

ĐOÀN PHAN TÂN

THÔNG TIN HỌC

GIÁO TRÌNH DÀNH CHO SINH VIÊN
NGÀNH THÔNG TIN - THƯ VIỆN VÀ QUẢN TRỊ THÔNG TIN

Enter



NHÀ XUẤT BẢN ĐẠI HỌC QUỐC GIA HÀ NỘI

PGS.TS. ĐOÀN PHAN TÂN

THÔNG TIN HỌC

**GIÁO TRÌNH DÀNH CHO SINH VIÊN
NGÀNH THÔNG TIN - THƯ VIỆN VÀ QUẢN TRỊ THÔNG TIN**
(In lần thứ hai, có sửa chữa và bổ sung)

NHÀ XUẤT BẢN ĐẠI HỌC QUỐC GIA HÀ NỘI

LỜI NÓI ĐẦU

Năm 1990, giáo trình “Cơ sở Thông tin học” của tác giả lần đầu tiên được xuất bản. Sự ra đời của cuốn sách đã kịp thời đáp ứng với yêu cầu đổi mới nội dung chương trình và nâng cao chất lượng đào tạo cán bộ ngành thông tin - thư viện ở trường Đại học Văn hoá Hà Nội lúc bấy giờ. Từ đó đến nay cuốn sách trở thành một trong những tài liệu phục vụ cho hoạt động học tập, giảng dạy của các hệ đào tạo cán bộ ngành thông tin - thư viện.

Mười năm qua, sự phát triển mạnh mẽ của cuộc cách mạng khoa học và công nghệ, đặc biệt là việc ứng dụng rộng rãi công nghệ thông tin vào hoạt động thông tin - thư viện đã làm nảy sinh nhiều vấn đề mới, tạo ra bước phát triển mới trong lý luận và thực tiễn của thông tin học và hoạt động thông tin khoa học. Ngoài ra sự nghiệp đào tạo cán bộ thông tin - thư viện cũng đang đặt ra những yêu cầu mới, cao hơn về chất lượng. Vì vậy chúng tôi thấy cần phải viết lại cuốn “Cơ sở Thông tin học” dưới góc độ của ngày hôm nay. So với lần xuất bản năm 1990, lần này cuốn sách được cập nhật, bổ sung thêm nhiều nội dung mới, đồng thời các chương mục được cấu trúc lại cho khoa học, chặt chẽ, logic hơn và với nội

dung đó, cuốn sách có thể mang tên mới, gọn hơn: "Thông tin học".

"Thông tin học" cung cấp những kiến thức cơ bản, hệ thống về thông tin và thông tin học, cũng như những vấn đề về lý luận và phương pháp của các quá trình xử lý và khai thác thông tin. Nội dung cuốn sách được cơ cấu thành ba phần, bao gồm 7 chương.

Phần một gồm chương 1 và chương 2, giới thiệu tổng quát về khái niệm thông tin, các quá trình thông tin, thông tin học và những vấn đề liên quan đến quan hệ giữa thông tin và tiến bộ xã hội.

Các khái niệm về thông tin, thuộc tính của thông tin và các quá trình thông tin được trình bày khá chi tiết ở chương 1 và được xem là đối tượng và phạm vi nghiên cứu chủ yếu của thông tin học. Lịch sử phát triển của thông tin được trình bày trong sự gắn bó hữu cơ với lịch sử phát triển của kỹ thuật và công nghệ truyền tin. Các khoa học liên quan như lý thuyết thông tin, điều khiển học, ngôn ngữ học, tin học được giới thiệu với những nội dung tổng quát để nói lên rằng khoa học thông tin là một khoa học đa ngành ở trình độ cao. Dây chuyền thông tin tư liệu được trình bày như là một quá trình thông tin trong hoạt động thông tin khoa học, đồng thời cũng là giới thiệu mở đầu cho các quá trình xử lý và khai thác thông tin sẽ được trình bày chi tiết ở phần hai. Ngoài ra tác giả cũng điểm qua lịch sử phát triển của thông tin học và hoạt động thông tin khoa học ở Việt Nam.

Ở chương 2, bên cạnh việc khẳng định vai trò của thông tin là nguồn lực phát triển của kinh tế và xã hội, là cơ sở của

hoạt động quản lý điều hành, còn giới thiệu những vấn đề xã hội mới nổi lên dưới tác động của bùng nổ thông tin và công nghệ thông tin hiện đại như: kinh tế thông tin, thị trường thông tin, hiện tượng tin học hoá xã hội và xã hội thông tin.

Phần hai gồm các chương 3, 4, 5, 6, trình bày các quá trình xử lý, lưu trữ, khai thác, phổ biến và sử dụng thông tin. Đó cũng chính là những vấn đề lý luận và phương pháp luận cơ bản của thông tin học. Nội dung được trình bày ở đây tuân theo một logic chặt chẽ: các loại hình tài liệu gắn với nguồn tin và vấn đề bổ sung vốn tư liệu, xử lý thông tin gắn với sản phẩm thông tin mà nó tạo ra, lưu trữ thông tin là để phục vụ cho việc tìm tin, các cơ quan đơn vị thông tin có trách nhiệm đảm bảo các dịch vụ thông tin phục vụ cho đối tượng người dùng tin. Nội dung của các chương này được trình bày trên quan điểm cơ bản hệ thống, đồng thời cập nhật những vấn đề mới nhất trong hoạt động thông tin dưới tác động của công nghệ thông tin hiện đại như: các sản phẩm thông tin điện tử, khổ mẫu biên mục đọc máy MARC, hệ thống thông tin online, mạng thông tin toàn cầu Internet...

Phần ba là nội dung của chương 7, trình bày những kiến thức cơ bản về các hệ thống thông tin. Xuất phát từ chức năng, nhiệm vụ của tổ chức và quản lý, vai trò của thông tin trong hoạt động quản lý, chương này giới thiệu các hệ thống thông tin quản lý với sự trợ giúp của máy tính điện tử. Đó là các hệ thống thông tin tác nghiệp, các hệ thống thông tin hỗ trợ quyết định, hệ thống thông tin điều hành và hệ chuyên gia.

Cuốn sách này được viết trên cơ sở cuốn sách cùng tên của tác giả xuất bản năm 1990, các bài giảng về Thông tin học mà tác giả trình bày cho sinh viên ngành thông tin - thư viện bậc đại học và cao học, ở trường Đại học Văn hoá Hà Nội và một số kết quả nghiên cứu của tác giả từ năm 1991 đến nay.

Trong quá trình biên soạn cuốn sách chúng tôi nhận được sự đóng góp chân tình của các đồng nghiệp ở bộ môn Thông tin học, khoa Thông tin - thư viện, các chuyên gia đang làm việc tại các trung tâm thông tin, thư viện. Cho phép tôi được bày tỏ lời cảm ơn chân thành.

Mặc dù đã có nhiều cố gắng, song cuốn sách không thể tránh khỏi những thiếu sót. Rất mong sự góp ý của đồng nghiệp và bạn đọc để cuốn sách được tốt hơn.

Hà Nội, tháng 5 năm 2000

TÁC GIẢ

LỜI NÓI ĐẦU

CHO LẦN XUẤT BẢN THỨ HAI

Năm 2001 giáo trình “Thông tin học” của tác giả được xuất bản. Tác giả rất cảm kích trước thịnh tình mà các bạn đồng nghiệp và sinh viên ngành thông tin – thư viện cả nước dành cho cuốn sách. Từ đó đến nay đã được 5 năm. Từ hơn một năm nay, nhiều bạn đồng nghiệp ở các cơ sở đào tạo của ngành đã khích lệ tác giả cho tái bản cuốn giáo trình này. Khi đó tác giả đã nghĩ đến việc phải viết những dòng bổ sung, để hoàn thiện thêm và cập nhật nội dung cho cuốn sách.

Trong lần xuất bản này, ở chương 1, tác giả viết rõ thêm về mối quan hệ giữa các khái niệm thông tin, dữ liệu và tri thức; thế nào là một thông tin tốt; viết bổ sung thêm một mục về nghề thông tin và các yêu cầu về năng lực đối với nghề thông tin hiện nay. Trong chương 2, tác giả thấy cần bổ sung thêm phần nói về bối cảnh nền kinh tế thế giới mới trong môi trường bùng nổ thông tin và hội nhập kinh tế toàn cầu, dẫn đến việc ra đời thị trường thông tin và kinh tế thông tin. Trong chương 3, các nguồn thông tin truyền thống được bổ sung thêm một mục nói về nguồn thông tin điện tử – thông tin số và công nghiệp nội dung số, một khái niệm chỉ mới xuất hiện trong mấy năm gần đây. Trong chương 4, do ý

nghĩa quan trọng của nó, phần nói về AACR2 được giới thiệu kỹ hơn và tách thành một tiểu mục riêng. Ngày nay thư viện điện tử không chỉ còn là một khái niệm, mà đã trở thành một thực thể hiện hữu, sinh động với cấu trúc rõ ràng và mô hình hoạt động hiệu quả. Vì vậy, trong chương 6 tác giả đã tách ra một tiểu mục để viết riêng về thư viện điện tử. Ngoài ra ở đoạn này, đoạn kia trong cuốn sách, tác giả cũng dành nhiều thời gian viết lại, viết thêm để làm rõ hơn, sâu hơn những nội dung đã đề cập tới. Đó là kết quả của những tiếp thu, nghiên cứu, trải nghiệm của tác giả về lý luận và thực tiễn của thông tin học và hoạt động thông tin khoa học trong mấy năm qua.

Rất mong cuốn sách đóng góp được phần nhỏ bé của mình trong sự nghiệp đào tạo cán bộ của ngành.

Do điều kiện thời gian và khả năng có hạn, cuốn sách không tránh khỏi những thiếu sót. Rất mong được sự đóng góp của các bạn đồng nghiệp và đông đảo bạn đọc gần xa.

