng loạt, bằng tay hoặc bằng máy như biểu trưng, hàng thủ công mỹ nghệ, hình thức thể hiện trên sản phẩm, bao bì sản phẩm (Điều 15.2 Nghị định 100/2006/NĐ-CP).

- Tác phẩm nhiếp ảnh là tác phẩm thể hiện hình ảnh của thế giới khách quan trên vật liệu bắt sáng hoặc trên phương tiện mà hình ảnh được tạo ra hay có thể tạo ra bằng bất cứ phương tiện kỹ thuật nào. Có thể là phương pháp hoá học, điện tử, hoặc phương pháp khác (Điều 16 Nghị định 100/2006/NĐ-CP).

- Bản hoạ đồ, bản vẽ, sơ đồ, bản đồ có liên quan đến địa hình, các loại công trình khoa học (Điều 18 Nghị định 100/2006/NĐ-CP).

- Chương trình máy tính, sưu tập dữ liệu.

- Tác phẩm văn học, nghệ thuật dân gian (Điều 14 luật SHTT).

Tác giả của tác phẩm: Những người trực tiếp sáng tạo ra một phần hoặc toàn bộ tác phẩm văn học, nghệ thuật và tác phẩm khoa học là tác giả của tác phẩm. Tác giả bao gồm:

- Cá nhân người Việt Nam có tác phẩm được bảo hộ.

- Cá nhân người nước ngoài có tác phẩm được sáng tạo và thể hiện dưới hình thức vật chất nhất định tại Việt Nam; có tác phẩm được công bố lần đầu tiên ở Việt Nam; có tác phẩm được bảo hộ tại Việt Nam (Điều 8 Nghị định 100/2006/NĐ-CP).

- Ngoài ra, những người sau đây cũng được công nhận là tác giả như: người dịch tác phẩm từ ngôn ngữ này sang ngôn ngữ khác được coi là tác giả tác phẩm dịch; người phóng tác từ tác phẩm đã có, người cải biên, chuyển thể tác phẩm từ loại hình nghệ thuật này sang loại hình khác được coi là tác giả của tác phẩm phóng tác, cải biên, chuyển thể; người biên soạn, chủ giải, tuyển chọn tác phẩm của người khác thành tác phẩm có tính sáng tạo được coi là tác giả biên soạn, chủ giải, tuyển chọn (Điều 13 Luật SHTT).

Chủ sở hữu tác phẩm bao gồm:

- Tác giả là chủ sở hữu toàn bộ hoặc một phần tác phẩm do mình sáng tạo, trừ trường hợp tác phẩm sáng tạo theo nhiệm vụ được giao hoặc theo hợp đồng.

- Các đồng tác giả là chủ sở hữu chung tác phẩm do họ cùng sáng tạo, trừ trường hợp tác phẩm sáng tạo theo nhiệm vụ được giao hoặc theo hợp đồng.

- Các cơ quan, tổ chức giao nhiệm vụ cho tác giả là chủ sở hữu toàn bộ hoặc một phần tác phẩm do tác giả tạo ra theo nhiệm vụ mà cơ quan hoặc tổ chức giao.

- Cá nhân hoặc tổ chức giao kết hợp đồng sáng tạo với tác giả là chủ sở hữu một phần hoặc toàn bộ tác phẩm do tác giả sáng tạo theo hợp đồng.

- Người thừa kế hợp pháp của tác giả là chủ sở hữu tác phẩm được thừa kế trong trường hợp tác giả đồng thời là chủ sở hữu tác phẩm đó.

- Những người được chủ sở hữu tác phẩm chuyển giao các quyền của mình theo hợp đồng thì những người này là chủ sở hữu quyền được chuyển giao.

Quyền tác giả bao gồm quyền nhân thân và quyền tài sản của tác giả đối với tác phẩm do mình sáng tạo.

- Quyền nhân thân gồm các quyền: Đặt tên cho tác phẩm; đứng tên thật hoặc bút danh trên tác phẩm, nêu tên thật hoặc bút danh khi tác phẩm được công bố, sử dụng; công bố, phổ biến hoặc cho người khác công bố, phổ biến tác phẩm của mình; bảo vệ sự toàn vẹn của tác phẩm, không cho phép người khác sửa chữa, cắt xén hoặc xuyên tạc tác phẩm dưới bất kỳ hình thức nào gây phương hại đến danh dự và uy tín của tác giả.

- Tác giả không đồng thời là chủ sở hữu tác phẩm có các quyền nhân thân đối với tác phẩm mà mình sáng tạo gồm: đặt tên cho tác phẩm; đứng tên thật hoặc bút danh trên tác phẩm; bảo vệ sự toàn vẹn tác phẩm,

cho phép hoặc không cho phép người khác sửa đổi nội dung của tác phẩm.

- Chủ sở hữu tác phẩm không đồng thời là tác giả có các quyền nhân thân đối với tác phẩm gồm: công bố, phổ biến hoặc cho người khác công bố, phổ biến tác phẩm thuộc sở hữu của mình, trừ trường hợp giữa tác giả và chủ sở hữu có thỏa thuận khác; cho hoặc không cho người khác sử dụng tác phẩm thuộc quyền sở hữu của mình, trừ trường hợp giữa tác giả và chủ sở hữu có thỏa thuận khác (Điều 19 Luật SHTT, Điều 22 Nghị định 100/106/NĐ-CP).

- Quyền tài sản bao gồm: đối với tác giả đồng thời là chủ sở hữu tác phẩm có các quyền như được hưởng nhuận bút; được hưởng thù lao khi tác phẩm được sử dụng; được hưởng lợi ích vật chất từ việc cho người khác sử dụng tác phẩm dưới các hình thức như xuất bản, tái bản, trưng bày, triển lãm, biểu diễn, phát thanh, truyền hình, ghi âm, ghi hình, chụp ảnh; dịch, phóng tác, cải biên, chuyển thể, cho thuê; được nhận giải thưởng đối với tác phẩm mà mình là tác giả.

- Đối với tác giả không đồng thời là chủ sở hữu tác phẩm có các quyền tài sản đối với tác phẩm mà mình là tác giả gồm: được hưởng nhuận bút; được hưởng thù lao khi tác phẩm được sử dụng; được nhận giải thưởng đối với tác phẩm mà mình là tác giả.

- Chủ sở hữu không đồng thời là tác giả được hưởng lợi ích vật chất từ việc sử dụng tác phẩm dưới các hình thức xuất bản, tái bản, trưng bày, triển lãm, biểu diễn, phát thanh, truyền hình, ghi âm, ghi hình, chụp ảnh; dịch, phóng tác, cải biên, chuyển thể, cho thuê (Điều 22 Luật SHTT, Điều 23 Nghị định 100/2006/NĐ-CP).

Các hành vi sau đây bị coi là hành vi xâm phạm quyền tác giả:

- Chiếm đoạt quyền tác giả đối với tác phẩm văn học, nghệ thuật, khoa học.

- Mạo danh tác giả.

- Công bố, phân phối tác phẩm mà không được phép của tác giả, đồng tác giả trong trường hợp có đồng tác giả.
 - Sửa chữa, cắt xén hoặc xuyên tạc tác phẩm dưới bất kỳ hình thức nào gây phương hại đến danh dự, uy tín của tác giả.
 - Sao chép tác phẩm mà không được phép của tác giả, chủ sở hữu quyền (trừ trường hợp có quy định khác).
 - Làm tác phẩm phái sinh nhưng không được phép của tác giả, chủ sở hữu quyền (trừ trường hợp có quy định khác).
 - Sử dụng tác phẩm mà không được phép của chủ thể quyền, không trả nhuận bút, thù lao, quyền lợi vật chất (trừ trường hợp có quy định khác).
 - Cho thuê tác phẩm mà không trả tiền nhuận bút, thù lao và quyền lợi vật chất khác cho tác giả, hoặc chủ sở hữu quyền tác giả.
 - Nhân bản, sản xuất bản sao, phân phối trưng bày hoặc truyền đạt tác phẩm đến công chúng qua mạng truyền thông hoặc các phương tiện kỹ thuật số mà không được phép của chủ sở hữu quyền tác giả.
 - Xuất bản tác phẩm không được phép của chủ sở hữu quyền tác giả.
 - Cố ý huỷ bỏ, làm vô hiệu các biện pháp kỹ thuật do chủ sở hữu quyền thực hiện để bảo vệ quyền.
 - Cố ý xoá bỏ, thay đổi thông tin quản lý quyền dưới hình thức điện tử có trong tác phẩm.
 - Sản xuất, lắp ráp, biến đổi, phân phối, nhập khẩu, xuất khẩu hoặc cho thuê thiết bị khi biết hoặc có cơ sở để biết thiết bị đó làm vô hiệu biện pháp kỹ thuật để bảo vệ quyền của tác giả đối với tác phẩm của mình.
 - Làm và bán tác phẩm mà chữ ký của tác giả bị giả mạo.
 - Xuất khẩu, nhập khẩu bản sao tác phẩm mà không được chủ sở hữu quyền tác giả cho phép (Điều 28 Luật SHTT).
- Tham khảo thêm các tài liệu:
- Luật sở hữu trí tuệ Việt Nam

- Công ước Bern

5.2 CÁC VẤN ĐỀ VỀ BẢN QUYỀN

BẢN QUYỀN. Là quyền duy nhất do chính phủ trao cho tác giả hoặc thừa nhận đối với tác giả một tác phẩm nhằm ngăn chặn những người khác không được phép in sao, sửa đổi, phát hành ra công chúng, biểu diễn hay trình diễn trước công



chúng. Bản quyền không bảo vệ những ý tưởng trừu tượng; bản quyền chỉ bảo vệ những hình thức diễn đạt cụ thể trong một tác phẩm. Để được bảo hộ thì tác phẩm được cấp bản quyền phải có tính nguyên bản và một chút tính sáng tạo.

Trong một thư viện truyền thống, quyền sở hữu tài liệu là quan trọng; nhưng trong lĩnh vực lưu hành tài nguyên điện tử, quyền sở hữu trí tuệ, cụ thể là quyền tác giả hay bản quyền là quan trọng hơn. Suu tầm thông tin và làm cho thông tin đó trở nên phổ biến hơn đối những người khác là một điều liên quan đến vấn đề xã hội, và những người xây dựng thư viện số phải am hiểu quyền sở hữu trí tuệ để hành động một cách có trách nhiệm và đúng luật xung quanh những ứng dụng cụ thể của họ.

Thư viện số có thể làm cho việc truy cập trở nên rộng rãi hơn thư viện truyền thống. Và chính điều này đã nảy sinh ra nhiều vấn đề: truy cập thông tin trong thư viện số, nói chung ít bị kiểm soát hơn truy cập sưu tập in ấn trong thư viện thường. Đưa thông tin vào thư viện số là có khả năng làm cho thông tin đó trở nên phổ biến ngay đối với một số lượng độc giả hầu như vô hạn.

Sở hữu một cuốn sách chắc chắn không phải là xác lập được quyền sở hữu đối với tài liệu đó theo nghĩa của bản quyền. Mặc dù có nhiều bản của một tài liệu nhưng chỉ có một bản quyền. Điều này không chỉ áp dụng cho bản in mà cả cho bản điện tử, dù được số hoá từ bản in hay được tạo nên dưới dạng điện tử từ đầu. Khi mua một cuốn sách, ta có thể bán lại, nhưng

chắc chắn không mua quyền tái phân phối. Quyền đó tùy thuộc vào bản quyền. Ai làm chủ một tác phẩm cụ thể? Bản quyền đầu tiên là của người sáng tác trừ phi tác phẩm được thuê sáng tác. Trong trường hợp này bản quyền thuộc về cơ quan hay tổ chức thuê theo hợp đồng; bản quyền có thể được sang nhượng hay chuyển cho một đơn vị khác thông qua một hợp đồng cụ thể, được thực hiện bằng văn bản do người chủ ký tên.

Luật bản quyền là phức tạp. Tình trạng luật pháp đối với tập tin máy tính và tài liệu cụ thể được xuất bản trên World Wide Web lại mù mờ. Muốn xây dựng thư viện số thì phải cần số hoá tài liệu. Chúng ta phải làm như thế nào để tránh vi phạm bản quyền?

Trước hết chúng ta phải xem xét:

- Nếu tác phẩm được số hoá ở trong miền (domain) công cộng thì chúng ta không phải xin phép. Dĩ nhiên kết quả số hoá của chúng ta cũng không được bảo vệ bản quyền, trừ phi kết quả của ta nhiều hơn bản gốc;
- Nếu tài liệu được tặng cho cơ sở của ta để số hoá và người tặng có bản quyền, thì chúng ta tiến hành số hoá, tuy nhiên cần phải yêu cầu người tặng cung cấp cho mình quyền được số hoá - có thể bằng một mẫu giấy có ghi "quyền sử dụng tác phẩm với bất kỳ mục đích chung của cơ sở, dưới bất kỳ phương tiện nào".

Nếu ta muốn số hoá tài liệu mà không rơi vào hai trường hợp trên thì ta phải cân nhắc việc số hoá của chúng ta có phải là một việc làm có lợi ích chung mà không xâm phạm quyền lợi của người khác hay không. Đây là một điều khó về mặt pháp lý. Cuối cùng nếu chúng ta không chắc chắn với điều cân nhắc trên thì ta phải tiến hành xin phép để được cấp phép thực hiện số hoá.

