

**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

<b>LỜI NÓI ĐẦU.....</b>	<b>5</b>
<b>CHƯƠNG 1: TỔNG QUAN VỀ THƯ VIỆN SỐ.....</b>	<b>6</b>
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

---

<b>CÂU HỎI ÔN TẬP CHƯƠNG 1 .....</b>	<b>46</b>
<b>CHƯƠNG 2: CÁC THÀNH PHẦN CỦA THƯ VIỆN SỐ .....</b>	<b>47</b>
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
<b>CÂU HỎI ÔN TẬP CHƯƠNG 2 .....</b>	<b>121</b>
<b>CHƯƠNG 3: SỬU TẦM VÀ KHAI THÁC THÔNG TIN SỐ.....</b>	<b>122</b>
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
<b>CÂU HỎI ÔN TẬP CHƯƠNG 3</b> .....	<b>151</b>
<b>CHƯƠNG 4: BỘ SƯU TẬP VÀ BIÊN MỤC TÀI LIỆU SỐ</b> .....	<b>152</b>
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. Các bộ sưu tập theo loại hình xuất bản.....	156
4.3.2. Các bộ sưu tập theo dạng lưu trữ tài liệu .....	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. Biên mục Analog.....	174
4.6.2. Biên mục Digital.....	174
4.6.3. Chuyển đổi MARC sang Dublin Core .....	175
<b>CÂU HỎI ÔN TẬP CHƯƠNG 4</b> .....	<b>177</b>
<b>CHƯƠNG 5: CÁC VẤN ĐỀ VỀ SỞ HỮU TRÍ TUỆ VÀ BẢN QUYỀN</b> .....	<b>178</b>
5.1 MỘT SỐ KHÁI NIỆM VỀ SỞ HỮU TRÍ TUỆ .....	178
5.2 CÁC VẤN ĐỀ VỀ BẢN QUYỀN.....	185
<b>CÂU HỎI ÔN TẬP CHƯƠNG 5</b> .....	<b>188</b>
<b>CHƯƠNG 6: XÁC ĐỊNH CHÍNH SÁCH VÀ XÂY DỰNG KẾ HOẠCH PHÁT TRIỂN THƯ VIỆN SỐ</b> .....	<b>189</b>
6.1 XÁC ĐỊNH CHÍNH SÁCH PHÁT TRIỂN THƯ VIỆN SỐ .....	189
6.1.1 Cấu trúc của thư viện số.....	189
6.1.2 Hạ tầng cơ sở kỹ thuật.....	190
6.1.3 Kho tư liệu số hóa.....	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
<b>CÂU HỎI ÔN TẬP CHƯƠNG 6.....</b>	<b>196</b>
<b>CHƯƠNG 7: CÔNG NGHỆ NỘI DUNG VÀ CÔNG NGHIỆP NỘI</b>	
<b>DUNG.....</b>	<b>197</b>
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
<b>CÂU HỎI ÔN TẬP CHƯƠNG 7.....</b>	<b>227</b>
<b>TÀI LIỆU THAM KHẢO.....</b>	<b>228</b>

## 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 lôgích", "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]

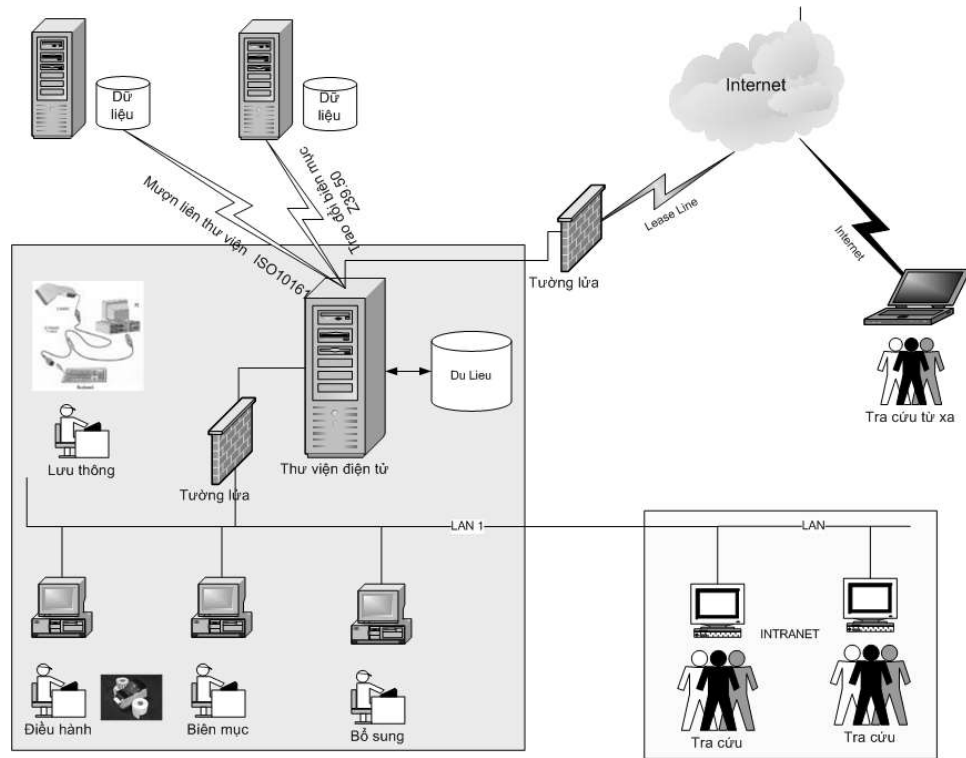
Tuy ý kiến chưa hoàn toàn thống nhất, nhưng tựu chung lại, ta có thể nhận dạng một số đặc điểm của thư viện điện tử như sau:

- 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