Hà Nội tháng 8 năm 2006

TÁC GIẢ

MỤC LỤC

	trang
LỜI NÓI ĐẦU	3
LỜI NÓI ĐẦU CHO LẦN XUẤT BẢN THỨ HAI	7
MỤC LỤC	9
<i>Chương 1. THÔNG TIN, CÁC QUÁ TRÌNH THÔNG TIN VÀ THÔNG TIN HỌC</i>	17
1. Thông tin	17
1.1. Khái niệm thông tin. Dữ liệu, thông tin và tri thức	17
1.2. Các đặc trưng của một thông tin tốt	23
1.3. Các yếu tố cơ bản trong xử lý thông tin	24
2. Các thuộc tính của thông tin	26
2.1. Giao lưu thông tin	26
2.2. Khối lượng thông tin	27
2.3. Chất lượng thông tin	27
2.4. Giá trị của thông tin	29
2.5. Giá thành của thông tin	29
3. Phân loại thông tin	31
4. Sơ lược lịch sử phát triển của kỹ thuật truyền tin	33
4.1. Tiếng nói	33

4.2. Chữ viết	34
4.3. Kỹ thuật ấn loát - Nghề in	36
4.4. Công nghệ thông tin hiện đại	37
5. Các quá trình thông tin	41
5.1. Quá trình thông tin	41
5.2. Thông tin khoa học và thông tin đại chúng	43
5.3. Dây chuyền thông tin tư liệu	45
6. Thông tin học và các khoa học liên quan	51
6.1. Thông tin học	51
6.2. Các khoa học liên quan	53
6.2.1. Lý thuyết thông tin	54
6.2.2. Lý thuyết mã hóa	62
6.2.3. Điều khiển học	63
6.2.4. Ngôn ngữ học	64
6.2.5. Tin học	65
7. Vài nét về lịch sử phát triển của thông tin học	68
8. Nghề thông tin	73
9. Hoạt động thông tin khoa học ở Việt Nam	83
Chương 2. THÔNG TIN VÀ TIẾN BỘ XÃ HỘI	91
1. Vai trò của thông tin	91
1.1. Thông tin là nguồn lực phát triển và là nguồn tài nguyên đặc biệt của mỗi quốc gia	92
1.2. Thông tin là yếu tố quan trọng thúc đẩy phát triển kinh tế và sản xuất	94
1.3. Thông tin giữ vai trò hàng đầu trong sự	96

phát triển của khoa học	
1.4. Thông tin là cơ sở của lãnh đạo và quản lý	97
1.5. Vai trò của thông tin trong văn hoá, giáo dục và đời sống	98
2. Thị trường thông tin và kinh tế thông tin	101
2.1. Bối cảnh mới của nền kinh tế thế giới	101
2.2. Thị trường thông tin	103
2.3. Kinh tế thông tin	110
3. Vài nét về hiện tượng tin học hoá xã hội và xã hội thông tin	113
3.1. Tiến trình phát triển của khoa học và công nghệ	113
3.2. Hiện tượng tin học hoá xã hội	115
3.3. Xã hội thông tin	119
4. Thông tin và các nước đang phát triển	125
5. Bùng nổ thông tin và những biện pháp khắc phục	128
5.1. Hiện tượng bùng nổ thông tin	128
5.2. Hệ quả của hiện tượng bùng nổ thông tin	131
5.3. Những phương hướng và biện pháp khắc phục	132
Chương 3. CÁC LOẠI HÌNH TÀI LIỆU - NGUỒN TIN	137
1. Tài liệu và những đặc trưng cơ bản của tài liệu	137
1.1. Tài liệu là gì?	137
1.2. Những đặc trưng của tài liệu	138
2. Các tài liệu tra cứu	142
3. Tài liệu khoa học kỹ thuật và quy luật phát triển của nó	149

3.1. Tài liệu khoa học kỹ thuật và vai trò của nó	149
3.2. Quy luật phát triển của tài liệu khoa học kỹ thuật	150
4. Nguồn thông tin điện tử – Thông tin số	155
5. Xây dựng vốn tài liệu - Chọn lọc và bổ sung	159
5.1. Chính sách bổ sung	160
5.2. Cách tiếp cận các nguồn tài liệu	162
5.3. Các phương thức và thủ tục bổ sung	164
Chương 4. XỬ LÝ TÀI LIỆU VÀ CÁC SẢN PHẨM THÔNG TIN	167
1. Mô tả thư mục	167
1.1. Khái niệm chung về mô tả thư mục	167
1.2. Các vùng dữ liệu	168
1.3. Phương pháp mô tả thư mục	170
1.4. Quy tắc mô tả thư mục theo tiêu chuẩn quốc tế ISBD	171
1.5. Quy tắc biên mục Anh – Mỹ AACR2	176
1.6. Khổ mẫu MARC và UNIMARC	180
1.7. Chỉ số ISBN và ISSN	192
2. Mô tả nội dung tài liệu	194
2.1. Khái niệm chung về mô tả nội dung	194
2.2. Ngôn ngữ tự nhiên và ngôn ngữ tư liệu	195
2.3. Các hệ thống phân loại	198
2.4. Từ điển từ chuẩn	210
2.5. Phân loại, đánh chỉ số, làm tóm tắt	222
3. Các sản phẩm thông tin	231

3.1. Ấn phẩm thông tin thư mục	231
3.2. Tạp chí tóm tắt	232
3.3. Tổng luận khoa học	234
3.4. Cơ sở dữ liệu	235
Chương 5. LƯU TRỮ VÀ TÌM KIẾM THÔNG TIN	242
1. Lưu trữ thông tin	242
1.1. Nguyên tắc lưu trữ thông tin	242
1.2. Các phương tiện lưu trữ thông tin truyền thống	247
1.3. Các phương tiện lưu trữ thông tin bán tự động	250
1.3.1. Phiếu lỗ mép	250
1.3.2. Phiếu lỗ soi	252
1.4. Các phương tiện lưu trữ thông tin tự động hoá	255
1.4.1. Lưu trữ thông tin trên MTĐT	255
1.4.2. Lưu trữ thông tin trên CD-ROM	261
2. Tìm tin	263
2.1. Khái niệm chung về tìm tin	263
2.2. Phương thức tìm tin	264
2.3. Quá trình tìm tin	269
2.4. Một số dạng tìm tin cơ bản	272
2.4.1. Tìm tin trên các bộ phiếu thủ công	272
2.4.2. Tìm tin trên MTĐT	274
2.4.3. Tìm tin trên CD-ROM	276
2.4.4. Tìm tin on-line	278

3. Mô hình hoá quá trình tìm tin	280
4. Đánh giá hiệu quả tìm tin	283
<i>Chương 6. CÁC ĐƠN VỊ THÔNG TIN VÀ DỊCH VỤ THÔNG TIN</i>	285
1. Các loại hình đơn vị thông tin và mạng lưới thông tin	285
1.1. Các thư viện và cơ quan lưu trữ	286
1.2. Thư viện điện tử	292
1.3. Các trung tâm thông tin tư liệu	296
1.4. Ngân hàng dữ liệu	301
1.5. Mạng thông tin	302
1.6. Mạng thông tin toàn cầu Internet	307
2. Các dịch vụ phổ biến thông tin	315
2.1. Khái niệm chung về phổ biến thông tin	315
2.2. Các dịch vụ cung cấp tài liệu cấp một	318
2.3. Dịch vụ biên dịch tài liệu	323
2.4. Các dịch vụ phổ biến thông tin cấp hai	324
2.5. Dịch vụ tìm tin on-line	326
2.6. Phổ biến thông tin trên CD-ROM	330
2.7. Tìm tin trên Internet	332
2.8. Các dịch vụ phổ biến thông tin cấp ba	334
2.9. Phổ biến thông tin có chọn lọc	335
3. Người dùng tin và nhu cầu thông tin	337
<i>Chương 7. CÁC HỆ THỐNG THÔNG TIN</i>	344
1. Mở đầu	344

2. Quản lý và thông tin trong quản lý	345
2.1. Tổ chức và quản lý	345
2.2. Vấn đề ra quyết định và nhu cầu thông tin	351
3. Các hệ thống thông tin	354
3.1. Khái niệm hệ thống	354
3.2. Hệ thống thông tin	357
3.3. Vai trò của hệ thống thông tin	360
4. Một số hệ thống thông tin	362
4.1. Các hệ thống thông tin tác nghiệp	364
4.2. Các hệ thống thông tin quản lý	369
4.2.1. Hệ thống thông tin báo cáo - IRS	370
4.2.2. Hệ thống thông tin hỗ trợ quyết định - DDS	371
4.2.3. Hệ thống thông tin điều hành - IRS	374
4.3- Các hệ chuyên gia	377
5. Vai trò của hệ thống thông tin trong xã hội hiện đại	380
TÀI LIỆU THAM KHẢO	383

Chương 1

THÔNG TIN, CÁC QUÁ TRÌNH THÔNG TIN VÀ THÔNG TIN HỌC

1. THÔNG TIN

1.1. Khái niệm thông tin. Dữ liệu, Thông tin và tri thức

Ngày nay, trong đời sống hàng ngày, ở đâu ta cũng thấy người ta nói tới từ “thông tin”: thông tin là nguồn lực của sự phát triển; chúng ta đang sống trong thời đại thông tin; một nền công nghiệp thông tin, xã hội thông tin đang hình thành v.v...

Quả thật thông tin (information) là khái niệm cơ bản của khoa học cũng là khái niệm trung tâm của xã hội trong thời đại chúng ta. Mọi quan hệ, mọi hoạt động của con người đều dựa trên một hình thức giao lưu thông tin nào đó. Mọi tri thức đều bắt nguồn bằng một thông tin về những điều đã diễn ra, về những cái người ta đã biết, đã nói, đã làm. Điều đó luôn xác định bản chất và chất lượng của những mối quan hệ của con người. Cùng với sự phát triển của khoa học và công nghệ, ngày nay thông tin trở thành nguồn tài nguyên đặc biệt và nguồn lực phát triển của mỗi quốc gia, là công cụ điều hành sản xuất và quản lý xã hội, là cơ sở của các hoạt động

chuyển giao tri thức, và trở thành đối tượng nghiên cứu của nhiều ngành khoa học.

Vậy thông tin là gì?

Có rất nhiều cách hiểu về thông tin. Thậm chí ngay các từ điển cũng không thể có một định nghĩa thống nhất. Ví dụ: Từ điển Oxford English Dictionary thì cho rằng thông tin là "điều mà người ta đánh giá hoặc nói đến; là tri thức, tin tức"; Từ điển khác thì đơn giản đồng nhất thông tin với kiến thức: "thông tin là điều mà người ta biết" hoặc "thông tin là sự chuyển giao tri thức làm tăng thêm sự hiểu biết của con người" v.v...

Nguyên nhân của sự khác nhau trong việc sử dụng thuật ngữ này chính là do thông tin không thể sờ mó được. Người ta bắt gặp thông tin chỉ trong quá trình hoạt động, thông qua tác động trừu tượng của nó.

Từ La tinh "informatio", gốc của từ hiện đại "information" có hai nghĩa. Một, nó chỉ một hành động rất cụ thể là tạo ra một hình dạng (forme). Hai, tùy theo tình huống, nó có nghĩa là sự truyền đạt một ý tưởng, một khái niệm hay một biểu tượng. Tuy nhiên cùng với sự phát triển của xã hội, khái niệm thông tin cũng phát triển theo.

Ngày nay, bên cạnh thông tin nói, thông tin văn bản, thông tin hình ảnh còn có thông tin số, thông tin đa phương tiện. Tuy nhiên dù thông tin ở dạng nào, bao giờ thông tin cũng có hướng tới đáp ứng yêu cầu về tri thức và nâng cao sự hiểu biết của con người. Vì vậy, theo *nghĩa thông thường*, có thể coi thông tin là tất cả các sự việc, sự kiện, ý tưởng, phán đoán làm tăng thêm sự hiểu biết của con người. Thông tin

hình thành trong quá trình giao tiếp: một người có thể nhận thông tin trực tiếp từ người khác thông qua các phương tiện thông tin đại chúng, từ các ngân hàng dữ liệu hoặc từ tất cả các hiện tượng quan sát được trong môi trường xung quanh.

Ta cũng biết rằng, mọi đối tượng vật chất đều có thuộc tính cơ bản là có khả năng tự phản ánh, và con người nhận biết được nội dung của sự phản ánh đó qua giác quan. Do đó, theo *quan điểm triết học*, thông tin là sự phản ánh của tự nhiên và xã hội (thế giới vật chất) bằng ngôn từ, ký hiệu, hình ảnh v.v...hay nói rộng hơn bằng tất cả các phương tiện tác động lên giác quan của con người.