Vấn đề áp dụng bản quyền còn tùy thuộc quan niệm, chính sách của từng quốc gia. Trong khi vấn đề bản quyền còn đang gây trở ngại cho việc lưu thông sách điện tử ở Châu Âu, thì nhiều liên minh thư viện ở Mỹ lại đang hợp tác lại với nhau để giải quyết một rào cản bản quyền khác –

đó là không có khả năng triển khai dịch vụ mượn liên thư viện đối với sách điện tử. Chiến dịch “Quyền đọc điện tử” (The Right to e-Read) do Phòng Hiệp hội Thông tin và Tài liệu Thư viện Châu Âu (European Bureau of Library Information and Documentation Association-EBLIDA) khởi xướng là một sáng kiến nhằm nâng cao mức độ hiểu biết đối với những bên liên quan như nhà xuất bản, thư viện, bạn đọc/người dùng tin, cũng như người xây dựng chính sách về thách thức mà thư viện gặp phải khi cung cấp sách điện tử và tài liệu đọc số bởi những quy định về bản quyền. Mặc dù vậy, nguyên tắc “First Sales Doctrine” trong luật bản quyền Mỹ đã được thiết lập từ lâu lại cho phép mọi người có thể có quyền bán, cho mượn, hoặc tùy ý sử dụng một ấn phẩm có bản quyền mà không cần sự cho phép của người nắm giữ bản quyền [37].

Tóm lại để tiến hành xây dựng thư viện số, ta phải lưu ý đến vấn đề bản quyền. Những người thực hiện phải cam kết hiểu biết đầy đủ về bản quyền và nhận thức rằng giấy phép là rất cần thiết để chuyển đổi tài liệu số hóa không thuộc trong miền công cộng.

CÂU HỎI ÔN TẬP CHƯƠNG 5

1. Sở hữu trí tuệ được hiểu như thế nào?
2. Phân tích nội dung của vấn đề bản quyền!

CHƯƠNG 6: XÁC ĐỊNH CHÍNH SÁCH VÀ XÂY DỰNG KẾ HOẠCH PHÁT TRIỂN THƯ VIỆN SỐ

Mục tiêu đào tạo: Hướng dẫn người học phương pháp, nguyên tắc xác định chính sách và kế hoạch phát triển thư viện số, trong đó tập trung vào hai nội dung chính là kế hoạch phát triển nguồn tin và kế hoạch tài chính.

6.1 XÁC ĐỊNH CHÍNH SÁCH PHÁT TRIỂN THƯ VIỆN SỐ

Phát triển từ thư viện truyền thống thành thư viện số (TVS) đang là xu hướng tất yếu ở tất cả các nước. Để xây dựng được một TVS theo đúng nghĩa, cần có một số quan điểm thống nhất có cách tiếp cận đúng và lựa chọn những bước đi thích hợp.

Để xây dựng TVS, chúng ta cần quan tâm nhiều vấn đề mà nổi bật là 4 khía cạnh chủ yếu:

- Cấu trúc của TVS;
- Hạ tầng cơ sở kỹ thuật;
- Kho tư liệu số hóa;
- Các vấn đề bảo quản, khai thác và bản quyền.

6.1.1 Cấu trúc của thư viện số

Các TVS đều được bố trí trên “Giao diện web”, trên đó, ngoài những vùng chung như: Giới thiệu về cơ quan, về hệ thống, về thư viện; hướng dẫn sử dụng và các công cụ trợ giúp thì phần chủ yếu là nội dung, tức “Tài nguyên thông tin”

- Phần thứ nhất của “Tài nguyên thông tin” thông thường là danh mục chủ đề. Phần này được cấu trúc theo trình tự từ chung đến riêng, từ tổng quát đến chi tiết, từ ngoài vào trong theo thứ bậc. Cách tổ chức như vậy nhằm tạo thuận tiện cho người dùng trong khai thác thông tin. Thông thường trong TVS các xuất bản phẩm điện tử (tạp chí, bản tin, kỷ yếu ...)

được bố trí sắp xếp theo kiểu này và để tìm kiếm tài liệu theo chủ đề ta có thể vào mục tương ứng tuần tự từ ngoài vào trong. Nếu muốn tìm/truy cập nhanh tới tài liệu cần phải có sự hỗ trợ của công cụ tìm kiếm thông qua các lệnh tìm cụ thể ...

- Phần thứ hai là các tổ hợp CSDL, biểu hiện danh mục các CSDL, thường được sắp xếp theo chủ đề hoặc theo vần chữ cái. Người dùng có thể tiếp cận tới các CSDL này để khai thác thông tin theo các cấp độ khác nhau: từ thư mục tới toàn văn; khai thác riêng rẽ từng CSDL hay khai thác theo nhóm CSDL ... Mức độ khai thác đến đâu tùy thuộc vào khả năng của hệ thống và đặc biệt là sự cho phép của cơ quan chủ quản, các lệ phí tương ứng.

- Phần thứ ba là phần liên kết tới các nguồn tài nguyên thông tin bên ngoài. Đây là thế mạnh của TVS. Tuy nhiên, mức độ và khả năng liên kết đến đâu phục thuộc vào sự hợp tác với các cơ quan khác và việc khai thác các tầng thông tin số hóa đó cũng có những khác biệt: có vùng thông tin khai thác tự do, miễn phí; nhưng cũng có những vùng phải có mật khẩu, phải trả tiền ...

Như vậy, TVS không chỉ có một hệ thống mà có thể gồm nhiều hệ khác nhau. Tuy nhiên, các tài nguyên thông tin, các CSDL đó liên kết được với nhau trong một chế độ phục vụ thống nhất. Tức là khi được yêu cầu, chúng xuất hiện đối với người dùng như thể chúng cùng trong một hệ thống.

6.1.2. Hạ tầng cơ sở kỹ thuật

Một TVS phải có hạ tầng cơ sở đủ mạnh đó là:

- Mạng Intranet có tốc độ kết nối nhanh với Internet
- Hệ thống máy chủ lớn thực hiện việc quản trị các dịch vụ khác nhau: Máy chủ web, máy chủ FTP, Mail, các máy chủ lưu trữ dữ liệu, máy chủ Firewall, máy chủ cho các ứng dụng khác ...

- Hệ thống máy trạm để cập nhật, khai thác thông tin
- Các thiết bị công nghệ chuyên dụng cho TVS: mã vạch, thẻ từ, RFID, máy quét, máy sao dữ liệu ...
- Phần mềm phục vụ cho việc xây dựng và phát triển TVS: Phần mềm TVS, phần mềm hệ thống, hệ điều hành, hệ quản trị CSDL, phần mềm xuất bản điện tử, xuất bản CD-ROM ...

6.1.3. Kho tư liệu số hóa

(Tham khảo mục 2 chương 6)

6.1.4. Các vấn đề bảo quản, khai thác và bản quyền

(Tham khảo mục 2 chương 5)

6.2 XÂY DỰNG KẾ HOẠCH PHÁT TRIỂN NGUỒN TIN SỐ

Một phần quan trọng trong TVS chính là kho tư liệu số hóa của bản thân cơ quan thông tin/ thư viện chủ quản. Có 3 cách để tạo lập Kho này, đó là:

- Tự tiến hành số hóa nguồn tư liệu trên giấy của thư viện. Tức là chuyển tài liệu hiện có sang dạng số bằng phương pháp quét hay nhập lại thông tin từ bàn phím ... Đây là hướng phải đầu tư lớn, đầu tư liên tục và tốn kém thời gian, tiền của, công sức
- Bổ sung/ tích hợp nguồn tin điện tử thông qua việc mua, trao đổi tài liệu điện tử đang được xuất bản (bản tin, tạp chí điện tử, các chế bản điện tử trước khi in ra trên giấy). Chúng ta đều biết: hầu hết các ấn phẩm hiện nay đều vừa xuất bản trên giấy vừa tồn tại dưới dạng điện tử và nếu tận dụng được nguồn này, ta sẽ tiết kiệm được nhiều công sức, thời gian
- Xây dựng các liên kết (tạo khả năng truy cập) đến các nguồn tài liệu trên Internet, nhất là nguồn của các cơ quan có cùng diện chuyên đề bao quát

Tạo lập và phát triển kho tư liệu số của riêng mỗi cơ quan thông tin/thư viện là vấn đề lớn nhất trong xây dựng TVS. Công việc này đòi hỏi

phải đầu tư lớn và liên tục. Để làm tốt công việc này, các cơ quan xây dựng TVS cần có cách tiếp cận hợp lý, khả thi và kinh tế. Cụ thể là:

- Nhất thiết phải lập kế hoạch sát sao và ưu tiên đầu tư cho việc thu thập, xử lý và số hóa nguồn tin cơ bản, nguồn tin tiềm năng của riêng mình. Coi đây là nhiệm vụ trọng tâm, thậm chí phải triển khai mạnh trước khi bắt tay vào xây dựng TVS.

- Nếu không có sự đi trước này, khi ta xây dựng xong hạ tầng mạng và có các phần mềm hệ thống, phần mềm thư viện số đầy đủ nhưng đến lúc đó cơ quan vẫn không có hoặc có rất ít tài liệu số hóa của bản thân chắc chắn TVS đó không thể phát huy được hiệu quả; và như vậy không tương xứng với kinh phí đầu tư nhiều tỷ đồng để xây dựng TVS.

- Trong việc lựa chọn tài liệu để số hóa, ta phải ưu tiên các tài liệu đặc thù của thư viện, các tài liệu duy nhất và có giá trị lâu dài để trao đổi, ví dụ: các tài liệu quý hiếm, các sưu tập có giá trị (không ở đâu có) ...; ưu tiên số hóa trước hết đối với tài liệu chưa ở đâu số hóa, tài liệu tiếng hiếm, tiếng Việt ... Song song với việc số hóa là việc xây dựng các siêu dữ liệu đối với từng tài liệu và cập nhật tài liệu đã được số hóa này vào CSDL tương ứng để phục vụ kịp thời cũng như làm cơ sở cho việc xây dựng TVS sau này.

- Ngoài ra, ta cũng cần quan tâm đến chất lượng của việc số hóa tài liệu cũng như cần sao lưu đầy đủ, kịp thời các tài liệu số hóa để tránh rủi ro cũng như tránh phải làm đi làm lại (lãng phí công sức, tiền của). Điều này phụ thuộc nhiều vào công tác tổ chức, phụ thuộc vào cán bộ thực hiện cũng như thiết bị và quy trình số hóa.

- Phải phối hợp, tận dụng sản phẩm số hóa của các cơ quan thông tin/thư viện khác, nhất là của những cơ quan có cùng diện chuyên đề bao quát. Trên cơ sở đó mới có thể tăng nhanh “nguồn tin” của mình, tiết kiệm được thời gian, công sức và tiền của.

Các thư viện cần có chiến lược đầu tư và triển khai kế hoạch tổng thể

trong việc số hóa đối với các nguồn tin quan trọng, có ý nghĩa lâu dài. Các thư viện cần:

Xác lập chính sách ưu tiên số hóa

1. Về dạng tài liệu

- Các đề tài nghiên cứu khoa học
- Luận văn thạc sỹ, luận án tiến sỹ
- Tài liệu hội nghị, hội thảo khoa học
- Các giáo trình, bài giảng
- Sách, tài liệu tham khảo quý hiếm

2. Về ngôn ngữ

- Tài liệu Hán nôm
- Tài liệu tiếng Việt, hạn chế số hóa tài liệu tiếng Anh

3. Về lĩnh vực

- Ưu tiên số hóa tài liệu phục vụ các ngành, chuyên ngành đào tạo tín chỉ, đào tạo đẳng cấp quốc tế, tài liệu KH-CN những ngành mũi nhọn

Xác lập và hình thành một tổ chức số hóa tài liệu

- Trong mạng lưới các cơ quan thông tin, thư viện cần có sự phân công, phối hợp giữa các cơ quan trong việc số hóa tài liệu, cơ quan nào có kho tư liệu chuyên môn hóa với số lượng tài liệu nhiều nhất về lĩnh vực nào thì sẽ đảm nhận số hóa nguồn tư liệu về lĩnh vực đó, các cơ quan khác hỗ trợ trong việc số hóa

- Bản thân trong các thư viện nên tổ chức một bộ phận chuyên trách cho việc thu thập, số hóa, xử lý các nguồn tin điện tử

Nghiên cứu, lựa chọn và xác lập những chuẩn, quy định chung trong việc số hóa tài liệu

- Các cơ quan thông tin thư viện tiến hành số hóa cần có những chuẩn cũng như các quy định thống nhất đối với việc số hóa tài liệu. Những chuẩn này được xác định trên cơ sở nghiên cứu, áp dụng chuẩn của nước ngoài vào điều kiện cụ thể của Việt Nam. Mỗi cơ quan tham

gia mạng lưới số hóa đều phải tuân thủ nghiêm ngặt những chuẩn này sao cho các tài liệu số hóa luôn đảm bảo về mặt chất lượng cũng như về tiến độ thời gian; đồng thời chúng được tổ chức trong các CSDL có cấu trúc tương hợp hoặc dễ dàng trong chuyển đổi.

- Có như vậy các tài liệu số hóa đó mới có thể phục vụ rộng rãi trong toàn mạng lưới và mang lại hiệu quả cao.

Tổ chức số hóa tài liệu ở quy mô công nghiệp

- Hoạt động số hóa ở quy mô lớn được gọi là “Công nghiệp nội dung” (CNND). Người ta coi CNND bao gồm: thu thập thông tin; phân tích nội dung để phân loại, lưu trữ; bao gói thông tin thành các CSDL (trên CD, DVD, thiết bị lưu trữ ...), nhân sao và cung cấp/bán các CSDL đó.

- Nguyên vật liệu đầu vào của ngành CNND là thông tin và đầu ra cũng là thông tin. Nhưng thông tin đầu ra là CSDL, là thông tin có cấu trúc, có nội dung cụ thể và được bán/cung cấp cho những đối tượng cụ thể nhằm phục vụ cho một/một số hoạt động nhất định

- Để xây dựng được một nền CNND cần phải:

+ Có một hành lang pháp lý đủ mạnh

+ Có sự phối hợp chặt chẽ về mặt tổ chức, liên kết, phân công phân mảng giữa các cơ quan có hoạt động số hóa tài liệu

+ Có lực lượng cán bộ chuyên trách thu thập, số hóa, bao gói thông tin

+ Trang bị các thiết bị hiện đại, ví dụ các máy quét chuyên dụng (nhanh, chất lượng cao, quét được các khổ lớn, màu sắc đẹp ...), các máy chủ lưu trữ và sao lưu chuyên dụng với dung lượng lớn; các thiết bị sao chuyên dụng

6.3 XÂY DỰNG KẾ HOẠCH TÀI CHÍNH

Việc xây dựng và phát triển TVS là một việc làm lâu dài và tốn kém, cần có sự tập trung đầu tư, sự quyết tâm của các cấp lãnh đạo và có một kế hoạch triển khai cụ thể, đặc biệt là một kế hoạch tài chính đảm bảo.

TVS không hoàn toàn tách rời thư viện truyền thống, vì vậy việc xây dựng TVS bao gồm các bước sau:

- Xây dựng hạ tầng cơ sở đủ mạnh bao gồm: hạ tầng cơ sở vật chất, hạ tầng kỹ thuật (các trang thiết bị kỹ thuật, CNTT và phần mềm)
- Xây dựng Kho tài nguyên thông tin (số hóa)
- Xây dựng và phát triển đội ngũ nhân sự để triển khai, vận hành hệ thống

Tùy theo quy mô của từng thư viện và các nhu cầu của từng giai đoạn cụ thể để xây dựng các kế hoạch tài chính cho phù hợp

CÂU HỎI ÔN TẬP CHƯƠNG 6

1. Phân tích nội dung của chính sách phát triển thư viện số!
2. Phân tích kế hoạch phát triển nguồn tin số và kế hoạch tài chính!

CHƯƠNG 7: CÔNG NGHỆ NỘI DUNG VÀ CÔNG NGHIỆP NỘI DUNG

Mục tiêu đào tạo: Giúp người học nắm được các khái niệm, đặc điểm, bản chất của công nghệ nội dung và công nghiệp nội dung. Những ứng dụng rộng rãi, đa dạng và quá trình phát triển của công nghệ nội dung và công nghiệp nội dung trong đời sống xã hội ở Việt Nam và nước ngoài.

7.1 KHÁI NIỆM

7.1. Khái niệm nội dung

Có nhiều cách gọi từ thuật ngữ “content” tiếng Anh trong lĩnh vực khoa học thông tin - thư viện như: nội dung, nội dung số hoặc nội dung thông tin số.

Nội dung (Content): “Thông tin được cung cấp trực tuyến, bao gồm âm nhạc, video, trò chơi, sách, và những xuất bản phẩm khác, tin tức, nghệ thuật, và những thông tin khác. Thông thường có thể phân biệt nội dung (media) chuyên nghiệp (nội dung trực tuyến do các công ty cung cấp) và nội dung do người dùng tạo ra (user-generated content); nội dung trực tuyến được cung cấp bởi cá nhân thông qua trang web cá nhân, mạng xã hội, nhận xét của người dùng hoặc những đăng tải trên blog. [39]

Theo Từ điển tiếng Việt: Nội dung là “mặt bên trong của sự vật, cái được hình thức chứa đựng hoặc biểu hiện. *Nội dung của tác phẩm*”. [40]

Theo Từ điển Khoa học Thông tin - Thư viện (Dictionary for Library and Information Science): Nội dung là chủ đề quan trọng hoặc thực chất của một tác phẩm viết hoặc bài viết, trái ngược với hình thức hoặc phong cách của tác phẩm. Theo nghĩa rộng, “nội dung” là mọi ý tưởng, chủ đề, dữ kiện, hoặc tuyên bố chứa trong tài liệu hoặc tác phẩm viết; đồng nghĩa với chủ đề. [20]

Nội dung (*trong ngữ cảnh Công nghiệp nội dung*) là mọi loại dịch vụ thông tin, giải trí, truyền thông và giao dịch được cung cấp cho nhóm người dùng tin xác định, không phân biệt dịch vụ được tạo ra để bán (dịch vụ thương mại) hoặc được cấp miễn phí (thí dụ thông tin quảng cáo và xuất bản phẩm của công ty).[14]

Khái niệm "*nội dung*" được dùng để chỉ tất cả các dịch vụ thông tin, giải trí, truyền thông và giao dịch có kết hợp các tài liệu, các đồ hoạ hoặc mô phỏng thông minh, hình ảnh động và hình ảnh tĩnh và được tạo ra để đáp ứng nhu cầu của nhóm người dùng xác định, bất luận chúng được tạo ra để bán (dịch vụ thương mại) hoặc để cung cấp miễn phí (ví dụ, các xuất bản và quảng cáo của công ty).

Mặc dù một định nghĩa chung cho các ngành công nghiệp nội dung số 'không tồn tại, nó thường được mô tả là 'nơi nghệ thuật, kinh doanh và công nghệ hội tụ', song, theo một báo cáo của DCITA 'Nội dung kỹ thuật số' được mô tả như sau: “Một tập hợp các kết quả đầu ra của sản phẩm và dịch vụ trong các ngành công nghiệp sáng tạo” (các ngành công nghiệp sáng tạo sản xuất nội dung số) [41]

7.1.2. Khái niệm công nghệ nội dung

Theo Luật KH&CN 2013, *Công nghệ* là giải pháp, quy trình, bí quyết kỹ thuật có kèm theo hoặc không kèm theo công cụ, phương tiện dùng để biến đổi nguồn lực thành sản phẩm.

Công nghệ nội dung là công nghệ liên quan đến sản xuất, tạo lập, xử lý, lưu giữ, truyền dẫn, phổ biến nội dung.

Trong khi bắt đầu chủ yếu là thông tin định hướng văn bản, cho đến nay “nội dung điện tử” đã đi một chặng đường dài với sự phát triển, đổi mới nhanh chóng của công nghệ thông tin và truyền thông. Trên Internet nó đã thay đổi khá nhanh chóng đối với nội dung đa phương tiện, trình

bày lại văn bản, đồ họa, âm thanh và video cũng như sự kết hợp những phương tiện này.

Công nghệ nội dung không chỉ thực hiện trên máy tính mà nó triển khai trên tất cả các thiết bị nối mạng kỹ thuật số, như điện thoại di động và các thiết bị kỹ thuật số cá nhân (PDA). Công nghệ nội dung đưa lại những thay đổi trong sản xuất và phổ biến của hoạt động kinh doanh, quá trình sáng tạo và phương thức trình bày.

Công nghệ nội dung là công nghệ thể hiện nội dung bằng thông tin kỹ thuật số được cung cấp qua các thiết bị điện tử trên mạng. Tức là các biểu tượng có thể được sử dụng và diễn giải bởi các tác nhân con người trong quá trình truyền thông

7.1.3. Khái niệm công nghiệp nội dung

Công nghiệp theo nghĩa rộng của kinh tế học là hoạt động kinh tế quy mô lớn, sản phẩm (có thể là phi vật thể) tạo ra trở thành hàng hóa. Như vậy công nghiệp nghĩa chung bao gồm cả dịch vụ. Công nghiệp nghĩa hẹp là lĩnh vực sản xuất hàng hóa vật chất mà sản phẩm được "chế tạo, chế biến" cho nhu cầu tiêu dùng hoặc phục vụ hoạt động kinh doanh tiếp theo. Đây là hoạt động kinh tế, sản xuất quy mô lớn, được sự hỗ trợ thúc đẩy mạnh mẽ của các tiến bộ công nghệ, khoa học và kỹ thuật.

Công nghiệp nội dung (còn gọi là công nghiệp nội dung kỹ thuật số) là một khái niệm rất mới và trên thế giới, người ta cũng chưa tìm ra được một định nghĩa thống nhất. Mỗi quốc gia có cách hiểu về lĩnh vực này khác nhau. Có thể đưa ra một khái niệm đơn giản là nội dung cộng với công nghệ số bằng công nghiệp nội dung số. (nội dung + công nghệ số = công nghiệp nội dung số).

Ở VN, dự thảo luật công nghệ thông tin từng đưa ra khái niệm công nghiệp nội dung số nhưng không nói rõ nó gồm những lĩnh vực gì. Tuy nhiên, quan điểm của Bộ Bưu chính Viễn thông (nay là Bộ Thông tin &

Truyền thông) thì cho rằng đây là ngành thiết kế, sản xuất, xuất bản, lưu trữ, phân phối, phát hành các sản phẩm nội dung số và dịch vụ liên quan. Nó bao gồm nhiều lĩnh vực như: tra cứu thông tin, dữ liệu số, giải trí số, nội dung giáo dục trực tuyến, học tập điện tử, thư viện và bảo tàng số, phát triển nội dung cho mạng băng rộng, mạng di động 3G....

Công nghiệp nội dung là một thành phần của công nghiệp thông tin (Information Industry). Ngoài công nghiệp nội dung (thông tin), công nghiệp thông tin còn công nghiệp phân phối/phổ biến thông tin và công nghiệp xử lý thông tin [37]

Công nghiệp nội dung bao gồm mọi lĩnh vực nhằm tạo ra giá trị thông qua tạo lập hoặc phổ biến văn bản số hoá hoặc tương tự, hình ảnh, video, và/hoặc âm thanh dựa trên dịch vụ cho nhóm người dùng xác định.

Công nghiệp nội dung gồm những công ty media tập trung vào tạo ra và truyền đạt nội dung, những nhà cung cấp phần mềm và phần cứng, những nhà vận hành hệ thống cáp truyền đưa nội dung đến người dùng đầu cuối hoặc cung cấp những thiết bị để người dùng đầu cuối truy cập và sử dụng những dịch vụ được cung cấp.[13]

Công nghiệp nội dung theo OECD phải đáp ứng tiêu chí là:

“Việc sản xuất (hàng hoá và dịch vụ) của lĩnh vực công nghiệp phải chủ yếu là nhằm thông tin, giáo dục hoặc giải trí con người thông qua các phương tiện truyền thông. Những công nghiệp này bao gồm sản xuất, xuất bản và/hoặc phân phối nội dung (những sản phẩm giải trí, văn hoá và thông tin) trong đó nội dung là thông báo cáo tổ chức dành cho con người”. [1]

7.2 CÔNG NGHỆ NỘI DUNG

7.2.1. Số hóa tài liệu

(Tham khảo mục 3.3 chương 3)

7.2.2. Xuất bản điện tử

Xuất bản điện tử là việc ứng dụng công nghệ thông tin, viễn thông trong hoạt động xuất bản. Kết quả là xuất bản điện tử tạo ra các xuất bản phẩm điện tử. Theo từ điển Bách Khoa toàn thư mở, xuất bản điện tử gồm việc xuất bản dưới dạng số hóa đối với sách, tạp chí, phát triển các thư viện số và các bảng giới thiệu (catalogue). Mặc dù việc phân phối các sản phẩm xuất bản số trên internet ngày nay đã có liên hệ rất chặt chẽ với xuất bản điện tử, nhưng vẫn còn có hoạt động xuất bản điện tử không liên quan đến mạng máy tính như các CD-ROM Bách khoa toàn thư, các sách tra cứu trên CD-ROM... Như vậy xuất bản điện tử hiểu một cách đơn giản là việc sản xuất và phát hành các xuất bản phẩm dưới dạng số bằng các phương tiện điện tử. Sản phẩm của xuất bản điện tử là sách điện tử, báo điện tử và các tài liệu dạng số khác. Việc sử dụng các sản phẩm của hoạt động xuất bản điện tử cũng phải bằng các phương tiện điện tử.

Đặc điểm của xuất bản điện tử: Vì là hoạt động xuất bản nên xuất bản điện tử cũng mang những đặc điểm của xuất bản truyền thống nhất là khâu tổ chức sản xuất. Xuất bản điện tử cũng bao gồm các qui trình, công đoạn để biến một ý tưởng, một kết quả hoạt động trong lĩnh vực văn học, nghệ thuật, khoa học, kỹ thuật công nghệ... được thể hiện dưới dạng sản phẩm hoàn chỉnh tuân thủ luật pháp và các qui định của các cơ quan quản lý có thẩm quyền. Các đặc trưng, yêu cầu vốn có của xuất bản truyền thống đều được thể hiện trong xuất bản điện tử. Điểm khác biệt của xuất bản điện tử là sử dụng công nghệ thông tin, viễn thông vào các công đoạn tạo ra sản phẩm, không còn công đoạn in như trong hoạt động xuất bản truyền thống; hình thức tồn tại của sản phẩm thông qua các vật mang thông tin đã được số hóa; là các công nghệ để tạo ra các xuất bản phẩm điện tử; là trình tự thực hiện các công đoạn trong quá trình xuất bản; sử dụng các công nghệ hiện có cho phép các nhà xuất bản chuyển giao

nội dung đến độc giả nhanh chóng và tiện lợi...