Sự vật luôn vận động, ở trạng thái bất định và chứa đựng tính ngẫu nhiên. Tăng lượng tin tức về một hiện tượng nào đó cũng là giảm độ chưa biết hoặc độ bất định của nó. Vì vậy theo *quan điểm của lý thuyết thông tin* thì thông tin là sự loại trừ tính bất định của hiện tượng ngẫu nhiên.

Tính trật tự đối lập với cái bất định và ngẫu nhiên là thuộc tính cơ bản của thông tin mà khoa học phát hiện. Theo đó thông tin phản ánh cái xác định, trật tự trong các mối quan hệ của các sự vật và hiện tượng. Với ý nghĩa đó thông tin là lượng đo trật tự nhân tạo chống lại sự hỗn độn của tự nhiên. Chính điều đó giải thích ý nghĩa to lớn của thông tin trong mọi lĩnh vực hoạt động của con người.

Trong đời sống con người, thông tin là một nhu cầu rất cơ bản. Nhu cầu đó không ngừng tăng lên cùng với sự gia tăng các mối quan hệ trong xã hội. Mỗi người sử dụng thông tin lại tạo ra thông tin mới. Các thông tin đó lại được truyền cho người khác trong quá trình thảo luận, truyền đạt mệnh lệnh, trong thư từ và tài liệu hoặc qua các phương tiện truyền

thông khác. Thông tin được tổ chức tuân theo một số quan hệ logic nhất định, trở thành một bộ phận của tri thức, đòi hỏi phải được khai thác và nghiên cứu một cách hệ thống.

Trong hoạt động của con người thông tin được thể hiện qua nhiều hình thức đa dạng và phong phú như: con số, chữ viết, âm thanh, hình ảnh v.v... Thuật ngữ thông tin dùng ở đây không loại trừ các thông tin được truyền bằng ngôn ngữ tự nhiên. Thông tin cũng có thể được ghi và truyền thông qua nghệ thuật, bằng nét mặt và động tác, cử chỉ. Hơn nữa con người còn được cung cấp thông tin dưới dạng mã di truyền. Những hiện tượng này của thông tin thấm vào thế giới vật chất và tinh thần của con người, cùng với sự đa dạng phong phú của nó đã khiến khó có thể đưa ra một định nghĩa thống nhất về thông tin.

Dữ liệu, thông tin và tri thức

Thông tin có nhiều mức độ chất lượng khác nhau. Các số liệu, sự kiện, hình ảnh ban đầu thu thập được qua điều tra, khảo sát là các thông tin nguyên liệu, còn gọi là *dữ liệu (data)*. Thuật ngữ dữ liệu - data có xuất xứ từ chữ Latin *datum*, có nghĩa là "cái đã cho" hay sự kiện, nó có thể có dạng một con số, một sự trình bày hoặc hình ảnh. Dữ liệu có thể có hai dạng: dạng có cấu trúc như các biểu ghi, các cơ sở dữ liệu; dạng phi cấu trúc như các tệp văn bản, dạng này thường chiếm đến 80% lượng dữ liệu của một tổ chức. Đặc trưng cơ bản của dữ liệu là chúng có thể tổ chức, lưu trữ và lưu truyền trong các hệ thống và mạng lưới thông tin.

Khi dữ liệu qua xử lý, phân tích, tổng hợp và được cho là có ý nghĩa cho một đối tượng, một công việc nào đó thì chúng sẽ trở thành *thông tin (information)*. Thông tin là sự phản

ánh về một vật, một hiện tượng, một sự kiện hay quá trình nào đó của tự nhiên và xã hội thông qua khảo sát trực tiếp hoặc lý giải gián tiếp. Các thông tin này nếu tiếp tục được xử lý sẽ tạo ra các sản phẩm thông tin có giá trị cao hơn, còn gọi là thông tin có giá trị gia tăng (value added information). Trong trường hợp này thông tin đã thực sự trở thành hàng hoá. Dữ liệu mô tả sự việc chứ không đánh giá sự việc còn thông tin luôn mang một ý nghĩa xác định và nói chung gồm nhiều giá trị dữ liệu. Nói cách khác thông tin là dữ liệu có ý nghĩa.

Ở mức độ cao hơn nữa là các thông tin quyết định trong quản lý và lãnh đạo - kết quả xử lý của những nhà quản lý có năng lực và kinh nghiệm, các thông tin chứa đựng trong các quy luật, nguyên lý của khoa học - kết quả của những công trình nghiên cứu, thử nghiệm của các nhà khoa học và chuyên môn, v.v... Khi đó thông tin trở thành *tri thức* (*knowledge*). Người ta có thể thu nhận được tri thức bằng tư duy trên các thông tin đã qua xử lý. Chỉ có những thông tin được trí óc của con người tiếp nhận và được xử lý tích cực qua quá trình suy nghĩ, học hỏi để nhận thức mới trở thành tri thức. Tri thức bao gồm tất cả những hiểu biết của con người và được sở hữu bởi con người. Thông tin là "cái của người", còn tri thức là "cái của mình". Có thể nói tri thức là tập hợp các thông tin hữu ích đã được trí tuệ của con người xác nhận và đưa vào sử dụng trong thực tiễn.

Theo tổ chức hợp tác và phát triển kinh tế OECD (Organization for Economic Cooperation and Development) thì: tri thức là toàn bộ kết quả về trí lực của con người sáng tạo ra từ trước đến nay, trong đó tri thức về khoa học, về kỹ thuật, về quản lý là các bộ phận quan trọng nhất.

Trong bài phát biểu với tiêu đề “Về con đường xây dựng kinh tế tri thức ở nước ta” tại Hội thảo Quốc gia “Công nghệ mới vì sự phát triển con người”, do Bộ Khoa học – Công nghệ và Môi trường tổ chức tháng 7 năm 2001 tại Hà Nội, GS. Phan Đình Diệu đã khái quát khái niệm tri thức và các loại tri thức như sau: “Nói chung, tri thức là những hiểu biết của con người. Có nhiều loại hiểu biết ở những mức độ và phạm vi khác nhau: biết *cái gì* là ở mức thông tin, biết *tại sao* là ở mức các tri thức khoa học, biết *làm thế nào* là những hiểu biết về công nghệ, biết *ai* và *với ai* là những tri thức về xã hội, và biết *ở đâu* và *lúc nào* thì nói chung là những hiểu biết về tổ chức và làm quyết định. Mỗi loại tri thức đều có tầm quan trọng của mình. Ngày nay bên cạnh các tri thức về khoa học và công nghệ, các tri thức về xã hội, về tổ chức và quản lý ngày càng phát triển phong phú và có ý nghĩa to lớn, nhiều trường hợp là quyết định, trong việc tạo nên sự giàu có của một nền kinh tế”.

Sự phân chia khái niệm dữ liệu, thông tin và tri thức chỉ là tương đối. Dữ liệu của người này có thể là thông tin đối với người khác; tương tự, thông tin đối với người này có thể là tri thức đối với người khác.

Ngay từ thời cổ đại con người đã biết “Tri thức là sức mạnh”. Ngày nay tri thức đang trực tiếp tạo ra quyền lực, tiền bạc và sức cạnh tranh. Tri thức là thông tin mang dấu ấn của một con người, cho nên khả năng nắm bắt và áp dụng tri thức chuyên môn của các thành viên trong tổ chức đóng vai trò cơ bản để duy trì và tạo ra lợi thế cạnh tranh của tổ chức đó. Điều đó lý giải tại sao nguồn lực con người bao giờ cũng là quan trọng nhất trong bất kỳ hệ thống tổ chức nào.

1.2. Các đặc trưng của một thông tin tốt

Không phải mọi thông tin đều có giá trị như nhau. Thông tin này có thể tốt hơn thông tin kia. Trong công việc hàng ngày bạn có thể tiếp nhận hàng ngàn thông tin khác nhau. Vậy thông tin như thế nào là thông tin tốt, đặc trưng của chúng là gì?

Thứ nhất, thông tin phải **thích hợp (pertinence)**. Điều đó có nghĩa là thông tin phải đáp ứng được với yêu cầu của người sử dụng, thông tin phải giúp người dùng tin giải quyết được vấn đề mà họ đang đặt ra trong công việc của họ.

Thứ hai, thông tin phải **kịp thời (timeliness)**. Điều đó có nghĩa là thông tin phải được cung cấp đúng lúc mà người dùng tin cần. Ví dụ, những thông tin về giá cả một mặt hàng nào đó sẽ mất hết ý nghĩa nếu nó không đến kịp thời, vì giá cả luôn thay đổi từng ngày, thậm chí từng giờ.

Thứ ba, thông tin phải **chính xác (accuracy)**. Tính chính xác là yêu cầu bắt buộc đối với mọi thông tin. Nếu các số liệu trong một bản thống kê mà không chính xác, thì mọi kết luận rút ra từ các số liệu đó có thể dẫn đến sai lầm hoặc có hại.

Ngoài ra thông tin tốt còn là thông tin có tác dụng **giảm bớt tính bất định (reduced uncertainty)**. Và như vậy thông tin đem lại sự ổn định cho sự vật. Thông tin tốt còn là thông tin chứa đựng **yếu tố bất ngờ (element of surprise)**. Điều đó thể hiện tính mới của thông tin.

Sự quan tâm đến hiện tượng thông tin gia tăng đột biến vào thế kỷ XX và ngày nay chúng trở thành đối tượng nghiên cứu của nhiều ngành chuyên môn, trong đó có triết học, vật

lý, sinh học, ngôn ngữ học, thông tin học và tin học, kỹ thuật điện tử và truyền thông, khoa học quản lý và nhiều ngành khoa học xã hội. Về phương diện thương mại, công nghiệp dịch vụ thông tin đã trở thành một nền công nghiệp mới mẻ mang tính toàn cầu. Ngày nay hầu như không một ngành công nghiệp, sản xuất và dịch vụ nào lại không quan tâm đến thông tin. Những quan điểm và hiện tượng khác nhau của lĩnh vực này đã dẫn đến những khái niệm và định nghĩa khác nhau về thông tin.

1.3. Các yếu tố cơ bản trong xử lý thông tin

Con người nhận thông tin thông qua các giác quan: âm thanh qua thính giác; hình ảnh và văn bản qua thị giác; trạng thái, nhiệt độ, cảm xúc qua xúc giác; mùi vị qua khứu giác. Để giải thích và hiểu được các tín hiệu nhận được từ các giác quan, con người phải phát triển và học các hệ thống ngôn ngữ phức hợp, nó bao gồm một "bộ chữ cái" các tín hiệu và các tác nhân kích thích cùng với các quy tắc sử dụng chúng. Điều đó cho phép người ta nhận ra các đối tượng mà họ nhìn thấy, hiểu được các thông báo mà họ đọc hoặc nghe, cảm nhận được các tín hiệu nhận được qua xúc giác và khứu giác. Chẳng hạn, chính nhờ sáng tạo ra *hệ thống ký âm* trong âm nhạc, các nhạc sĩ đã sáng tác ra những bản giao hưởng tuyệt vời thể hiện những cung bậc tình cảm khác nhau và người nghe có am hiểu về âm nhạc có thể cảm thụ được cái hay, cái đẹp của nó; *hệ thống văn tự* của hàng trăm ngôn ngữ nói do con người sáng tạo ra đã giúp con người ghi lại những thông tin, lưu giữ và lưu truyền chúng qua không gian và thời gian,...

Các vật mang thông tin chuyển tải tín hiệu tới người nhận có thể là giấy, sóng điện từ, sóng ánh sáng, sóng âm.

Cho đến trước khi máy tính điện tử ra đời, các tín hiệu truyền đi thông qua các vật mang tin trên là những tín hiệu được lưu trữ và xử lý dưới dạng *tương đồng*, dựa trên công nghệ in, chụp ảnh và điện thoại. Với công nghệ thông tin hiện đại, thông tin được biểu diễn dưới dạng các tín hiệu số nhị phân, dựa trên *kỹ thuật số*. Đó có thể coi là bước chuyển biến mang ý nghĩa lịch sử vào cuối thế kỷ XX trong cách thức mà con người sáng tạo, tiếp cận và sử dụng thông tin. Với kỹ thuật số, máy tính điện tử trở thành công cụ chủ yếu xử lý thông tin. Máy tính điện tử thu nhận dữ liệu, xử lý rất nhanh các dữ liệu đó để biến chúng thành thông tin hữu ích, phục vụ cho nhu cầu thông tin đa dạng của con người.

Việc *ghi thông tin bằng kỹ thuật số*, là một bước tiến rất dài so với *kỹ thuật tương đồng* (analogue) dựa vào tín hiệu cơ học hoặc tín hiệu điện. Lúc đầu người ta dùng băng dục lỗ để lưu trữ dữ liệu dưới dạng số nhị phân. Ngày nay chúng đã được thay thế bằng các phương tiện dựa trên cơ sở công nghệ điện tử và điện quang, đó là các băng từ, đĩa từ, đĩa quang. Trong đó, đặc biệt phải kể đến đĩa mềm ra đời những năm 1970 được sử dụng rất thuận lợi cho các máy tính cá nhân, và đĩa quang ra đời đầu những năm 1980 cho khả năng lưu trữ thông tin đa phương tiện với dung lượng lớn.

Kỹ thuật số đã mở rộng khái niệm truyền thống về tài liệu. Ngày nay tài liệu không chỉ là các đối tượng vật chất ghi cố định các thông tin mà bao gồm cả các vật mang tin đa phương tiện. Nhờ được ghi dưới dạng số hoá, các đối tượng này rất dễ thao tác; chúng có thể chia nhỏ, kết hợp với các phần khác, thay đổi dạng thức, hiển thị dưới nhiều hình thức trên nhiều thiết bị khác nhau.

Việc sử dụng kỹ thuật số để biểu diễn thông tin đã dẫn đến việc ra đời một nguồn thông tin mới bên cạnh các nguồn thông tin truyền thống lưu trữ trong các thư viện hoặc cơ quan lưu trữ: *nguồn thông tin số* (Digital Information Resources) hay còn gọi là *nguồn thông tin điện tử* (Electronic Information Resources). Đó là các cơ sở dữ liệu, cơ sở tri thức, các bản tin điện tử, sách báo và tạp chí điện tử, các Website trên Internet, các CD-ROM,...

2. CÁC THUỘC TÍNH CỦA THÔNG TIN

2.1. Giao lưu thông tin

Thông tin tiềm tàng khắp nơi trong xã hội. Đó là các nguồn thông tin về lao động, đất đai, tài nguyên, môi trường; thông tin về các tổ chức và các hoạt động kinh tế, xã hội; thông tin về khoa học và công nghệ; thông tin về sản xuất, kinh doanh v.v... Nhưng thông tin chỉ có giá trị và ý nghĩa khi nó được truyền đi, phổ biến và được sử dụng. Có thể nói bản chất của thông tin nằm trong sự giao lưu của nó.

Để phân biệt nội dung thông tin cần truyền đi và cách thức chuyển giao thông tin người ta tách nội dung thông tin ra khỏi hình thức biểu diễn nó. Các hình thức biểu diễn thông tin (các ký hiệu, dấu hiệu, hình ảnh...) là hữu hạn nhưng nội dung của thông tin (khái niệm, ý tưởng, sự kiện, tên v.v...) thì vô hạn. Trong trường hợp thông tin có một hình thức biểu diễn, quá trình chuyển giao thông tin chính là quá trình truyền các ký hiệu biểu diễn nó. Những ý tưởng mới sẽ được truyền đi bằng một tổ hợp mới của một số hữu hạn các ký hiệu (chữ cái, chữ số...). Trong đời sống hàng ngày thông tin được biểu diễn bằng ngôn ngữ. Khi đó thông tin được diễn tả

bằng cú pháp và ngữ nghĩa của ngôn ngữ mà người ta sử dụng.

2.2. Khối lượng thông tin

Lý thuyết thông tin xác định số lượng thông tin như sau: Người ta thừa nhận rằng càng nhiều tín hiệu sinh ra từ nguồn tin thì càng có nhiều thông tin được truyền đi. Khi đó thông tin được mô tả bằng sự thống kê và tổ hợp các dấu hiệu phát ra từ nguồn.

Thông báo được chuyển đi bằng cách ghi tín hiệu lên một dạng vật chất trung gian, tức là một cái giá, gọi là vật mang tin. Vật mang tin có thể là giấy, sóng điện từ, băng từ, v.v... Về mặt lý thuyết mỗi vật mang tin đều có khả năng xác định giới hạn số lượng các tín hiệu mà nó có thể chứa đựng trên đơn vị không gian hay thời gian. Lý thuyết thông tin sẽ cho chúng ta biết cách ước lượng giới hạn này và quan trọng hơn là làm thế nào để trình bày các tín hiệu sao cho khi chúng được truyền đi trên các vật mang tin, thông tin ít bị sai lệch nhất.

Các kỹ sư truyền thông có trách nhiệm truyền đi chính xác các tín hiệu nhưng họ không cần quan tâm đến nội dung cũng như chất lượng của thông tin. Rõ ràng là việc truyền đi chính xác một thông tin không chính xác không làm cho thông tin này trở nên "tốt hơn".

2.3. Chất lượng thông tin

Tất cả chúng ta đều mong muốn thông tin phải xác thực và đúng đắn. Nói cách khác thông tin phải phù hợp với thực tế. Độ tin cậy của thông tin được tăng lên nếu nó được kiểm

tra bằng các phương tiện độc lập. Để đạt mục đích sử dụng, thông tin phải được cập nhật một cách đầy đủ và chính xác. Điều đó cho phép người nhận có thể lựa chọn thông tin theo yêu cầu của họ. Nếu thông tin không đầy đủ thì mức độ không chắc chắn cần phải được chỉ ra, hoặc phải tuân theo một quy ước nào đó. Ví dụ, khi nói "trời nắng", ta đã ngầm hiểu là trời nắng tại thời điểm và nơi đang nói. Thông tin phải dễ hiểu, tức là người nhận có thể lĩnh hội được. Và lại, có những quy tắc (của ngôn ngữ, ký hiệu...) cho phép ta bảo đảm khía cạnh này của chất lượng thông tin.

Thông tin chất lượng thấp là những thông tin sai lạc hoặc bị bóp méo (do hành động có chủ tâm của nguồn hoặc do quá trình truyền tin). Nó có thể trái ngược với các thông tin khác. Nó có thể trình bày một cách nghèo nàn hoặc thậm chí khó lĩnh hội đối với người nhận. Điểm đáng chú ý ở đây là nhiều sản phẩm của công nghệ thông tin có mục đích phát hiện và nếu có thể, cải thiện thông tin chất lượng thấp trước khi nó được đưa tới người nhận.

Ngoài những đặc trưng chung của thông tin chất lượng cao, với từng trường hợp sử dụng thông tin cụ thể người ta còn đòi hỏi một số đặc trưng khác. Ví dụ, khi trả lời một câu hỏi, thông tin phải thích hợp và kịp thời; phải ở dạng mà người tiếp nhận có thể xử lý một cách thuận lợi (mô tả, phân loại, lưu trữ, tìm kiếm, cập nhật v.v...); phải có mức độ chi tiết thoả đáng và khi cần thiết phải được bảo vệ. Người ta cũng đánh giá cao nếu thông tin được trình bày hấp dẫn và dễ sử dụng.

2.4. Giá trị của thông tin

Trong một nghiên cứu mới đây, người ta thấy có bốn yếu tố tác động đến chất lượng thông tin và đem lại giá trị cho nó. Đó là: tính chính xác, phạm vi bao quát của nội dung, tính cập nhật và tần số sử dụng. Trong đó quan trọng nhất là nội dung, thứ đến là tính chính xác.

Trên bình diện tổng quát, ta thấy rằng thông tin có giá trị là những thông tin có tính chất riêng biệt và thông tin có tính chất dự báo. Tính chất riêng biệt làm cho thông tin phù hợp với yêu cầu của người sử dụng. Còn tính chất dự báo cho phép người ta có thể lựa chọn một quyết định trong nhiều khả năng cho phép. Có thể nói thêm rằng giá trị nhận thức của thông tin dự báo liên quan mật thiết đến tính đúng đắn của việc lựa chọn quyết định.

Trên một ý nghĩa rộng hơn, giá trị của thông tin nằm trong quyền lực tổ chức của nó. Thông tin phản ánh cái xác định, trật tự trong các mối quan hệ của tổ chức. Thông tin có giá trị cao cho phép người ta có thể làm môi trường tốt lên và có thể ra những quyết định đối phó được với sự thay đổi của hoàn cảnh. Tính chất quyền lực này của thông tin còn nằm trong cách nó có thể tượng trưng cho những kiến trúc vật chất và tinh thần và được phản ánh trong các định nghĩa của từ điển về động từ "thông tin": "Thông tin là sắp xếp, hình thành, tạo thành (tri tuệ và tính cách...) bằng cách truyền đạt kiến thức" (Oxford English Dictionary).

2.5. Giá thành của thông tin

Giá thành của thông tin có thể quy về hai bộ phận chính:

- Thứ nhất là Lao động trí tuệ, bao gồm việc hình thành ra thông tin và xử lý nội dung của nó.

- Thứ hai là Các yếu tố vật chất, đó là các phương tiện xử lý và lưu trữ thông tin, các phương tiện truyền tin...

Đối với các yếu tố vật chất việc định giá thường dễ dàng hơn: các vật mang tin, các phương tiện truyền tin, năng lượng dùng để chuyển tải thông tin v.v... được định giá bằng giá trị của thị trường.

Khi thông tin được lặp lại với số lượng lớn và được ghi trên các phương tiện vật chất như trường hợp của báo chí, sách, băng ghi âm v.v... thì giá của phương tiện vật chất sẽ chi phối giá thành của một đơn vị thông tin. Trong trường hợp như thế, sản phẩm thông tin trở thành hàng hoá theo đúng nghĩa kinh tế của từ này. Ở đây thông tin có thể mua, bán, trao đổi và như thế rõ ràng quyền sở hữu thông tin bị chia sẻ.