7.2.3. Hệ quản trị nội dung

Hệ quản trị nội dung (Content Management System - CMS) là chương trình máy tính cho phép xuất bản, biên tập và hiệu đính nội dung cũng như duy trì nội dung từ một giao diện trung tâm. Những hệ thống như vậy cung cấp quy trình để quản quản lý dòng công việc trong môi trường hợp tác

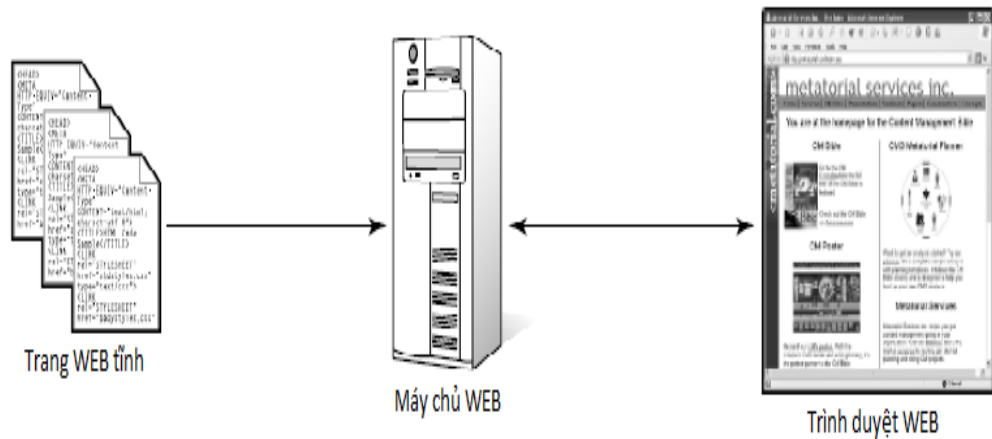
Một số tính năng cơ bản của CMS:

- Tạo, duyệt hoặc thay đổi nội dung trực tuyến
- Chế độ Soạn thảo "Nhìn là biết" WYSIWYG (What You See is What You Get)
- Quản lý người dùng
- Tìm kiếm và lập chỉ mục
- Lưu trữ
- Tùy biến giao diện
- Quản lý ảnh và các liên kết (URL)

Trước kia, để thêm bài viết cho website tĩnh - website html, cần phải có kiến thức nhất định về ngôn ngữ html. Cách thức để phát triển, vận hành website html đơn thuần giống như việc của người in sách. Viết nội dung chỉ là một phần công việc, sau đó người ta phải cần tới giấy và các máy in để in ra trang sách, cuối cùng là đóng các trang sách lại thành một cuốn sách. Các website html cũng bao gồm những trang html nhỏ, và việc tạo ra từng html nhỏ đòi hỏi phải biết về ngôn ngữ html, và phải qua nhiều bước để biến một văn bản words thành một trang html. Công việc quản lý, cập nhật website sẽ khá vất vả và tốn thời gian. Nếu có thay đổi về địa chỉ email, sẽ phải thay đổi trên tất cả các trang html. Nếu muốn có thêm một bài viết mới, phải tạo link tới bài viết, phải viết 1 trang html hoàn chỉnh theo giao diện của website... Khi phát triển

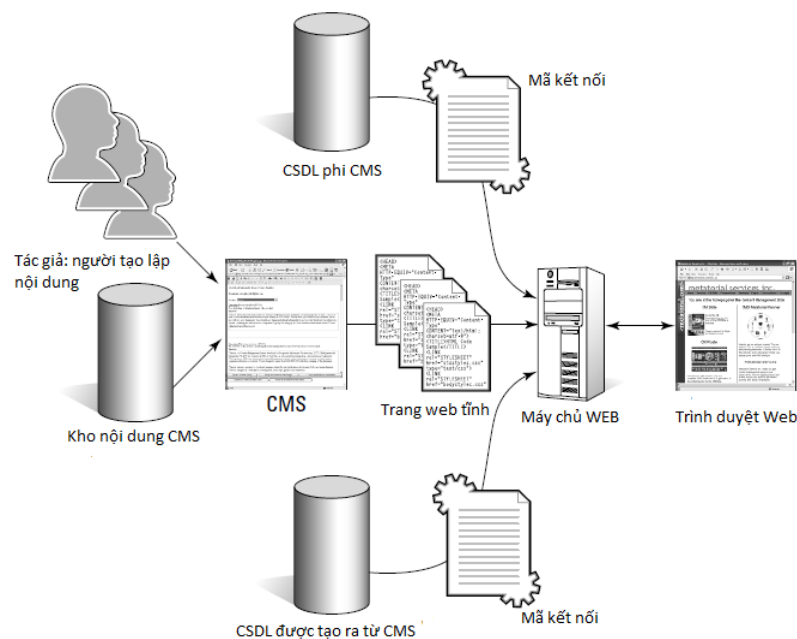
website, điều tất yếu là khối lượng bài viết nhiều lên, hoặc trang web có nhiều phần nội dung, thì việc quản lý theo kiểu từng trang một như vậy sẽ không thể thực hiện được. Và rất may, bây giờ chúng ta đã không còn phải làm như thế nữa, nhờ CMS - hệ quản trị nội dung.

Web tĩnh được tạo ra bằng các trang HTML tĩnh, được tạo ra bằng các phần mềm soạn thảo HTML



Hình 7.1 Mô hình cung cấp nội dung Web tĩnh

Web động được tạo ra từ những trang HTML động, được tạo ra bằng các phần mềm CMS



Hình 7.2 Mô hình cung cấp nội dung Web tĩnh và Web động qua CMS

CMS – Content management (hệ quản trị nội dung) là một phần mềm ứng dụng có chức năng quản trị nội dung cho website. CMS cung cấp những công cụ trực quan nhất, đơn giản nhất để những người dùng phổ thông có thể quản lý dễ dàng các hoạt động của 1 website mà không cần tới những hiểu biết chuyên môn về lập trình.

CMS là bước đi tiên phong và phát triển nhất cho xu hướng của các phần mềm nền tảng web. Bạn sẽ chỉ cần các click chuột, các ô check box, các phần nhập nội dung thông tin... mà hoàn toàn không phải quan tâm tới các dòng lệnh, các logic hệ thống đằng sau. CMS sẽ đảm nhiệm tất cả để cung cấp cho bạn một môi trường làm việc dễ dàng, tiện lợi, và thân thuộc (giống như cách các bạn thao tác với hệ điều hành máy tính)

Các CMS thường được viết với cùng ngôn ngữ lập trình chung cho website, phổ biến nhất hiện nay là PHP và MySQL. Tuy nhiên, trên lý thuyết một CMS có thể viết bởi bất cứ ngôn ngữ lập trình và quản lý dữ liệu nào “trên nền tảng web” (có thể chạy trên môi trường web).

Những yêu cầu cho CMS – Hệ quản trị nội dung

Giấu đi đằng sau sự phức tạp, cung cấp cho người dùng phương thức sử dụng đơn giản nhất là tiêu chí phát triển CMS. Tuy nhiên, chính vì mục đích sử dụng của chúng ta khác nhau, nên cũng có rất nhiều các CMS khác nhau, phù hợp cho từng mục đích riêng. Ở đây, chúng ta hãy tiếp cận với những tiêu chí đánh giá 1 CMS tốt nói chung. Trên thực tế, tùy vào từng dự án, một hoặc một vài thành phần có thể được loại bỏ, rút gọn để phù hợp hơn với mục đích sử dụng.

- Tính liên tục của thông tin:

Mặc dù môi trường Internet có những giới hạn riêng, các nhà lập trình luôn tìm được cách để vượt qua những giới hạn đó, duy trì kết nối với người sử dụng thông qua nhiều hình thức. Có thể qua những bài viết, những messages, những tấm ảnh... mà chúng ta có thể gộp chung là “thông tin”. CMS cần đảm bảo những thông tin đó được truyền tải nhanh

nhất, chính xác nhất, an toàn nhất. Đặc biệt là những thông tin đó phải luôn sẵn sàng phục vụ, bởi môi trường Internet là môi trường luôn luôn kết nối, hoàn toàn không có giờ nghỉ

- *Khả năng quản lý người dùng*

Một website càng phát triển hay càng phức tạp, số lượng người dùng (user) càng nhiều. Và người quản trị website sẽ cần những cơ chế có thể phân nhóm, quản lý những nhóm người dùng khác nhau, nhằm cung cấp những thông tin, chức năng phù hợp nhất tới từng nhóm người dùng đó. Vì vậy, CMS phải có 1 hệ thống xác thực và quản lý người dùng có khả năng quản lý từ một tới vài nghìn (hay vài trăm nghìn – với những website phát triển) người dùng (users).

- *Khả năng quản lý truy cập*

Gắn liền với việc phân loại người dùng, là quy trình giới hạn truy cập tài nguyên với từng nhóm người dùng khác nhau. Ví dụ đơn giản nhất bạn có thể thấy là người dùng bình thường sẽ không có khả năng viết bài, hoặc những người chuyên viết bài không có khả năng thay đổi hiển thị của website... Không như trước kia, khi website còn đơn giản và chỉ cần 1 người quản trị. Ngày nay, các website cung cấp lượng thông tin khổng lồ, như một thư viện và cần một tập thể để duy trì thư viện ấy (những người thủ thư). Khả năng quản lý truy cập mà chúng ta nói tới ở đây chính là giúp cho việc quản lý tập thể đó, đảm bảo cho bộ máy bên trong của website của chúng ta hoạt động ăn khớp với nhau, không có trục trặc.

- *Khả năng quản lý mở rộng chức năng*

Thế giới Internet luôn thay đổi và website của bạn cũng cần thay đổi phù hợp với môi trường Internet. Do vậy, khả năng mở rộng chức năng là điều mà chúng ta luôn nghĩ tới khi thiết kế website. CMS là nền móng để phát triển website của bạn, do đó, nền móng này phải thích ứng tốt với các thay đổi, các điều chỉnh hoặc mở rộng có thể sẽ có trong tương

lai. Chúng ta đang nhắc tới một cơ chế cho phép người quản trị (dù không am hiểu về lập trình) cũng có khả năng quản lý, điều chỉnh các chức năng mở rộng của website. Một CMS càng phổ biến khi người quản trị càng có nhiều khả năng mở rộng website của mình một cách đơn giản, tiện dụng nhất .

- *Bảo mật và báo lỗi*

Bảo mật là yêu cầu cơ bản nhất đối với website. Do đó, nó cũng là vấn đề cơ bản nhất mà CMS phải giải quyết ngay từ khi bắt đầu dự án. Ngoài ra, các trường hợp bị tấn công, spam hay các lỗi phát sinh trong quá trình sử dụng cũng cần được quản lý, thông báo tới người quản trị để có những hành động xử lý tương thích.

- *Các chức năng hệ thống nâng cao khác*

Ngoài các chức năng cơ bản trên, CMS có thể có các chức năng hệ thống nâng cao khác. Do tính “nâng cao” của mình, việc tích hợp hay không các chức năng này phụ thuộc vào mục đích, đặc thù của từng dự án. Một vài chức năng nâng cao có thể kể đến như: Khả năng quản lý database tích hợp; Khả năng quản lý và chỉnh sửa giao diện tích hợp; Khả năng triển khai đa ngôn ngữ...

Với các website phát triển và phức tạp, thì chính những chức năng hệ thống nâng cao khác biệt là điểm mạnh và đem đến thành công cho website. Tuy nhiên, cần cân nhắc kỹ về tính “nâng cao” khi triển khai dự án, để có những chiến lược và kế hoạch cụ thể, do việc phát triển các chức năng này không hề đơn giản.

7.2.4. Công nghệ thư viện số

7.2.4.1. Giới thiệu tổng quan

Kể từ năm 2003, nhiều nhà nghiên cứu đã tham gia vào nghiên cứu việc sử dụng hiệu quả các website thư viện đại học và nghiên cứu nhằm đưa ra cách tốt nhất để tổ chức nội dung cho người dùng tin trong môi

trường số. Dù có tiến hành nghiên cứu nào đi nữa, thì tìm kiếm tài liệu cho khoá học, và xác định tài khoản người dùng cũng như thông tin chung về thư viện vẫn nằm trong số tính năng cần thiết cho bạn đọc. Một cá nhân truy cập thông tin này giờ đây đã thay đổi do sự phổ biến nhanh chóng của công nghệ di động. Thư viện đại học và nghiên cứu đã triển khai nhiều giải pháp để đáp ứng người dùng tin trên di động, bao gồm tạo ra phiên bản website di động, ứng dụng di động, và thiết kế website tự động thích nghi với mọi loại kích cỡ màn hình di động. Tất cả những triển khai này đã tổ chức lại nội dung trên các website nhằm hiển thị tối ưu trên các kích cỡ màn hình nhỏ hơn. Trong khi có sự ưu tiên chuyển giao các dịch vụ di động giành được sự quan tâm nhiều ở các thư viện, thì các thảo luận hiện nay đang đề cập đến làm thế nào để phục vụ bạn đọc hiệu quả nhất trên nhiều loại hình thiết bị di động khác nhau được sử dụng để truy cập nguồn lực thông tin thư viện.

Sự nổi lên của công nghệ di động cũng sẽ làm thay đổi luồng công việc học tập và nghiên cứu, giờ đây sinh viên và nhà nghiên cứu đang ngày càng hứng thú hơn với việc tìm kiếm trong một mục lục thư viện, đọc tóm tắt và toàn văn dưới dạng số ngay lập tức, và sau đó tìm và lưu tài liệu tham khảo thông qua website trên thiết bị di động hoặc ứng dụng di động của thư viện thay vì tới một cơ sở thư viện nào đó để mượn. Điều quan trọng thư viện không chỉ xây dựng các tính năng cho các thông tin cơ bản về thư viện, mà còn ưu tiên những nhiệm vụ giúp cho việc tiến hành nghiên cứu có năng suất cao thông qua thiết bị di động. Tối ưu năng lực tìm kiếm của mục lục thư viện cho thiết bị di động là một trong những tính năng hữu ích nhất, cùng với việc tích hợp các cơ sở dữ liệu trực tuyến sẵn có của thư viện vào cùng một ứng dụng web hoặc ứng dụng di động của thư viện đại học và nghiên cứu sẽ tạo nên một công cụ mạnh cho bạn đọc tìm kiếm sách và bài báo từ một vị trí duy nhất trên thiết bị di động của họ.

Với sự tăng về mức độ truy cập tài liệu số trên thiết bị di động, các thư viện đại học và nghiên cứu sẽ cần tổ chức chuyển giao những nội dung sẵn có trước kia tới máy tính bảng và điện thoại thông minh, và phân phối nhiều loại hình nội dung mới hơn như sách điện tử. Thói quen đọc đang thay đổi với mức tăng sử dụng thông tin trên thiết bị di động ở cả màn hình lớn và nhỏ hơn. Xu hướng này đòi hỏi nội dung cần được định dạng theo hình thức dễ đọc hơn cho người dùng tin. Trong khi tệp tin PDF đã và đang trở thành thói quen chung để truy cập nội dung số kể từ những năm 1990, thì chúng cũng trở nên phù hợp hơn cho máy tính cá nhân, máy tính xách tay và các thiết bị lớn khác. Những giải pháp chuyển giao nội dung trên định dạng EPUB 3 được xem như là một lựa chọn có thể phát triển tốt cho việc đọc trên các thiết bị di động nhỏ. Sinh viên và nhà nghiên cứu cũng mong muốn truy cập nhiều hơn tới tài liệu số, bao gồm sách điện tử, video, và những hình ảnh hoá dữ liệu trên thiết bị di động. Để bắt kịp xu hướng này, thư viện đại học và nghiên cứu đang làm việc với các nhà cung cấp và nhà xuất bản để phân phối sách điện tử, chủ yếu thông qua mô hình truy cập vĩnh viễn và thuê bao hàng năm. Trong khi vấn đề bản quyền còn đang gây trở ngại cho việc lưu thông sách điện tử ở Châu Âu, thì nhiều liên minh thư viện ở Mỹ lại đang hợp tác lại với nhau để giải quyết một rào cản bản quyền khác – đó là không có khả năng triển khai dịch vụ mượn liên thư viện đối với sách điện tử.

7.2.4.2. Những tác động đến chính sách, lãnh đạo và thực hành

Thư viện hiện nay đang bắt đầu phát triển những chính sách cho việc sử dụng thiết bị di động tại cơ sở thư viện của mình. Ví dụ, Hệ thống Thư viện Đại học Penn (Penn State University Libraries), nhận ra tầm quan trọng ngày càng tăng của điện thoại trong đời sống hàng ngày. Thư viện đã dành riêng một khu vực sử dụng điện thoại để tránh những cuộc

điện thoại mất sự yên tĩnh trong môi trường thư viện. Ngoài ra, sự cần thiết phải cập nhật chính sách bản quyền để nhận ra những tài sản số thân thiện với thiết bị di động, ví dụ như sách điện tử, đang tác động đến nhiều thư viện trên khắp thế giới. Chiến dịch “Quyền đọc điện tử” (The Right to e-Read) phát động bởi Phòng Hiệp hội Thông tin và Tài liệu Thư viện Châu Âu (European Bureau of Library Information and Documentation Association-EBLIDA) là một sáng kiến được biết trên khắp Châu Âu nhằm nâng cao mức độ hiểu biết đối với những bên liên quan như nhà xuất bản, thư viện, bạn đọc/người dùng tin, cũng như người xây dựng chính sách về thách thức mà thư viện gặp phải khi cung cấp sách điện tử và tài liệu đọc số bởi những quan ngại về bản quyền. Mặc dù vậy, một nguyên tắc “First Sales Doctrine” trong luật bản quyền Mỹ đã được thiết lập từ lâu cho phép mọi người có thể có quyền bán, cho mượn, hoặc tùy ý sử dụng một ấn phẩm có bản quyền mà không cần sự cho phép của người nắm giữ bản quyền.

Việc tạo thuận lợi cho phong trào nội dung và chuyển giao di động đòi hỏi một tầm nhìn lãnh đạo trong các hiệp hội thư viện, nhà cung cấp công nghệ chuyên sâu trong lĩnh vực, cũng như trong thư viện đại học và nghiên cứu. Ví dụ, một chương trình xây dựng Hạ tầng Di động cho Thư viện của JISC (JISC Mobile Infrastructure for Libraries Programme) đã thiết lập lên một dự án “Mobilising Academic Content Online” để phát triển một bộ công cụ thực hành tốt nhằm giải quyết những thách thức liên quan đến chuyển giao nội dung tài liệu học tập trên thiết bị di động. Tương tự như vậy, một dự án được cấp quỹ gần đây tại Đại học California Davis (University of California Davis) sẽ tạo ra một lộ trình cho các kế hoạch đầu tư có tính chiến lược vào phần mềm, tiêu chuẩn, và chuyên môn mới. Những phát triển chuyên ngành và nhiều tài nguyên thông tin trực tuyến khác hiện cũng sẵn có để giúp thủ thư hiểu biết và phát triển các ứng dụng, cũng như website thư viện phù hợp với thiết bị

di động hơn. Tạp chí Techsource của Hiệp hội Thư viện Mỹ (ALA) hiện cung cấp nhiều thông tin và đào tạo làm thế nào để cải tiến một website di động của thư viện. “23 Mobile Things” là một khoá học trực tuyến từng bước tìm hiểu tiềm năng của các công cụ di động đối với chuyển giao dịch thư viện.

Sự ưu tiên chuyển giao nội dung và dịch vụ thư viện trên thiết bị di động đã được thấy trong nhiều ví dụ thực hành. Tại hệ thống Thư viện Đại học Duke (Duke University Libraries), thư viện đang sử dụng ứng dụng xem tạp chí điện tử cho máy tính bảng để làm cho tài nguyên thông tin của thư viện trở lên thân thiện với thiết bị di động hơn, đồng thời cho phép bạn đọc của thư viện tìm duyệt, đọc, và theo dõi tạp chí khoa học hiện hành dễ dàng. Thư viện Đại học Exeter (University of Exeter) cũng triển khai một dịch vụ nội dung di động bằng việc cung cấp một danh mục tài nguyên thông tin có thể truy cập dễ dàng qua website di động và ứng dụng di động. Khi các site di động và ứng dụng di động phát triển phổ biến, thì các thư viện đại học và nghiên cứu cũng cần cung cấp những giải pháp phù hợp cho mọi thiết bị di động, như thiết kế hỗ trợ hay tương thích mọi màn hình di động. Các website tại thư viện Đại học Toronto (University of Toronto), Đại học Princeton (Princeton University), Đại học ETH (ETH Bibliothek) và Đại học Arizona (University of Arizona) đã đảm bảo rằng việc xem thông tin có thể được tối ưu hoá cho mọi kích cỡ màn hình, thậm chí cả những màn hình vẫn chưa được sáng tạo.[37]

7.3 CÔNG NGHIỆP NỘI DUNG

7.3.1. Đặc điểm chung

Nội dung Công nghiệp kỹ thuật số bao gồm việc sản xuất và tiếp thị các chương trình phim và truyền hình trong hình thức truyền hình kỹ thuật số và tương tác; Trò chơi trực tuyến; tái sử dụng nội dung giáo dục điện tử; tiếp thị và cung cấp các hiện vật/tài liệu của các bảo tàng, phòng trưng bày và các thư viện ở dạng kỹ thuật số; xuất bản âm nhạc, văn bản,

các bộ phim và các trò chơi trên Internet; và sự phát triển và tiếp thị của các phần mềm, trò chơi, và các dịch vụ trực tuyến tạo phương tiện truyền thông kỹ thuật số và hiệu ứng hình ảnh, hoặc hỗ trợ để quản lý và xuất bản chúng.

Theo Tổ chức Hợp tác và Phát triển Kinh tế (OECD) công nghiệp nội dung phải đáp ứng tiêu chí là: “Việc sản xuất (hàng hoá và dịch vụ) của lĩnh vực công nghiệp phải chủ yếu là nhằm thông tin, giáo dục hoặc giải trí con người thông qua các phương tiện truyền thông. Những công nghiệp này bao gồm sản xuất, xuất bản và/hoặc phân phối nội dung (những sản phẩm giải trí, văn hoá và thông tin) trong đó nội dung là thông báo cáo tổ chức dành cho con người”

Dưới đây là thống kê của OECD (dựa trên bảng phân loại Công nghiệp quốc tế ISIC) về các lĩnh vực của công nghiệp nội dung [41]

581	Xuất bản sách, tạp chí và những hoạt động xuất bản khác
5811	Xuất bản sách
5812	Xuất bản những danh bạ và danh mục
5813	Xuất bản báo chí, tạp chí, xuất bản phẩm định kỳ
5819	Các hoạt động xuất bản khác
59	Sản xuất phim điện ảnh, video và chương trình truyền hình, ghi âm và hoạt động xuất bản âm nhạc
591	Các hoạt động phim điện ảnh, video và chương trình truyền hình
5911	Các hoạt động sản xuất phim điện ảnh, video và chương trình truyền hình
5912	Các hoạt động sau sản xuất phim điện ảnh, video và chương trình truyền hình
5913	Các hoạt động phát hành, phổ biến phim điện ảnh, video và chương trình truyền hình
5914	Các hoạt động chiếu phim
592	Các hoạt động xuất bản âm nhạc và ghi âm

5920	Các hoạt động xuất bản âm nhạc và ghi âm
60	Programming and broadcasting activities
601	Phát thanh
6010	Phát thanh
602	Hoạt động phát và tạo lập chương trình truyền hình
6020	Hoạt động phát và tạo lập chương trình truyền hình
639	Các hoạt động dịch vụ thông tin khác
6391	Hoạt động cơ quan thông tấn
6399	Các hoạt động dịch vụ thông tin khác không xếp vào mục nào khác

Bảng 7.1 các lĩnh vực của công nghiệp nội dung (theo OECD)

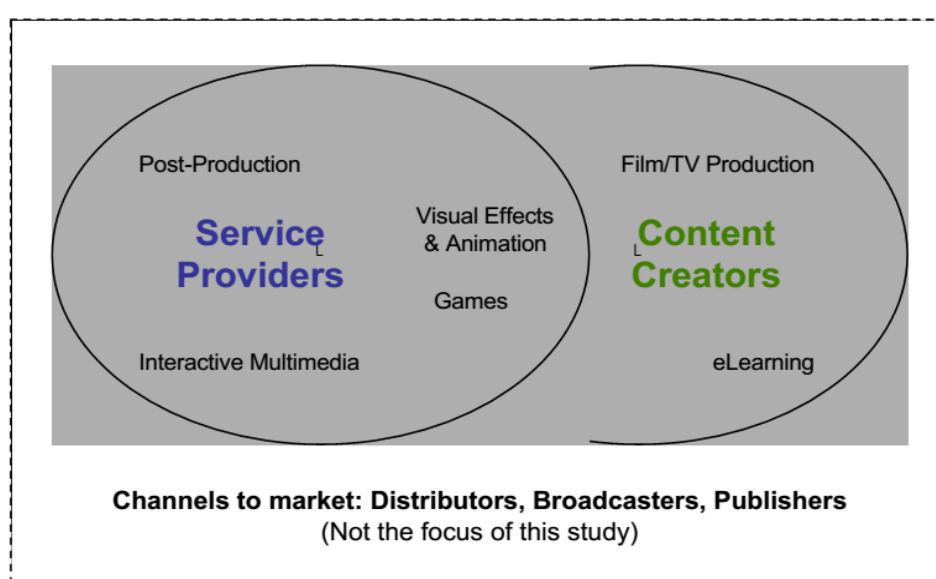
Công nghiệp nội dung số bao gồm các công ty hoạt động trong các lĩnh vực sau:

- Hiệu ứng hình ảnh và hình ảnh động (bao gồm cả thực tế ảo, thiết kế kỹ thuật và các sản phẩm 3D)
- Tương tác đa phương tiện và phát triển phần mềm,
- Máy tính và trò chơi trực tuyến
- Giáo dục đa phương tiện (eLearning)
- Sản xuất phim truyền hình kỹ thuật số

Theo ghi nhận của Trung tâm Kinh tế Quốc tế, sự phân mảnh trong ngành công nghiệp nội dung số là rất cao. Điều này tạo ra một số thách thức, bao gồm cả những khó khăn trong việc đảm bảo tài chính và khám phá thông tin về thị trường. Thị trường phân mảnh cũng có xu hướng tạo ra sự phụ thuộc vào các nhà phân phối nội dung. Ngành công nghiệp nội dung số không hoạt động trong một thị trường duy nhất. Thay vào đó, các lĩnh vực trong ngành công nghiệp này phải đối mặt với điều kiện thị trường khác nhau và đối mặt với những thách thức riêng rẽ. Hai thành phần chính của công nghệ nội dung số là sáng tạo nội dung và cung cấp dịch vụ.

Người sáng tạo nội dung (người sở hữu IP) bao gồm các công ty tạo ra và sở hữu trí tuệ riêng của họ. Các công ty này có xu hướng phụ thuộc nhiều vào hỗ trợ tài chính của chính phủ.

Nhà cung cấp dịch vụ (người không giữ lại quyền IP riêng của họ) bao gồm các công ty có xu hướng làm việc trên cơ sở thu phí dịch vụ và nói chung là không kiếm được doanh thu liên tục từ sở hữu trí tuệ của họ. Trong hầu hết trường hợp, các công ty này ít phụ thuộc vào nguồn tài trợ của chính phủ.