Nhưng điều đó không đúng với thông tin thuần khiết. Người sáng tạo ra thông tin có quyền sở hữu chúng và được bảo đảm bằng pháp luật (ví dụ như luật về phát minh sáng chế, luật bí mật nghề nghiệp, luật bản quyền v.v...), trong khi vẫn sẵn sàng cung cấp cho người khác. Thực tế đó, như là một bản chất vốn có của thông tin, làm cho nó khó có thể đối xử như là một sản phẩm hay hàng hoá. Do đó, pháp luật luôn gặp khó khăn trong việc giải thích và thi hành luật bản quyền. Phần nào để tránh vấn đề này, thông tin thuần túy trên thực tế thường được cung cấp tự do không kể đến giá thành của chúng (thông tin đại chúng). Ở một cực khác, khi thông tin có giá trị độc chiếm thì người cung cấp thông tin phải gánh chịu một khoản chi phí để bảo vệ thông tin trong

khi chờ bán nó với một giá cao (các báo cáo đặc biệt, văn bản hội nghị, các dịch vụ tư vấn cá nhân v.v...).

3. PHÂN LOẠI THÔNG TIN

Thông tin rất phong phú và đa dạng. Người ta có thể phân loại thông tin theo nhiều tiêu chí khác nhau.

+ *Theo giá trị và quy mô sử dụng*

- Thông tin chiến lược.
- Thông tin tác nghiệp.
- Thông tin thường thức.

+ *Theo nội dung của thông tin*

- Thông tin khoa học và kỹ thuật: các kết quả nghiên cứu phát minh, các phương pháp, các sản phẩm, các tính chất công nghệ, các tiêu chuẩn, các trang thiết bị,....

- Thông tin kinh tế: tài chính, giá cả, thị trường, quản lý, cạnh tranh.

- Thông tin pháp luật: luật, quy định, quy tắc...

- Thông tin văn hoá và xã hội: giáo dục, y tế, nghệ thuật, thể thao,...

+ *Theo đối tượng sử dụng*

- Thông tin đại chúng: dành cho mọi người.

- Thông tin khoa học: dành cho người dùng tin (khách hàng).

+ Theo mức độ xử lý nội dung

- Thông tin cấp một: thông tin gốc.
- Thông tin cấp hai: thông tin tin hiệu và chỉ dẫn.
- Thông tin cấp ba: tổng hợp các thông tin cấp một.

+ Theo hình thức thể hiện thông tin

- Thông tin nói.
- Thông tin viết.
- Thông tin bằng hình ảnh.
- Thông tin điện tử hay thông tin số.
- Thông tin đa phương tiện (multimedia).

Đề cập tới các loại thông tin ta không thể không đề cập đến các loại hình tài liệu, đó là cái giá vật chất mang tri thức của nhân loại. Mỗi tài liệu có hai đặc trưng chủ yếu: đặc trưng về mặt vật chất và đặc trưng về mặt tri thức.

Đặc trưng về mặt vật chất thể hiện ở chất liệu và các tín hiệu sử dụng, kích thước, trọng lượng, cách trình bày, phương thức sản sinh... của tài liệu.

Đặc trưng về mặt tri thức thể hiện ở nội dung chủ đề, giá trị sử dụng, đối tượng công chúng, mức độ xử lý biên tập, mức độ phổ biến của tài liệu.

Người ta phân loại tài liệu theo phương thức truyền tin và theo mức độ xử lý nội dung tài liệu. Dưới đây là ví dụ về một số loại hình tài liệu minh họa cho các khía cạnh trên:

Tài liệu cấp 1:

Tiếng nói	Chữ viết	Hình ảnh
Hội thảo	Sách	Ảnh
Giảng bài	Tạp chí	Bản vẽ
Gặp gỡ	Luận án	Sơ đồ

Tài liệu cấp 2:

Thư mục
Chỉ dẫn
Tóm tắt

Tài liệu cấp 3:

Tổng kết
Tổng luận
Ngân hàng thông tin

Hình 1. Các loại hình tài liệu

4. SƠ LƯỢC LỊCH SỬ PHÁT TRIỂN CỦA CÁC KỸ THUẬT TRUYỀN TIN

Bản chất của thông tin nằm trong sự giao lưu của nó. Do đó lịch sử phát triển của thông tin gắn liền với lịch sử phát triển của kỹ thuật truyền tin.

4.1. Tiếng nói

Tiếng nói hay ngôn ngữ nói đóng vai trò nền tảng trong truyền thông xã hội. Đó là phương tiện truyền tin thô sơ nhất, nó không lưu giữ được thông tin theo thời gian và khả

Tài liệu cấp 1:

<i>Tiếng nói</i>	<i>Chữ viết</i>	<i>Hình ảnh</i>
Hội thảo	Sách	Ảnh
Giảng bài	Tạp chí	Bản vẽ
Gặp gỡ	Luận án	Sơ đồ

Tài liệu cấp 2:

Thư mục
Chỉ dẫn
Tóm tắt

Tài liệu cấp 3:

Tổng kết
Tổng luận
Ngân hàng thông tin

Hình 1. Các loại hình tài liệu

4. SƠ LƯỢC LỊCH SỬ PHÁT TRIỂN CỦA CÁC KỸ THUẬT TRUYỀN TIN

Bản chất của thông tin nằm trong sự giao lưu của nó. Do đó lịch sử phát triển của thông tin gắn liền với lịch sử phát triển của kỹ thuật truyền tin.

4.1. Tiếng nói

Tiếng nói hay ngôn ngữ nói đóng vai trò nền tảng trong truyền thông xã hội. Đó là phương tiện truyền tin thô sơ nhất, nó không lưu giữ được thông tin theo thời gian và khả

năng truyền rất hạn chế trong không gian. Tuy nhiên loài người đã tồn tại rất lâu nhờ phương thức truyền tin này.

Các nhà ngôn ngữ học đếm thấy có khoảng ba ngàn ngôn ngữ nói ngày nay (bốn ngàn ngôn ngữ khác đã biến mất) và khoảng một vài trăm trong số đó được thực sự ghi lại nhờ chữ viết. Sự phát triển của ngôn ngữ là tiêu chí đánh giá sự phát triển và trình độ văn minh của một bộ tộc. Sự phong phú của tiếng Việt nói lên truyền thống văn hiến lâu đời của người Việt.

Theo cố Thủ tướng Phạm Văn Đồng tiếng Việt của chúng ta rất giàu, rất đẹp. " Chính cái giàu đẹp đó làm nên cái chất, giá trị, bản sắc, tinh hoa của tiếng Việt, kết quả của cả một quá trình và biết bao công sức dồi mài". Để minh hoạ cho nhận định đó, Ông đã lấy ví dụ một câu ca dao:

Hỡi cô tát nước bên đàng

Sao cô múc ánh trăng vàng đổ đi?

và một câu thơ trong truyện Kiều:

Long lanh đáy nước in trời

Thành xây khói biếc non phơ bóng vàng.

rồi bình luận: " Đó là những câu rất hay, là hai hòn ngọc. Một là hòn ngọc hây còn mộc mạc của dân gian, và một hòn ngọc đã qua tay của người thợ thiên tài".

4.2. Chữ viết

Chữ viết hay văn tự, là kỹ thuật ghi lại ngôn ngữ nói, đó là một thứ kỹ thuật cách mạng. Nhờ đó người ta có thể ghi lại được các thông tin.

Lịch sử phát minh ra văn tự diễn ra hai đợt nối tiếp nhau, ứng với hai phương thức văn tự khác nhau về hình thái:

- Văn tự tượng hình: ra đời tại Mesopotamie (một miền đất ở Tây Á giữa hai con sông Tigre và Euphrate) khoảng 4000 năm trước Công nguyên, thoát đầu mang tính thuần túy tượng hình, rồi dần dần mang tính trừu tượng hơn. Trong ngôn ngữ tượng hình, mỗi biểu tượng có thể diễn tả trọn vẹn một khái niệm. Các biểu tượng này trải qua nhiều thay đổi về hình dạng khiến sự tương đồng giữa chúng và các đối tượng mà nó thay thế dần dần mất đi, nhưng ý nghĩa của chúng thì ngày càng rõ ràng hơn. Các biểu tượng này được gọi là mẫu tự. Ngôn ngữ càng phát triển thì mẫu tự ngày càng nhiều. Hệ thống chữ tượng hình của Trung Quốc ngày nay có đến trên 50 000 mẫu tự. Chữ tượng hình ngày nay được sử dụng ở các nước Bắc Phi, Trung Cận Đông, Đông Á, Nam Á, Đông Nam Á... (Việt Nam là trường hợp duy nhất ở Đông Nam Á dùng hệ thống chữ cái La tinh để ghi lại tiếng Việt. Đó là do nhà truyền giáo người Bồ Đào Nha A. de Rhodes sáng tạo ra vào giữa thế kỷ XVII, nhằm phục vụ cho việc truyền giáo. Cuốn từ điển Việt - Bồ - La của A. de Rhodes xuất bản lần đầu tiên vào năm 1651).

- Văn tự chữ cái: Trong quá trình phát triển của ngôn ngữ viết, đến một lúc phương pháp biểu diễn đã thay đổi từ tượng hình sang tượng thanh, mà ở đó âm thanh nói được biểu diễn bằng một bộ chữ cái bao gồm các ký hiệu đồ thị. Các tổ hợp khác nhau của một bộ phận các ký hiệu này có thể thay thế cho những khái niệm phức hợp hơn như các từ, các câu. Văn tự chữ cái bắt nguồn từ người Pheniciens (cư trú ở vùng đất Syri ngày nay) vào khoảng từ 2000 đến 1000 năm trước Công

nguyên, rồi dần hình thành ở Hy Lạp vào những năm 800 đến 600 trước Công nguyên. Hệ thống mẫu tự Hy Lạp là sự khởi đầu của mẫu tự La tinh mà sau này được dùng làm cơ sở chung cho tất cả các hệ thống chữ viết ở phương Tây.

Việc phát minh ra chữ viết mang một tầm vóc xã hội rất to lớn. Ngay sau khi ra đời chữ viết đã được sử dụng làm kỹ thuật truyền thông và do đó đã biến đổi sâu sắc các phương thức giao lưu tư tưởng và truyền bá thông tin. Từ khi có chữ viết mới có nghề chép sử, con người thoát khỏi thời kỳ truyền thuyết; nghề bưu chính ra đời, giúp con người trao đổi thông tin từ xa; nghề báo chí hình thành như một công cụ thông tin đại chúng đầu tiên; giáo dục trở thành hoạt động xã hội có tổ chức, thực hiện chức năng chuyển giao thông tin giữa các thế hệ.

4.3. Kỹ thuật ấn loát - Nghề in

Kỹ thuật ấn loát ra đời vào thế kỷ XV với phát minh ra máy in bằng chữ rời của Gutenberg đã biến tài liệu viết thành sách in với khả năng nhân bản rất lớn.

Kể từ khi cuốn sách đầu tiên được in typo là bộ thánh thi Mayence vào năm 1457 đến năm 1500 đã có gần 20 triệu bản của 35 000 đầu sách được in ra. Sách in đã trở thành công cụ chủ yếu truyền bá các tư tưởng mới mẻ, trở thành một trong những nguyên nhân tạo ra những chuyển biến thời Phục hưng.