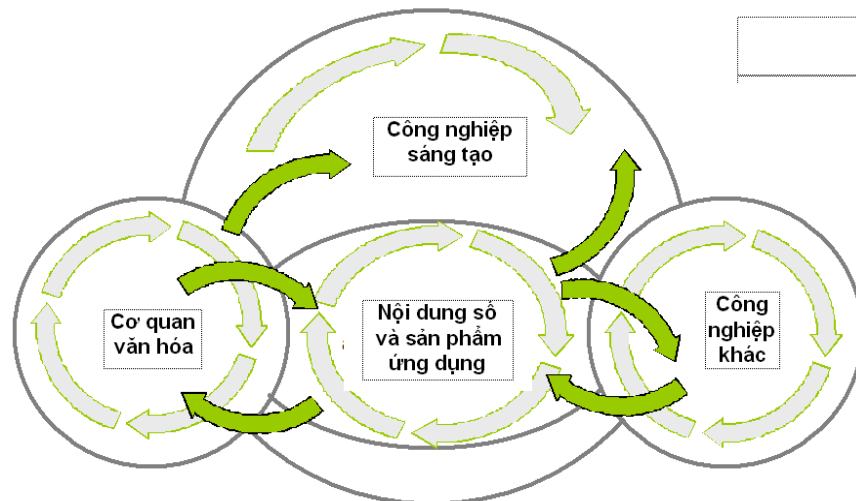
Hình 7.3 Hai thành phần của công nghiệp nội dung số

7.3.2. Công nghiệp nội dung (số) thế giới

7.3.2.1. Công nghiệp nội dung (số) ở Úc

Công nghiệp nội dung số là một đóng góp trực tiếp và gián tiếp quan trọng đến nền kinh tế và xã hội Úc. Nó góp phần trực tiếp thông qua tạo việc làm, vốn thu nhập đầu tư và xuất khẩu. Một cách gián tiếp, nó cung cấp các kỹ năng và khả năng phát triển công nghiệp truyền thống cũng như các ngành công nghiệp mới và đang nổi lên. Các ngành công nghiệp là nền tảng của hệ thống phân phối của các công ty, cung cấp các công nghệ và ứng dụng cần thiết để quản lý hiệu quả nội dung và giao diện người dùng cuối.

Công nghiệp nội dung số mở rộng ra các thành phần ứng dụng và dịch vụ của các ngành công nghiệp công nghệ thông tin như phim truyền thống, giải trí và các ngành công nghiệp văn hóa khác. Đặc biệt là các ngành dịch vụ như y tế, giáo dục và kiến trúc... Sự chùng chéo với các ngành công nghiệp văn hóa đa dạng và phong phú khẳng định công nghiệp nội dung số có ý nghĩa rất lớn đối với nền văn hóa Úc.



Hình 7.4: Liên kết công nghiệp nội dung số ở Úc

Các ngành công nghiệp nội dung số là chiến lược quan trọng trong định hướng phát triển giá trị gia tăng của Úc đối với sản phẩm và dịch vụ. Một vài ngành công nghiệp nội dung số được phát triển tương đối tốt, chẳng hạn như phim ảnh và sản xuất truyền hình, trong khi những lĩnh vực khác còn trong quá trình thăm dò như phát triển trò chơi và đa phương tiện tương tác trên nền tảng mới. Mặc dù ngành công nghiệp này có tầm quan trọng lớn đến vị thế kinh tế hiện tại và tương lai của Úc, song sự năng động và bản chất của nó ít được biết đến. Một phần do thực tế ngành công nghiệp này chuyển đổi rất nhanh chóng và công nghệ cũng như điều kiện thị trường toàn cầu thường xuyên thay đổi.[11]

Công nghiệp nội dung (kỹ thuật) số Úc được ước lượng đã sản xuất ra giá trị 21 tỷ đô la, gần 3,5 % GDP của đất nước. Tác động của công

ngành nội dung số mở rộng trên toàn nền kinh tế với sản xuất cốt lõi của ngành công nghiệp là tạo ra các nội dung kỹ thuật số trong các ngành công nghiệp sáng tạo.

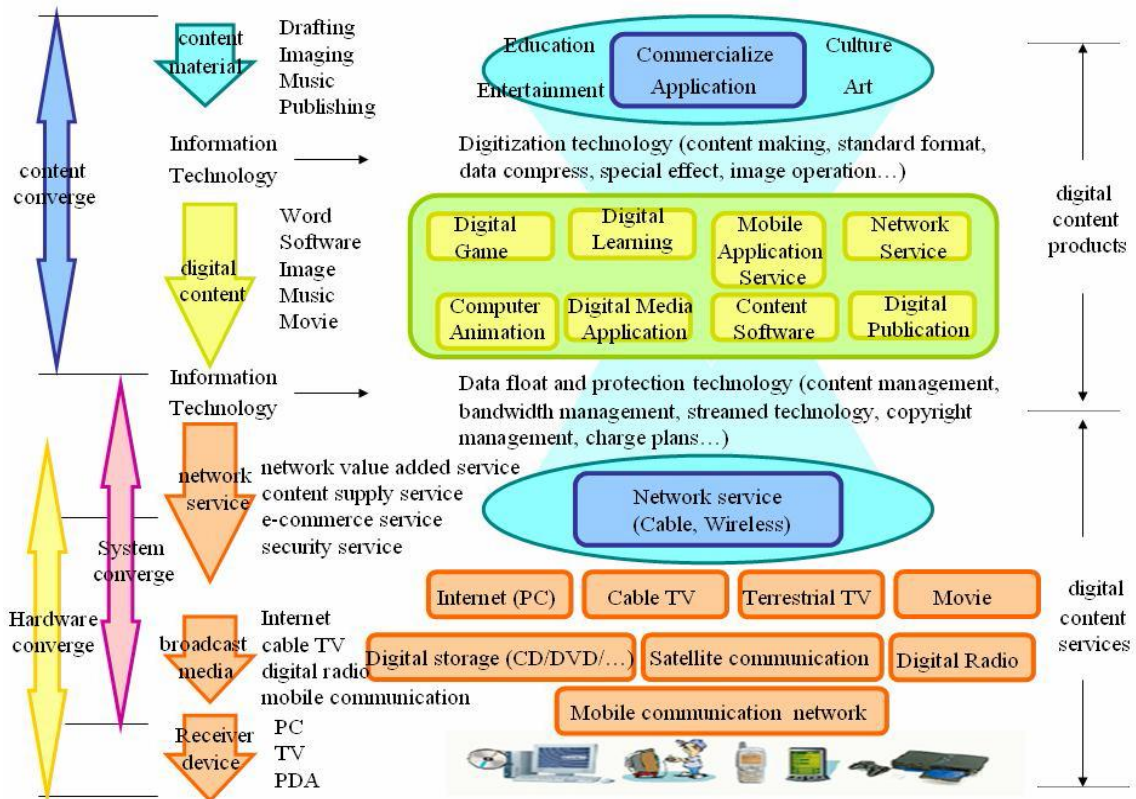
Tiềm năng phát triển thị trường xuất khẩu cho ngành công nghiệp nội dung số Úc là rất cao. Trong khi một số lĩnh vực như trò chơi được tập trung chủ yếu vào xuất khẩu thì thị trường chính vẫn dành cho nội địa. Công nghiệp nội dung kỹ thuật số Úc được dự kiến sẽ tăng trưởng từ 5,9 và 6,3 phần trăm mỗi năm, với sản lượng tăng lên khoảng 36 tỷ đô la Úc (AUD) và 37,6 tỷ AUD của tài khóa 2014-15, đứng trên các ngành như nông nghiệp và truyền thông. Giống như công nghệ thông tin nói chung, nội dung số và các ứng dụng có các đặc điểm của một 'công nghệ nói chung, Ứng dụng rộng rãi trong nhiều ngành công nghiệp nhằm giúp tăng năng suất trong nhiều ngành công nghiệp quan trọng hơn cốt lõi ngành công nghiệp sản xuất nội dung kỹ thuật số của chính nó. Những công nghệ này được áp dụng cho hầu như mọi hoạt động kinh tế, và có khả năng thay đổi các hoạt động này bằng cách cải thiện hiệu quả và tạo ra cơ hội cho các hệ thống mới và các mô hình kinh doanh

7.3.2.2. Công nghiệp nội dung ở Đài Loan

Theo định nghĩa của ngành công nghiệp nội dung số ở Đài Loan, công nghiệp nội dung số là chuyển dữ liệu truyền thông truyền thống sang dữ liệu số nhằm phát triển các cách thức sử dụng khác nhau. Những cách thức này có đặc điểm chung là tạo sự dễ dàng cho kết nối, truyền tải, sao chép, tìm kiếm, chỉnh sửa và sử dụng lại. Đồng thời, những cách thức này gắn kết với dịch vụ, băng thông, phí và bản quyền quản lý. Định dạng nội dung số được gửi từ Internet, mạng truyền thông di động, truyền hình cáp mặt đất, truyền hình vệ tinh, phim ảnh và đài phát thanh kỹ thuật số. Người sử dụng nhận nội dung kỹ thuật số bằng cách sử dụng phần cứng như truyền hình, điện thoại di động, thiết bị hỗ trợ kỹ thuật số cá nhân

(Personal Digital Assistant - PDA), máy tính cá nhân và MP3. Đây là những cấu trúc chính của ngành công nghiệp nội dung kỹ thuật số.

Nội dung số bao gồm nhiều ngành công nghiệp khác nhau, chẳng hạn như internet, phần mềm nội dung, mạng băng rộng không dây và ứng dụng dịch vụ di động.



Hình 7.5: Cấu trúc của ngành công nghiệp nội dung số Đài Loan

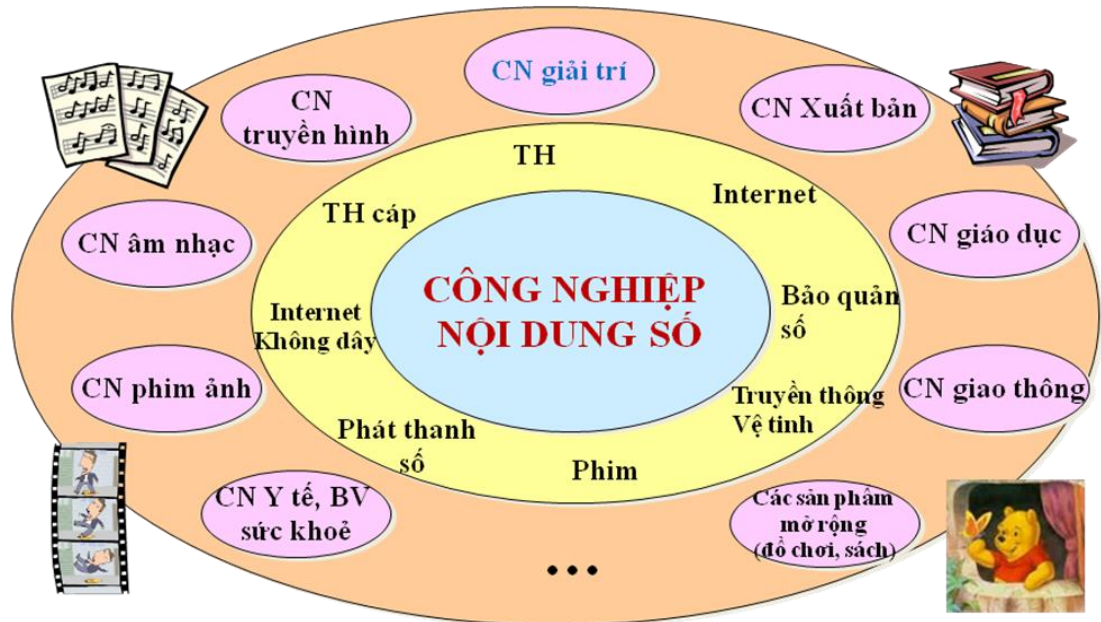
Ngành công nghiệp hoạt hình ở Đài Loan đã chuyển đổi hoàn toàn từ Sản xuất thiết bị gốc (Original Equipment Manufacturing – OEM) sang hợp tác phát triển và xây dựng thương hiệu mới. Ngành công nghiệp trò chơi điện tử kỹ thuật số cũng đã tham gia tích cực vào thị trường quốc tế. Hiện tại Đài Loan đẩy mạnh dịch vụ này với các quốc gia từ Châu Âu, Mỹ và Châu Á Thái Bình Dương để nâng cao năng lực của mình trong nghiên cứu sản phẩm, phát triển và quản lý hoạt động.

Theo Vụ Đầu tư, Bộ Kinh tế Đài Loan (Department of Investment Services, MOEA, ROC), trong thời gian qua, ngành công nghiệp nội

dung số đã phát triển mạnh do chính phủ đầu tư và tài trợ lớn. Kể từ giai đoạn đầu của Dự án "Tăng cường công nghệ nội dung số", giá trị thị trường nội dung tăng từ 133,4 tỷ Tân Đài tệ (NT) năm 2001 lên 360,9 tỷ NT trong năm 2007, mang lại 21% tăng trưởng hàng năm kép (CAGR).

Ở Đài Loan, công nghiệp nội dung số tập trung vào 2 vấn đề quan trọng là mô hình kinh doanh và sở hữu trí tuệ. Chính phủ có thể cung cấp kinh phí để khuyến khích sự hợp tác cung cấp dịch vụ công nghiệp / học viện / nội dung giữa các tổ chức.

Dưới đây là các lĩnh vực chính trong công nghiệp nội dung của Đài Loan:



Hình 7.6 Các lĩnh vực công nghệ nội dung số ở Đài Loan [20]

7.3.3. Công nghiệp nội dung Việt Nam

Chính phủ Việt Nam đã ban hành Nghị định số 71/2007/NĐ-CP, theo đó quy định Công nghiệp công nghệ thông tin có các lĩnh vực sau:

- Công nghiệp phần cứng
- Công nghiệp phần mềm
- Công nghiệp nội dung
- Hoạt động công nghiệp nội dung bao gồm:

- Hoạt động sản xuất sản phẩm nội dung thông tin số;
- Hoạt động cung cấp, thực hiện các dịch vụ nội dung thông tin số.

Nghị định số 71/2007/NĐ-CP cũng quy định rõ dịch vụ nội dung thông tin số gồm:

- Dịch vụ phân phối, phát hành sản phẩm nội dung thông tin số;
- Dịch vụ nhập, cập nhật, tìm kiếm, lưu trữ và xử lý dữ liệu số;
- Dịch vụ quản trị, duy trì, bảo dưỡng, bảo hành các sản phẩm nội dung thông tin số;
- Dịch vụ chỉnh sửa, bổ sung tính năng, bản địa hóa các sản phẩm nội dung thông tin số;
- Dịch vụ đào tạo từ xa; dịch vụ khám, chữa bệnh từ xa; dịch vụ truyền thông được cung cấp trên môi trường mạng;
- Các dịch vụ nội dung thông tin số khác

Ngành CNNDS VN đang tăng trưởng chậm lại: Năm 2009 đạt tốc độ tăng trưởng 56,81%; năm 2010 tăng trưởng đạt 39,71%; năm 2011 tăng trưởng giảm xuống còn 25%; song đến hết năm 2012, ước tính tốc độ tăng trưởng rơi xuống còn 12%, doanh thu ngành đạt khoảng 1,3 tỉ USD, tương đương khoảng 27.000 tỉ đồng.

Tốp 5 doanh nghiệp NDS hàng đầu tại VN hiện nay là VNG, VC Corp, Vatgia.com, FPT và VTC Online.

Công nghiệp nội dung số bao gồm nhiều lĩnh vực như:

- Xuất bản và phát hành nội dung số trên Internet
- Phát triển nội dung cho mạng di động
- Giải trí số (Trò chơi trực tuyến, trò chơi tương tác, nhắn tin trúng thưởng ...)
- Thương mại điện tử
- Thiết kế, quảng cáo, và tiếp thị trên Internet
- Các nội dung giáo dục trực tuyến, học tập điện tử
- Y tế điện tử, Chăm sóc sức khỏe qua mạng

- Thư viện số, bảo tàng số, cơ sở dữ liệu chuyên ngành
- Tra cứu thông tin, dữ liệu số
- Phim số, truyền hình số, hoạt hình, và các dịch vụ/sản phẩm liên quan
- Nội dung học tập, quản lý nội dung và các vấn đề liên quan của công nghiệp sáng tạo số.

Lưu ý: Công nghiệp nội dung số không bao gồm việc cung cấp các dịch vụ công, các trang thông tin miễn phí

Về giáo dục trực tuyến, học tập điện tử

Hiện có trên 50 công ty đăng ký sản xuất, kinh doanh các sản phẩm về học tập điện tử, trong đó có khoảng 30 website cung cấp dịch vụ giáo dục như:

- Cổng eLearning của Bộ GDĐT
- Website của một số trường đại học
- Website của một số công ty cung cấp dịch vụ: www.truongthi.com.vn, www.khoabang.com.vn

Sản phẩm/dịch vụ chính của các công ty gồm:

- Cung cấp bài giảng, tài liệu học tập trực tuyến
- Luyện thi đại học, thi ngoại ngữ
- Từ điển điện tử, tra cứu thông tin về giáo dục
- Các chương trình vừa học vừa chơi của học sinh tiểu học
- Các bài học, bài tập của học sinh theo các lớp học
- Các thí nghiệm ảo về vật lý, hoá học, sinh học phổ thông

Về trò chơi điện tử

Xuất hiện từ những năm 1990, phát triển mạnh trong vài năm trở lại đây. Hiện đang bùng nổ các loại hình như trò chơi trực tuyến, trò chơi tương tác qua truyền hình, trò chơi trên điện thoại di động... Tuy nhiên, các trò chơi nhập khẩu từ nước ngoài hiện đang chiếm lĩnh thị trường Việt Nam. Chỉ có khoảng hơn 10 doanh nghiệp Việt Nam sản xuất, phát

triển trò chơi như FPT, VASC. Đang xuất hiện xu hướng chuyên nghiệp hoá trong việc cung cấp dịch vụ và phát triển công nghiệp trò chơi (Khung pháp lý đang dần được hoàn thiện, đào tạo chuyên nghiệp chuyên gia thiết kế game, lập trình game, ...)

Về phát triển nội dung cho Internet

Xuất hiện và phát triển cùng với sự phát triển của Internet Việt Nam.

Các loại hình dịch vụ chính gồm:

- Thiết kế, xây dựng cổng thông tin điện tử (ePortal)
- Thiết kế website
- Cập nhật, quản lý website
- Báo chí điện tử
- Thiết kế, tiếp thị, quảng cáo trên Internet
- Mua bán qua mạng Internet

Về phát triển nội dung cho mạng di động

Xuất hiện từ năm 2002 với sự tiên phong của CDiT với dịch vụ dự đoán kết quả World Cup. Hiện có khoảng 20 công ty tham gia. Các loại hình dịch vụ chính gồm:

- Tải nhạc chuông, logo, hình nền cho thiết bị di động
- Tin nhắn trúng thưởng
- Tin nhắn thông tin kinh tế xã hội
- Tin nhắn có nội dung chuyên sâu: tư vấn sức khoẻ, tư vấn an toàn giao thông, ...

Mạng điện thoại di động thế hệ thứ tư (4G) được hứa hẹn sẽ làm thay đổi hoàn toàn cách sử dụng smartphone, và qua đó lối sống, của mọi người trong những năm tới. Nhờ khả năng truyền dữ liệu tốc độ cao lên tới 300 Mbps (trong điều kiện lý tưởng), LTE có thể giúp người dùng tải về một bộ phim chỉ trong vài giây, còn dữ liệu video streaming gần như đến tức thì sau thao tác nhấn phím.

Việt Nam là một trong những quốc gia cuối cùng của khu vực triển

khai 4G (dự kiến 2016). LTE vào Việt Nam chậm có nhiều lý do, trong đó chủ yếu là do các nhà mạng chưa muốn nhập cuộc do chưa thu hồi đủ vốn và có lợi nhuận từ việc đầu tư vào mạng 3G trước đây. Xây dựng tiếp mạng 4G đồng nghĩa với việc phải đổi mới nâng cấp hệ thống truyền dẫn, trạm BTS... tốn kém ước tới 1 tỉ USD. Ngoài các điều kiện về quy hoạch băng tần (nhà nước), mức độ sẵn sàng của hạ tầng mạng (nhà mạng), triển vọng phát triển 4G tại Việt Nam còn phụ thuộc rất nhiều vào sự năng động của ngành công nghiệp nội dung số. Nếu chỉ để lướt web đọc tin tức, tra cứu thông tin, trao đổi email... mạng 3G hiện nay hoàn toàn có thể làm tốt vai trò đó. Chỉ khi nào các sản phẩm nghe nhìn như IPTV, game, xem phim, dịch vụ dữ liệu lớn... đủ hấp dẫn người dùng Việt Nam để họ chấp nhận trả một khoản tiền cho gói dữ liệu 4G, khi ấy LTE mới có điều kiện phát triển.

Về thư viện điện tử, kho dữ liệu số

Nhiều công ty đã phát triển, Việt hóa và cung cấp phần mềm quản trị tài nguyên số. Một số sản phẩm chính gồm Libol, Ilib, Lạc Việt, Dspace, GreenStone.

Xuất hiện cùng với các chương trình, chiến lược của ngành, quốc gia. Nhiều công ty cung cấp dịch vụ thiết kế, xây dựng, cập nhật, quản lý nội dung cho các kho dữ liệu số. Một số kho dữ liệu tiêu biểu:

- Kho dữ liệu tra cứu văn bản pháp luật
- Kho dữ liệu chuyên ngành của các bộ, ngành.
- Kho dữ liệu dân số (UB dân số thực hiện)
- Kho dữ liệu tài nguyên đất (Bộ tài nguyên môi trường chủ trì)
- Kho dữ liệu cán bộ công chức (Bộ Nội vụ chủ trì).

Về phim số và đa phương tiện số

Dịch vụ này đang phát triển mạnh. Truyền hình Internet phát triển nhanh chóng với các nhà cung cấp chính: VTC, VASC, FPT Telecom, ... Âm nhạc số phát triển theo hướng chuyên nghiệp, tôn trọng bản quyền,

điển hình là FPT Music với nhacso.net, VASC với kho nhạc số trên 2500 bài hát Việt. Đã có một số nhà cung cấp sản phẩm/dịch vụ đa phương tiện, sản phẩm tạo hiệu ứng hình ảnh 3D, âm thanh nổi.

Mặc dù ngành công nghiệp nội dung số vẫn tiếp tục phát triển, song, trong chiến lược phát triển, kinh doanh, nhiều doanh nghiệp hoạt động trong lĩnh vực này còn gặp khó khăn, lúng túng do thiếu "một hệ thống thông tin đầy đủ".

CNNDS đang từng bước khẳng định là ngành kinh tế mới, có nhiều tiềm năng, đem lại giá trị cao, thúc đẩy sự phát triển của xã hội thông tin... Với mục tiêu thúc đẩy phát triển CNNDS, Nhà nước đã ban hành một số chính sách, chương trình về hoạt động này. Theo đánh giá của một số chuyên gia, CNNDS nước ta hiện nay phát triển còn chậm, chưa đáp ứng được nhu cầu thực tế. Hằng năm, Bộ Thông tin và Truyền thông (TT và TT), Hiệp hội Phần mềm và Dịch vụ công nghệ thông tin (CNTT) Việt Nam, một số tổ chức nghiên cứu thị trường... đều có điều tra, khảo sát về hoạt động ngành CNNDS. Bộ TT và TT đã xây dựng, xuất bản cuốn "Sách trắng" về CNTT và truyền thông phục vụ quản lý Nhà nước. Điều đáng nói là những số liệu điều tra về CNNDS khi công bố phần lớn tập trung vào doanh thu, số lượng lao động, mức lương bình quân... mà chưa phản ánh đầy đủ, chi tiết về CNNDS Việt Nam. Các cuộc điều tra, khảo sát vẫn mang tính cục bộ, không thường xuyên và thiếu tính hệ thống. Vì vậy, kết quả không phản ánh toàn diện về CNNDS, cũng như chưa được lưu trữ thành một hệ thống nhất quán. Mặt khác, việc tiếp cận thông tin về kết quả các cuộc khảo sát điều tra đối với những doanh nghiệp quan tâm còn gặp nhiều khó khăn.

Bên cạnh đó, Việt Nam hiện vẫn chưa có bộ chỉ tiêu chuẩn hóa để thu thập thông tin về ngành CNNDS, cũng như chưa xây dựng được hệ thống thu thập thông tin một cách chính thống. Nhiều chuyên gia cho rằng, phần lớn doanh nghiệp rất cần thông tin về ngành để giúp họ xác

định chiến lược, mục tiêu phát triển cho từng năm hoặc lâu hơn. Do thiếu sự hợp tác từ các doanh nghiệp cùng lĩnh vực, cho nên không chỉ doanh nghiệp mà cả cơ quan quản lý Nhà nước cũng gặp khó khăn do thiếu thông tin. Thực tế cho thấy, năm 2010, Hiệp hội Phần mềm và Dịch vụ CNTT Việt Nam đã làm được báo cáo toàn cảnh về công nghiệp phần mềm Việt Nam nhưng những năm sau đó không thực hiện được vì các doanh nghiệp không cung cấp số liệu. Những doanh nghiệp nếu có cung cấp số liệu thì cũng không đáng tin cậy. Theo các chuyên gia, một trong những nguyên nhân khiến các nhà đầu tư nước ngoài "ngại" đầu tư vào lĩnh vực CNMDS tại Việt Nam là do nước ta chưa có "một hệ thống thông tin đầy đủ".

Có thể nói, sau hơn 10 năm phát triển, khái niệm nội dung số vẫn chưa được hiểu một cách rõ ràng, thống nhất tại Việt Nam, dẫn đến việc những chính sách hỗ trợ, quản lý của Nhà nước chưa thực sự bám sát thị trường. Tại "Hội thảo Việt Nam - Hàn Quốc về Công nghiệp nội dung số" diễn ra sáng 27/3/2013, Vụ Công nghệ thông tin (Bộ Thông tin - Truyền thông) cho rằng, đây chính là một nguyên nhân quan trọng khiến cho tỷ trọng của nội dung số trong ngành công nghiệp CNTT nói chung vẫn còn ở mức khá khiêm tốn.

Cụ thể, Sách trắng CNTT 2012 do Bộ TT&TT phát hành cho biết, tổng doanh thu ngành Công nghiệp CNTT đạt 25 tỷ USD trong năm 2011, nhưng đóng góp của ngành nội dung số mới chỉ dừng ở mức 1,1 tỷ USD. Đây là con số chưa tương xứng với tiềm năng phát triển của Việt Nam, như nhiều chuyên gia trong và ngoài nước đều nhận định.

Thực tế là thị trường nội dung số đã tăng trưởng nhanh và đạt đến đỉnh cao trong những năm 2008-2009. Thế nhưng nhiều hiện tượng, mặt trái đã xuất hiện như tin nhắn rác, thư rác, nhiều loại hình nội dung không lành mạnh (nhắn tin để xem bói, cá độ ...). Những loại hình kinh doanh kiểu này có tác động rất lớn đến các nhà làm chính sách và gây bức xúc

trong xã hội. Ngoài ra, mặt trái của game online, của mạng xã hội cũng gây nhiều quan ngại. Kết hợp với khủng hoảng kinh tế thế giới trong giai đoạn này, thị trường nội dung số đã tăng trưởng chậm lại rõ rệt và dự kiến doanh thu 2012 chỉ đạt khoảng 1,3 -1,5 tỷ USD. So với tốc độ tăng trưởng 20-40% những năm đỉnh cao thì tốc độ tăng trưởng hiện tại chỉ còn trên dưới 10%.