Ý nghĩa to lớn của việc phát minh ra máy in là ở chỗ nó đã làm thay đổi chức năng của sách. Từ chỗ chỉ là một công cụ lưu trữ văn bản, sách đã trở thành một công cụ vô song

của truyền thông, rồi cũng từ đó sách trở thành một đối tượng của thương mại.

Có thể nói nghề in là một thành tựu kỹ thuật to lớn, tạo điều kiện cho con người giao lưu thông tin rộng khắp, thúc đẩy mạnh mẽ sự phát triển của xã hội về khoa học kỹ thuật cũng như văn hoá nói chung. Từ nay thông tin có thể ghi lại và phổ biến với khối lượng lớn. Hoạt động nghiên cứu khoa học trở thành một hoạt động xã hội có tổ chức và giao lưu khoa học phát triển trên phạm vi toàn thế giới.

Cùng với việc phát minh ra máy in, việc phát minh ra máy chữ vào thế kỷ XVIII đã góp phần cơ giới hoá việc ghi các tài liệu văn bản. Tiếp theo việc phát minh ra phim và kỹ thuật chụp ảnh vào thế kỷ XIX đã cho phép con người lần đầu tiên có thể ghi lại các thông tin bằng hình ảnh một cách chính xác.

4.4. Công nghệ thông tin hiện đại

Những kết quả nghiên cứu của Rudolf Hertz về sóng điện từ (1887) và việc phát minh ra “diode” nổi tiếng của Ambrose Fleming (1904) cho phép con người sử dụng điện tử vào truyền thông mà không cần dây dẫn. Ứng dụng đầu tiên của điện tử là radio, mà ống chân không là cốt lõi. Kỹ thuật mới này của truyền thông được phổ biến rất nhanh và phát triển mạnh mẽ từ những năm 20 của thế kỷ này.

Song song với phát triển radio, một phương tiện truyền thông mới cũng bắt đầu xuất hiện là vô tuyến truyền hình. Do tính phức tạp của kỹ thuật cũng do số vốn bỏ ra để sản xuất, thực tế phải đợi đến những năm 40 của thế kỷ này vô tuyến truyền hình mới trở thành một phương tiện truyền

thông đại chúng. Phát thanh và truyền hình đã thực sự tạo nên sự bùng nổ truyền thông trong những năm 40 của thế kỷ này.

Giai đoạn tiếp theo, trong cuộc phiêu lưu của điện tử mà lúc đó mới chỉ bắt đầu là ra-đa (Radio Detection and Ranging) và liền sau đó là máy tính điện tử, công cụ đứng đầu trong các kỹ thuật truyền thông của thế kỷ XX.

Những phát minh khoa học của thế kỷ XX này, đặc biệt các thành tựu nghiên cứu về điện tử, về quá trình vận động của các điện tử trong các môi trường vật chất khác nhau đã là nền tảng cho sự ra đời và phát triển nhanh chóng của kỹ thuật điện tử và vi điện tử. Nhờ đó mà con người giải quyết được những vấn đề phức tạp về đo lường, khuếch đại, biểu diễn và biến đổi tín hiệu. Đó là cơ sở của kỹ thuật xử lý thông tin, một ngành kỹ thuật phát triển nhanh như vũ bão trong mấy chục năm gần đây, mà thành tựu vĩ đại nhất là sự phát minh ra máy tính điện tử (MTĐT), yếu tố cơ bản của công nghệ thông tin (CNTT).

Cơ sở công nghệ của CNTT là công nghệ cơ sở nhị phân (Binary digital technology). Công nghệ cơ sở nhị phân hay còn gọi là kỹ thuật số cho phép chuyển các thông tin dưới dạng chữ viết, âm thanh, hình ảnh thành thông tin viết dưới dạng kết hợp hai con số 0 và 1. Sau khi được xử lý (lưu trữ, biến đổi, truyền thông) thông tin này lại được chuyển thành dạng thông tin mà người bình thường có thể tiếp nhận được mà không giảm chất lượng của thông tin ban đầu.

CNTT (Information Technology -IT) là một thuật ngữ tương đối mới và có ý nghĩa rất rộng. Khó có thể đưa ra một định nghĩa chính xác cho khái niệm này. Tuy nhiên theo

nghĩa thường dùng hiện nay, CNTT có thể coi là tập hợp các phương pháp khoa học, các phương tiện và công cụ kỹ thuật hiện đại - chủ yếu là MTĐT và các mạng viễn thông - nhằm cung cấp các giải pháp toàn thể để xử lý, tổ chức, khai thác và sử dụng có hiệu quả các nguồn tài nguyên thông tin phong phú và tiềm tàng trong mọi lĩnh vực hoạt động của con người.

Có thể nói hạt nhân của CNTT là tin học và viễn thông.

Tin học là khoa học và công nghệ xử lý thông tin một cách tự động và hợp lý bằng máy tính điện tử. Công cụ chủ yếu của tin học là máy tính điện tử (phần cứng) và các chương trình máy tính, gồm các chương trình hệ thống và các chương trình chuyên dụng (phần mềm).

Viễn thông là sự truyền chữ viết, âm thanh, hình ảnh hay dữ liệu dưới dạng các tín hiệu điện, điện từ hay các xung điện thông qua các phương tiện truyền tin. Các phương tiện truyền tin bao gồm điện thoại (dùng dây dẫn hoặc cáp quang), radio, truyền hình, sóng cực ngắn và vệ tinh. Truyền dữ liệu, lĩnh vực phát triển nhanh nhất của viễn thông là quá trình truyền dữ liệu dưới dạng số bằng dây dẫn hoặc radio.

Dữ liệu số có thể sản sinh trực tiếp dưới dạng mã nhị phân bởi máy tính điện tử hoặc cũng có thể sản sinh bằng cách mã hoá các tín hiệu âm thanh, hình ảnh, chữ viết. Mạng truyền dữ liệu là mạng thông tin được tạo thành bằng cách nối các nguồn tin với nhau sao cho các dữ liệu có thể lưu thông tự do giữa chúng. Các dữ liệu có thể là các mục thông tin, các nhóm tin hoặc các chương trình máy tính. Ví dụ: một thoả thuận ngân hàng, một bức thư, một cuốn sách, các địa chỉ thư tín, một chương trình máy tính.

Để tham gia mạng lưới viễn thông, người sử dụng cần phải trang bị một terminal dùng để truyền và nhận dữ liệu, hay một máy tính có trang bị modem, một máy in. Mỗi hệ thống viễn thông đều sử dụng các phần mềm để quản lý mạng và thực hiện việc truyền thông tin. Các kiểu kiến trúc khách hàng/dịch vụ (client/server) với mức độ chủ động khác nhau của các trạm đầu cuối là hình thức tổ chức các hệ thống thông tin ngày càng phổ biến trong các lĩnh vực ứng dụng đã mở ra khả năng mới cho người dùng tin (khách hàng) chủ động khai thác thông tin trên mạng.

Các dịch vụ viễn thông chủ yếu là: cung cấp thông tin, thư tín điện tử, truyền dữ liệu và trao đổi các chương trình máy tính.

Trong các thành tựu của CNTT và truyền thông, hai yếu tố mang tính sự kiện, có ý nghĩa quan trọng nhất là sự ra đời và phát triển mạnh mẽ của máy tính cá nhân và Internet.

Sự tiến bộ của công nghệ vi xử lý đã thúc đẩy máy tính phát triển nhanh chóng vượt xa mọi dự đoán của các chuyên gia hàng đầu. Kỹ thuật vi xử lý khởi đầu một cuộc cách mạng trong tin học, đó là sự ra đời của máy vi tính, còn gọi là máy tính cá nhân PC (Personal Computer) vào đầu những năm 1980. Ngày nay một máy tính cá nhân thông thường cũng đã đạt tới tốc độ hàng tỷ phép tính/giây. Các phần mềm chuyên dụng dùng cho máy vi tính ngày càng đa dạng và tinh xảo, cùng với giá thành ngày càng giảm đã làm cho máy vi tính trở thành một công cụ hấp dẫn, ngày càng được sử dụng rộng rãi trong các lĩnh vực hoạt động của con người.

Ngày nay máy tính cá nhân đã trở thành một phần quan trọng trong cuộc sống hàng ngày của nhiều người. Máy tính

cá nhân có thể hỗ trợ con người trong công việc hàng ngày, trong giao tiếp với bạn bè đồng nghiệp; là công cụ tuyệt vời để xử lý, lưu trữ, tìm kiếm thông tin; là công cụ thiết yếu trong quản lý xã hội, điều hành sản xuất kinh doanh; là công cụ hỗ trợ con người trong sáng tạo và hưởng thụ văn hoá, trong học tập và trau dồi tri thức,...

Lợi ích của máy tính tăng lên rất nhiều khi chúng có thể kết nối với các máy tính khác để chia sẻ thông tin và xử lý chúng. Từ đó xuất hiện các mạng cục bộ, mạng diện rộng, đặc biệt là liên mạng thông tin toàn cầu Internet. Internet, một phương tiện truy cập thông tin rẻ tiền và rộng khắp từ các máy tính cá nhân, đã giúp cho việc lưu chuyển thông tin được nhanh hơn và tự do hơn, phá vỡ mọi rào cản giữa các quốc gia, dân tộc và các nền kinh tế. Các hệ thống thông tin và các hệ thống thông tin quản lý ở các cấp độ khác nhau cũng ra đời và phát triển trên nền tảng đó.

Ngày nay tin tức thế giới, thư tín cá nhân, những yêu cầu về khoa học, giáo dục, văn hoá, nghệ thuật và sản xuất, kinh doanh luân chuyển qua mạng thông tin toàn cầu Internet, đã và đang kết nối mọi người ở khắp các miền trên thế giới lại với nhau. Internet không chỉ còn là một phương tiện kỹ thuật mà đã thực sự trở thành một môi trường mới cho mọi hoạt động kinh tế, văn hoá, xã hội.

5. CÁC QUÁ TRÌNH THÔNG TIN

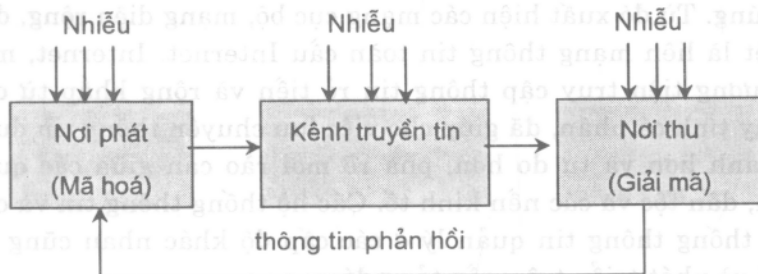
5.1. Quá trình thông tin

Trên quan điểm triết học thông tin là sự phản ánh của thế giới vật chất. Vì vậy có thể coi mọi đối tượng vật chất đều

là những nguồn thông tin (vì vật chất có thuộc tính tự phản ánh - Lênin). Song đó chỉ là thông tin ở dạng tiềm năng.

Để có được thông tin cần phải có đối tượng thu nhận tin. Quá trình tác động qua lại giữa nguồn tin và đối tượng thu nhận tin gọi là quá trình thông tin. Quá trình thông tin được thực hiện qua các phương tiện truyền tin.

Lược đồ chung của quá trình thông tin là:



Hình 2. Quá trình thông tin

Nơi phát hay nguồn tin có thể là một người, một nhóm người hay một tổ chức. Trong trường hợp thông tin truyền đi là có chủ đích, tín hiệu phải được phát đi dưới dạng mã nơi thu có thể hiểu được. Dạng đó gọi là mã (code). Ví dụ các thuật ngữ của tiếng Pháp, các chỉ dẫn giao thông...

Nơi thu hay đích là nơi nhận tín hiệu. Trái với nơi phát, nơi thu thường nhận được các tín hiệu truyền đi từ khắp nơi mà người phát không có chủ đích dành cho họ. Để nhận ra các tín hiệu, nơi thu phải chọn ra các thông tin phù hợp, giải mã các tín hiệu truyền đi để nhận ra các thông tin gốc.

Các kênh là các vật mang tin hay các phương tiện truyền thông. Chúng khác nhau tùy theo cách thức truyền tin. Có rất nhiều phương tiện truyền tin: sóng âm, sóng điện từ, các cử chỉ hành động, văn bản, vệ tinh viễn thông...

Những tín hiệu giữa nơi phát và nơi thu chỉ có thể hiểu được nếu chúng sử dụng cùng một hệ thống tín hiệu (mã). Tuy nhiên trong quá trình thông tin thường xảy ra sự sai lệch vào lúc phát đi hoặc lúc thu nhận tín hiệu. Nơi phát và nơi thu không hiểu được nhau do có những thông tin không nhận được hay do bị nhiễu.

Những trở ngại cho việc chuyển giao thông tin có thể gây ra do tổ chức, kỹ thuật, cũng có thể do tâm lý, nhận thức, do quan hệ giữa người dùng tin và cán bộ thông tin.

Sự chuyển giao thông tin không theo một chiều. Nơi nhận thường tác động lại bằng những thông tin phản hồi. Nghiên cứu phân tích các thông tin phản hồi cho phép đánh giá và điều chỉnh quá trình thông tin để được hiệu quả thông tin tối đa.

5.2. Thông tin khoa học và thông tin đại chúng

Quá trình phát triển của xã hội gắn liền với sự chuyển giao thông tin giữa các thế hệ. Nét mới hiện nay là quy mô phát triển và sự quan tâm to lớn mà người ta dành cho nó.

Các thông tin có thể truyền trực tiếp giữa người và người, có thể truyền qua các phương tiện truyền thông: báo chí, phát thanh, truyền hình. Đó là hoạt động thông tin đại chúng.

Thông tin đại chúng là thông tin dành cho đông đảo mọi thành viên trong xã hội, không phụ thuộc vào trình độ và

công việc của họ. Thông tin đại chúng thông báo những vấn đề, những hiện tượng và sự kiện của đời sống xã hội. Nói chung người ta không kiểm tra và biến đổi những thông tin ấy. Nét đặc trưng của thông tin đại chúng là khối lượng thông tin chuyển giao rất lớn và số lượng đông đảo công chúng sử dụng nó.

Bên cạnh hoạt động thông tin đại chúng thông qua các phương tiện truyền thông, có một loạt các thiết chế xã hội hoạt động với chức năng chuyển giao sự hiểu biết. Đó là các hệ thống giáo dục, các tổ chức nghề nghiệp, hoạt động hành chính, sinh hoạt gia đình... Trong các hoạt động đó ta phải kể đến hoạt động thông tin khoa học, có nhiệm vụ thu thập, xử lý, chuyển giao thông tin từ nguồn tới người sử dụng.

Sự giao lưu thông tin giữa con người trong thực tế bị chi phối bởi không gian và thời gian, nó chỉ có thể bền vững nếu để lại một dấu vết, tức là nó phải được ghi trên một giá mang thông tin, còn gọi là vật mang tin như: sách vở, tranh ảnh, băng từ, đĩa từ, đĩa quang v.v... Tóm lại đó là một tư liệu (document).

Chọn lọc ra từ khối lượng thông tin mang trên những giá đó những tri thức mà người ta yêu cầu, cung cấp cho người dùng tin vào lúc mà họ cần và giữ cho nó không lạc hậu bằng cách bổ sung cập nhật hàng ngày. Đó là mục tiêu của hoạt động thông tin tư liệu hay thông tin khoa học.

Thông tin khoa học, còn gọi là thông tin tư liệu, là dạng thông tin xã hội, hình thành trong quá trình hoạt động thực tiễn của con người và dành cho những đối tượng nhất định, nhằm giúp họ giải quyết những nhiệm vụ cụ thể trong từng lĩnh vực hoạt động của họ. Đối tượng đó được gọi chung là

người dùng tin (users). Họ có thể là các nhà lãnh đạo và quản lý, các nhà nghiên cứu, các giáo sư, các sinh viên, các kỹ sư, các nhà sản xuất và kinh doanh v.v...

5.3. Dây chuyền thông tin tư liệu

Nhiệm vụ của hoạt động thông tin tư liệu không chỉ là cất giữ một kho tài liệu, mà phải chọn lọc, đánh giá phân tích, phân phối những thông tin chính xác cần thiết theo yêu cầu của người dùng tin. Những yêu cầu này thay đổi tùy theo lĩnh vực, tình trạng của tri thức và đối tượng người dùng tin, nhưng chúng có một nét chung là các thông tin nhận được phải thích hợp và kịp thời.

Điều đó đòi hỏi công tác thông tin tư liệu phải thực hiện một loạt các công đoạn có cấu trúc một cách hợp lý mà người ta gọi là dây chuyền thông tin tư liệu (chaine documentaire).

Dây chuyền thông tin tư liệu bao gồm các công đoạn sau đây:

- Chọn lọc và bổ sung.
- Mô tả thư mục.
- Mô tả nội dung.
- Lưu trữ và bảo quản.
- Tìm và phổ biến thông tin.

Tính đa dạng và phức tạp của các công đoạn liên tiếp mà việc xử lý thông tin đặt ra trong dây chuyền thông tin tư liệu nói lên rằng khoa học thông tin là một khoa học đa ngành ở trình độ cao. Dưới đây là một số nét tổng quát về các công đoạn đó.

Chọn lọc và bổ sung

Là bước đầu tiên của dây chuyền tư liệu, chọn lọc và bổ sung cho phép ta xây dựng và nuôi dưỡng vốn tư liệu của một đơn vị thông tin. Nó gồm các bước sau: thăm dò nguồn tài liệu, lựa chọn tài liệu và làm các thủ tục bổ sung tài liệu.

Việc chọn lọc phải dựa trên nhiều nguồn khác nhau: các ấn phẩm thương mại, các tử phiếu và mục lục, thư mục quốc gia và thư mục các loại. Đối với các tài liệu không công bố, người cán bộ tư liệu cần phải nắm bắt được các hoạt động của các cơ quan, tổ chức khoa học và tổ chức mạng lưới trao đổi.

Việc chọn lọc và bổ sung không thể làm tùy tiện mà phải tuân theo một chính sách liên quan chặt chẽ đến lợi ích và mục tiêu của đơn vị thông tin.

Mô tả thư mục

Tiếp theo việc bổ sung tài liệu là những công đoạn giúp ta kiểm tra và tìm được ngay tài liệu khi cần thiết. Đó là mô tả thư mục.

Mục đích của mô tả thư mục là lập một phiếu cho tài liệu, trên đó ghi những đặc trưng bên ngoài của tài liệu: tên tác giả, tên tài liệu, nguồn gốc và dạng của tài liệu, ngôn ngữ, nhà xuất bản, nơi xuất bản, năm xuất bản v.v...

Mô tả thư mục cung cấp cho ta một cách biểu diễn tài liệu duy nhất, không mơ hồ. Nó giúp ta dễ dàng kiểm tra, định vị và tìm kiếm tài liệu.

Để đáp ứng với yêu cầu hợp tác và giao lưu thông tin quốc tế, năm 1960 với sự cố gắng của cơ quan tiêu chuẩn quốc tế ISO, quy tắc mô tả thư mục quốc tế ISBD được biên soạn.

Ngày nay quy tắc này được sử dụng ở nhiều nước tạo điều kiện cho việc trao đổi quốc tế các ấn phẩm thông tin thư mục. Trong các nước nói tiếng Anh người ta sử dụng quy tắc mô tả thư mục Anh - Mỹ AACR, được xây dựng trên cơ sở của ISBD nhưng chi tiết và sâu sắc hơn.

Mô tả nội dung

Việc mô tả nội dung tài liệu, còn gọi là phân tích tài liệu, có nhiệm vụ mô tả những thông tin có trong tài liệu, thể hiện nó bằng một hình thức trình bày mà hệ thống thông tin sử dụng.

Điều gì sẽ xảy ra nếu người ta dùng ngôn ngữ tự nhiên để mô tả nội dung tài liệu. Người ta sẽ không nắm được thực chất nội dung tài liệu do tính mơ hồ, phong phú, đa nghĩa của ngôn ngữ tự nhiên.

Để khắc phục những khó khăn về mặt ngữ nghĩa, người ta dùng ngôn ngữ tư liệu để mô tả nội dung tài liệu. Đó là ngôn ngữ nhân tạo, trong đó mỗi thuật ngữ có một ý nghĩa duy nhất đối với tất cả những ai sử dụng nó.

Tùy theo từng trường hợp, việc mô tả nội dung tài liệu được thực hiện ở các mức độ khác nhau.

Ở mức độ sơ cấp nhất, người ta tiến hành phân loại tài liệu. Ở đây người ta xác định chủ đề chính của tài liệu và thể hiện bằng một thuật ngữ thích hợp của ngôn ngữ tư liệu (khung phân loại). Mục đích của phân loại là sắp xếp thông tin thành một số ít các lớp để tổ chức các bộ phiếu và tổ chức kho, giúp cho việc tìm tài liệu theo nội dung một cách nhanh chóng và chính xác.

Ở mức độ sâu sắc hơn, việc mô tả nội dung tài liệu được thực hiện bằng cách đánh chỉ số, hay còn gọi là định từ khoá tài liệu. Nó bao gồm việc xác định những khái niệm và nội dung mà tài liệu đề cập tới và thể hiện bằng một số thuật ngữ của ngôn ngữ tư liệu. Nhờ đó ta có thể lưu trữ thông tin và trả lời câu hỏi của người dùng tin.

Mức độ cao hơn là cô đọng nội dung tài liệu bằng một bản tóm tắt với độ dài thay đổi, tùy theo trình độ phân tích và giá trị của tài liệu được sử dụng. Tóm tắt có lợi ích kép là dễ ghi nhớ, rút gọn thời gian tra cứu tài liệu và cho phép người dùng tin dễ hình dung ra nội dung chứa trong tài liệu.

Lưu trữ và bảo quản

Sau bước mô tả thư mục, mô tả nội dung, mỗi tài liệu cùng những thông tin mà nó chứa trong đó được biểu diễn bằng một chỉ dẫn (notice) cho phép ta có thể lưu trữ các thông tin chứa trong tài liệu, tức là đưa nó vào các công cụ cất giữ và tìm kiếm của hệ thống. Đó là:

- Các bộ phiếu truyền thống hay mục lục (phương tiện thủ công).
- Các bộ phiếu lỗ mép, phiếu lỗ soi (phương tiện bán tự động).
- Các băng từ, đĩa từ, đĩa quang (phương tiện tự động hóa).