Về chính sách ưu đãi doanh nghiệp nội dung số cũng gặp nhiều khó khăn khi ở Việt Nam, doanh nghiệp vừa sản xuất nội dung, vừa cung cấp dịch vụ là mô hình rất phổ biến. Rất hiếm doanh nghiệp chuyên sản xuất nội dung như ở nước ngoài nên tách bạch ra để xây dựng chính sách hỗ trợ là việc không đơn giản với cơ quan quản lý.

TS Hoàng Lê Minh, Viện trưởng Công nghiệp Phần mềm và Nội dung số Việt Nam cho rằng trong thời gian tới, Việt Nam cần xây dựng một hệ thống thông tin toàn diện về ngành CNNDS. Điều này phải bắt đầu từ việc xây dựng bộ chỉ tiêu chuẩn mực làm căn cứ thu thập dữ liệu, phục vụ theo dõi, phân tích, đánh giá và dự báo ngành CNNDS. Bộ chỉ tiêu phải bao gồm các chỉ tiêu về nhân lực, hạ tầng, thị trường, công nghệ, chính sách và các thông tin thống kê cần thiết khác. Trên cơ sở bộ chỉ tiêu, Bộ TT và TT cần tổ chức hệ thống thu thập thông tin về CNNDS một cách chính thống thông qua hệ thống báo cáo, các cuộc điều tra, khảo sát, trên cả nước về lĩnh vực công nghiệp này. Đây là một trong các cơ sở quan trọng để Nhà nước lấy đó làm căn cứ, đưa ra các chiến lược, quy hoạch, kế hoạch phát triển, chính sách và giải pháp phù hợp. Để giải quyết việc này, Nhà nước cần quan tâm hơn đến việc xây dựng một hệ thống thông tin về ngành CNNDS tại Việt Nam. Việc xây dựng hệ thống thông tin đầy đủ không chỉ góp phần thúc đẩy phát triển CNNDS mà còn sớm đưa Việt Nam trở thành nước mạnh về CNTT và truyền thông.

7.3.4. Công nghiệp nội dung trong hoạt động thông tin - thư viện

CNND trong lĩnh vực thông tin thư viện là ngành công nghiệp liên

quan đến sản xuất, phân phối/phát hành, cung cấp những sản phẩm và dịch vụ thông tin, bao gồm xuất bản tài liệu (sách, báo, tạp chí,...); xây dựng và cung cấp CSDL; cung cấp các dịch vụ thông tin trực tuyến.

Các loại hàng hoá thông tin gồm:

- Tài liệu KH&CN;
- Dịch vụ thông tin trực tuyến

Đối với các cơ quan thông tin – thư viện hiện nay thông tin nhìn chung được đóng gói chủ yếu là:

- *Tạp chí khoa học, công nghệ*: Các tạp chí khoa học kỹ thuật được chọn cho từng ngành và được đóng gói trong một CSDL cho từng ngành.
- *Sáng chế (Patent)*: Các sáng chế là một tư liệu công nghệ cực kỳ quan trọng. Số lượng các sáng chế được cấp bằng hàng năm trên thế giới lên tới hàng triệu. Công nghiệp đóng gói sáng chế vào các kiểu CSDL khác nhau là một ngành có doanh số lớn ở những nước có nền khoa học công nghệ và công nghiệp phát triển.
- *Kết quả nghiên cứu khoa học*: Là những công trình nghiên cứu khoa học thuộc các chương trình, đề tài, hội nghị hội thảo...
- *Bản đồ*: Do sự phát triển của công nghiệp phần mềm, các bản đồ đã được số hoá và được đóng gói vào những CSDL. Đây là những CSDL rất cần thiết cho việc quy hoạch phát triển kinh tế.
- *Văn bản pháp quy của một nước*
- *CSDL về kinh tế – xã hội như dân số, đất đai, tài nguyên...* phục vụ cho các nhu cầu phát triển kinh tế – xã hội .
- *Phim ảnh, tranh,...* các CSDL này được đóng gói, được bán, cấp cho những đối tượng có nhu cầu.
- *Sách*: hiện nay các nhà xuất bản đã bắt đầu chú trọng công tác xây dựng các CSDL sách điện tử (e-book) để cung cấp cho người dùng.

Đối với nhiều cơ quan thông tin – thư viện thì công nghiệp nội dung còn được hiểu chính là phần dịch vụ cung cấp thông tin theo chủ đề tới các nhóm đối tượng người dùng tin xác định. Ví dụ: Bảo mật thông tin là chủ đề lớn được cung cấp cho các doanh nghiệp, sinh viên, giảng viên...hoạt động trong lĩnh vực công nghệ thông tin. Đây là hoạt động mà ở các nước tiên tiến có tốc độ phát triển nhanh, đồng thời cũng là những hoạt động đem lại nhiều giá trị gia tăng. Khách hàng của hoạt động này là những doanh nghiệp, tổ chức, cá nhân và các cơ quan nhà nước có nhu cầu. Các thông tin được cung cấp thường rất đa dạng, từ nhu cầu thị trường trong và ngoài nước, thông tin về sản phẩm, về đối thủ cạnh tranh... đến các nhu cầu đầu tư trong nước và ngoài nước, các lợi thế đầu tư, các bằng sáng chế, bí quyết kỹ thuật...

CÂU HỎI ÔN TẬP CHƯƠNG 7

1. Phân tích khái niệm công nghệ nội dung và công nghiệp nội dung!
2. Trình bày các khía cạnh chủ yếu của công nghệ nội dung
3. Giới thiệu đặc điểm chung của công nghiệp nội dung
4. Phân tích tình hình phát triển công nghiệp nội dung tại Việt Nam, Úc và Đài Loan!

TÀI LIỆU THAM KHẢO

1. Andra Leurdijk , [et al]. Statistical, Ecosystems and Competitiveness Analysis of the Media and Content Industries: A Quantitative overview. EU JRC technical report, 2012
2. Barker, Joe. Finding Information on the Internet: A Tutorial. University of California at Berkeley. Retrieved August 18, 2006, from ([http://www.lib.berkeley.edu/ TeachingLib/Guides /Internet/ FindInfo.html](http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/FindInfo.html))
3. Barnes, S. J . Becoming a digital library. New York, Marcel Dekker, 2004.
4. Borgman, C.L. What are digital libraries – competing visions. Information Processing and Management, vol. 35, 1999.
5. Cao Minh Kiêm. Bài giảng thư viện số và công nghệ nội dung. Hà Nội, 2010
6. Chowdhury, G. G., Chowdhury, S. Introduction to Digital Libraries. London: Facet, 2003
7. Cohn, John M., Ann L.Kelsey, Managing the library automation project, 2001
8. Deakin University Library. 2006. Searching the Internet. Retrieved October 3, 2006, from [http://www.deakin.edu.au/ library/tutorials/ smartsearcher/ internet.html](http://www.deakin.edu.au/library/tutorials/smartsearcher/internet.html)
9. Deanna B.Marcum. Development of Digital Libraries, Greenwood press, 2001
10. Digital Content Industry Roadmapping Study. Mel., Australian Interactive Media Industry Association, 2005
11. Digital Libraries: Design, Development, and Impact/ Yin-Leng Theng, Schubert Foo, Dion Goh and Jin-Cheon Na; Information

- Science Reference, 2009
12. Dobрева, M., et al. User studies for digital library development. London, Facet, 2012
 13. Encyclopedia of library and information science. Volume 1, 2nd ed. New York : Marcel Dekker, 2003
 14. The Future Content : Discussions on the future of European Electronic Publishing Version I. Sep 1997
 15. Hanson, T., Day. Managing the electronic library. Lond. : Browker Sauer, 1998
 16. Hoàng Thị Thu Hương. Công nghiệp nội dung và một vài giải pháp phát triển ở Việt Nam. <http://vietnamlib.net>
 17. [http://en.wikipedia.org/wiki/Content_\(media\)](http://en.wikipedia.org/wiki/Content_(media))
 18. <http://vi.wikipedia.org>
 19. (<http://www.plos.org/about/open-access><http://www.greenstone.org/>)
 20. Joan M. Reitz. Dictionary for Library and Information Science. Westport, Connecticut: Libraries Unlimited, 2004
 21. Kresh, D. (2007). The whole digital library handbook. Chicago, American Library Association.
 22. Lesk, M. Understanding Digital Libraries (Second ed.). San Francisco, CA: Morgan Kaufman Publishers, 2004.
 23. Marcum, D. B., George, G. Digital library development : the view from Kanazawa. Englewood, Colo., Libraries Unlimited ; Oxford : Harcourt Education [distributor], 2006
 24. Nguyễn Huy Chương (chủ nhiệm đề án). Nghiên cứu, thiết kế mô hình và xây dựng thử nghiệm nguồn học liệu trực tuyến phục vụ đào tạo chất lượng cao cho một số ngành, chuyên ngành tại Đại học Quốc gia Hà Nội. *Đề án cấp Đại học Quốc gia Hà Nội*, 2009
 25. Nguyễn Huy Chương. Nguyên lý và nội dung cơ bản của thư viện điện tử. H., ĐHQGHN, 2013

26. Nguyen Huy Chuong. The Digitization Activities of Academic Libraries in Vietnam. *PNC Annual Conference and Joint Meetings*, 2008
27. Nguyễn Huy Chương. Nguyên lý và chính sách phát triển thư viện số (trong sách chuyên khảo *Xây dựng và phát triển thư viện số Việt Nam: Quá khứ - hiện tại – tương lai*). H.: NXB Đại học Quốc gia Hà Nội, 2017
28. Nguyễn Huy Chương. Sử dụng phần mềm mã nguồn mở - giải pháp tối ưu cho thư viện quy mô vừa và nhỏ ở Việt Nam. *Tạp chí Thông tin và tư liệu*, số 1, 2017.
29. Nguyễn Huy Chương. Xây dựng thư viện điện tử và phát triển nguồn tài nguyên số trong hệ thống thư viện đại học Việt Nam. *Kỷ yếu Hội thảo Khoa học Phát triển và chia sẻ nguồn tài nguyên số trong các thư viện đại học và nghiên cứu*, 2009
30. Nguyễn Huy Chương. Xây dựng và phát triển thư viện điện tử trong hệ thống thư viện đại học Việt Nam. *Kỷ yếu hội thảo khoa học Tiếp cận xây dựng thư viện số ở Việt Nam: Hiện trạng và vấn đề*, 2007
31. Nguyễn Huy Chương, Nguyễn Tiến Hùng. DSPACE – Giải pháp tạo lập, lưu trữ và phổ biến tài nguyên điện tử cho các thư viện ở Việt Nam. *Kỷ yếu Hội thảo khoa học Xây dựng và chia sẻ nguồn lực thông tin địa phương dạng số phục vụ bảo tồn di sản và phát triển kinh tế - xã hội*, 2011
32. Nguyễn Huy Chương, Nguyễn Tiến Hùng. Học liệu mở và hướng phát triển tài nguyên số tại các thư viện đại học Việt Nam. *Kỷ yếu Hội thảo khoa học 50 năm đào tạo nguồn nhân lực thông tin – thư viện, Trường Đại học Văn hóa Hà Nội*, 2011
33. Nguyễn Huy Chương, Trần Mạnh Tuấn, Các vấn đề về chính sách phát triển nguồn học liệu trực tuyến tại Đại học Quốc gia Hà Nội,

Kỷ yếu Hội thảo Khoa học Phát triển và chia sẻ nguồn tài nguyên số trong các thư viện đại học và nghiên cứu, Hà Nội, 2009

34. Nguyễn Minh Hiệp, Sử dụng phần mềm nguồn mở Thư viện số Greenstone để xây dựng kho tài nguyên học tập nhằm nâng cao chất lượng đào tạo nguồn nhân lực. Tp HCM, Đại học KHTN, 2006
35. Nguyễn Thị Bắc, Bài giảng Lập kế hoạch tự động hóa trong thư viện . Tp HCM, 2008
36. Nick More. The information society. In “World Information Report 1997-1998”. UNESCO Publishing
37. NMC Horizon Library Project, 2014
38. Suber P. *Open Access*. The MIT Press Essential Knowledge series, 2006
39. Sullivan, Danny. Kids Search Engines. Search Engine Watch. Retrieved January 7, 2008, from <http://searchenginewatch.com/showPage.html?page=2156191>
40. Từ điển tiếng Việt. H., Từ điển bách khoa, 2010
41. USITC Publication 4415 : Digital Trade in the U.S. and Global Economies, Part 1. Investigation No. 332-531 July, 2013
42. Verheul, I., et al. Digital library futures: user perspectives and institutional strategies. The Hague, Walter D e Gruyter, 2010
43. Vũ Thị Nha, Tìm kiếm Thông tin trên Internet, Trung tâm Thông tin Phát triển Việt Nam.
44. Vũ Văn Sơn, Giáo trình biên mục mô tả. - Hà Nội: Nxb Đại học Quốc gia, 2002.
45. Witten, I. H., & Bainbridge, D. How to Build a Digital Library. San Francisco, CA: Morgan Kaufman Publishers, 2003
46. World Encyclopedia of Library and Information Services. 5th ed. , Chicago: American Library Association, 2001
47. World Information Report 1997-1998”. UNESCO Publishing.

48. Yin-Leng Theng, Schubert Foo, Design and Usability of Digital Libraries: Case Studies in the Asia Pacific, 2005