Tài liệu gốc được bảo quản trong kho. Ở đây các tài liệu được sắp xếp vào một vị trí xác định. Tùy theo yêu cầu sử dụng, người ta có thể sắp xếp theo loại hình tài liệu, theo kích cỡ của tài liệu, theo chủ đề (sắp xếp theo hệ thống) hoặc theo thứ tự nhập của tài liệu (sắp xếp theo thời gian).

Việc sắp xếp cho phép biết được tài liệu mà người ta yêu cầu ở đâu. Một hệ thống định vị được gắn trên tài liệu (cote) cho phép ta xác định vị trí của tài liệu trong kho.

Các tài liệu, ít nhất là các tài liệu văn bản, có thể được cất giữ ở trạng thái gốc của nó hay dưới dạng thu gọn (microforme). Việc ghi tài liệu trên các vi phim và vi phiếu ngày càng phổ biến. Lợi ích của nó là tiết kiệm được diện tích kho (thể tích tài liệu giảm đến 90%), trọng lượng không đáng kể, dễ nhân bản và phân phối tài liệu.

Tìm và phổ biến thông tin

Chính nhờ việc lưu trữ thông tin mà người ta có thể tiến hành được việc tìm kiếm thông tin. Tìm tin và hệ quả tiếp theo là phổ biến các thông tin tìm được là nhiệm vụ cơ bản của các dịch vụ thông tin, phục vụ yêu cầu của người dùng tin. Đó cũng chính là lý do tồn tại của các cơ quan thông tin.

Tìm tin hay tra cứu tin là tập hợp các công đoạn có mục đích cung cấp cho người dùng tin những chỉ dẫn và thông tin trả lời cho câu hỏi đột xuất hay thường xuyên của họ.

Các dịch vụ tra cứu tin là dịch vụ cơ bản của các đơn vị thông tin, nó giúp người dùng tin có thể sử dụng tốt nhất vốn tư liệu hiện có nhằm thoả mãn yêu cầu thông tin của họ. Các ấn phẩm thư mục, mục lục, tóm tắt, chỉ dẫn, các cơ sở dữ liệu thư mục, cơ sở dữ liệu on-line và đa phương tiện là các công cụ giúp người dùng tin có thể tiếp cận các nguồn thông tin.

Dù bằng thủ công (trên các bộ phiếu) hay bằng tự động hoá (trên các máy tính điện tử), việc tìm tin có hai dạng cơ bản là tìm tin hồi cố và tìm tin chọn lọc.

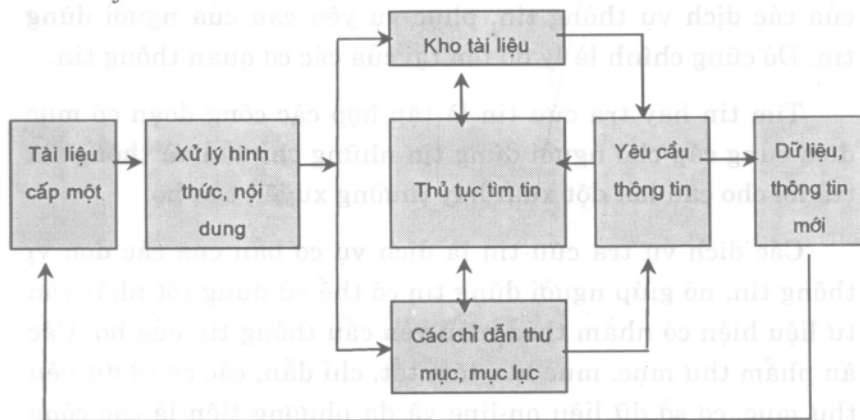
Tìm tin hồi cố là tìm các tài liệu trả lời các câu hỏi về các tài liệu hiện có, dựa trên cơ sở của kho tài liệu.

Tìm tin có chọn lọc là tìm các thông tin cần thiết phục vụ cho các cán bộ chuyên môn theo yêu cầu thường xuyên của họ.

Các sản phẩm thông tin cung cấp cho người dùng tin có thể ở nhiều dạng khác nhau: các tài liệu gốc, các bản tra cứu, các thông tin trích dẫn, các ấn phẩm thông tin, các cơ sở dữ liệu (CSDL) v.v...

Các hình thức phân phối thông tin cũng đa dạng: thường xuyên hay đột xuất, tại chỗ hay tại nhà, cá nhân hay tập thể.

Ta có thể mô tả dây chuyền thông tin tư liệu bằng sơ đồ sau đây:



Hình 3. Dây chuyền thông tin tư liệu

Ngày nay với sự phát triển của công nghệ thông tin, hầu hết các công việc trên đây của dây chuyền thông tin tư liệu đều có thể thực hiện một cách tự động bằng máy tính điện tử.

6. THÔNG TIN HỌC VÀ CÁC KHOA HỌC LIÊN QUAN

6.1. Thông tin học

Thông tin học là một ngành khoa học, có nhiệm vụ nghiên cứu bản chất, cấu trúc và quy luật phát triển của thông tin, cùng với những vấn đề về lý thuyết và phương pháp tổ chức, xử lý, khai thác và sử dụng các nguồn lực thông tin tiềm tàng trong xã hội.

Về mặt lý thuyết, thông tin học có nhiệm vụ tìm ra những quy luật chung nhất của việc sản sinh, thu thập, tổ chức, lưu trữ, tìm kiếm và phổ biến thông tin trong các lĩnh vực hoạt động của con người.

Về mặt ứng dụng, thông tin học có nhiệm vụ tìm ra những phương tiện và phương pháp thích hợp nhất để thực hiện các quá trình thông tin có hiệu quả, để xây dựng các hệ thống giao lưu thông tin hoàn thiện trong các tổ chức, các ngành khoa học và giữa khoa học với sản xuất.

Thư viện học và thông tin học

Là một khoa học mới hình thành từ đầu thế kỷ XX, thông tin học có mối liên hệ mật thiết với Thư viện học và Thư mục học.

Trước khi thông tin học ra đời, một số nhiệm vụ của nó đã được giải quyết trong phạm vi của thư viện học. Những cơ sở lý luận và phương pháp của thư viện học đã tạo tiền đề cho sự

phát triển của thông tin học. Tuy nhiên thư viện học và thông tin học có những nét riêng biệt trong chức năng, nhiệm vụ của mình. Song những nét riêng đó không làm cho chúng bị phân cách mà còn thúc đẩy quá trình tương tác giữa các khoa học này. Có thể nói chúng là những đồng minh tự nhiên của nhau.

Thư viện (Library) là nơi chứa các thông tin đã được ghi lại. Mặc dù xuất xứ của từ Thư viện là từ La tinh Liber có nghĩa là "sách", nhưng ngày nay thư viện là nơi chứa thông tin thuộc nhiều loại hình tài liệu khác nhau như: sách, báo, tạp chí, các đồ biểu, đĩa hát, băng ghi âm, ghi hình...

Thư viện học (Library science hay Librarianship) là một lĩnh vực nghiên cứu nghề nghiệp gắn gũi với thông tin học. Đến giữa những năm 80 của thế kỷ XX, các nhân viên thư viện còn xử lý chủ yếu với các tài liệu văn bản như sách, báo, tạp chí và các tài liệu không ở dạng sách như đĩa hát, băng ghi âm, băng ghi hình bằng các phương pháp truyền thống. Ngày nay hoạt động thư viện đang có những biến đổi sâu sắc dưới tác động của hiện tượng bùng nổ thông tin và dựa trên những thành tựu của công nghệ thông tin hiện đại, các trường thư viện đều đưa thông tin học vào chương trình đào tạo của mình. Tốt nghiệp trường thư viện người học đầu tiên phải làm quen với các nhiệm vụ như đánh giá, xử lý, lưu trữ và tìm kiếm thông tin, với việc phát triển vốn tài liệu, tạo ra các sản phẩm thông tin thư mục, các dịch vụ phục vụ người đọc và người dùng tin... Ngoài ra, các nhân viên thư viện còn phải học sử dụng các phương tiện nghe nhìn, máy tính điện tử và các chương trình ứng dụng. Nhân viên thư viện có quan hệ nhiều hơn với các công việc của cán bộ thông tin.

Thông tin có thể xác định như các tri thức hiện hữu trong bộ não của con người, trong các vật mang tin điện tử, vật mang tin viết hay rộng hơn trong bất kỳ một dạng vật chất nào. Thông tin học là khoa học của những thông tin như thế: bằng cách nào những thông tin ấy có thể được tạo ra, truyền đi, mã hoá, biến đổi, đo lường, sử dụng và đánh giá. Có thể nói thông tin học là một khoa học đa ngành vì nó chứa đựng những ý tưởng và công nghệ của nhiều lĩnh vực khoa học khác như: lý thuyết thông tin, điều khiển học, ngôn ngữ học, tin học, lý thuyết hệ thống, lý thuyết mã hoá và nhiều ngành khoa học xã hội khác. Các cán bộ thông tin xem xét, phân tích tất cả các hiện tượng tác động đến các khía cạnh của thông tin. Họ quan tâm đến việc xác định các vấn đề như: quy luật phát triển và giá trị sử dụng của các loại hình tài liệu, tác động của việc đọc với các nhóm người trong xã hội, vấn đề bản quyền và sở hữu trí tuệ ... Do đó đối với cán bộ thông tin, thư viện chỉ là một trong những khía cạnh của việc lưu trữ và phục vụ thông tin.

Hệ thống thông tin có thể được xây dựng trên cơ sở các cơ sở dữ liệu hay ngân hàng dữ liệu, các cơ quan lưu trữ, các trung tâm phân tích tổng hợp tin, các tổ chức, cơ quan như các trường học, các doanh nghiệp. Các cán bộ thông tin không chỉ làm việc trong các thư viện mà có thể trong các tổ chức khác như các trung tâm y tế, các trung tâm nghiên cứu của các trường đại học, các công ty máy tính, các trung tâm phân tích tổng hợp tin v.v...

6.2. Các khoa học liên quan

Ngoài thư viện học và thư mục học mà thông tin học có nhiều nét chung, thông tin học còn có quan hệ với nhiều

ngành khoa học khác. Sự hình thành của thông tin học được đặt trên cơ sở của một số ngành khoa học như: lý thuyết thông tin, điều khiển học, ngôn ngữ học, lý thuyết mã hoá, tin học v.v...

Dưới đây ta điểm qua vài nét về các ngành khoa học đó.

6.2.1. Lý thuyết thông tin

Một trong những nét đặc trưng của công nghệ thế kỷ XX là sự phát triển và bùng nổ các phương tiện truyền thông mới. Cùng với sự phát triển của các phương tiện truyền và xử lý thông tin, một ngành khoa học lý thuyết hình thành, phát triển và trở thành đối tượng của nhiều nghiên cứu sâu sắc, đó là Lý thuyết thông tin. Đây là ví dụ điển hình về một lý thuyết được khởi xướng đầu tiên bởi một người, kỹ sư điện người Mỹ Claude E. Shanon, mà ý tưởng đầu tiên của ông được đề xuất trong bài báo "A mathematical theory of communication", đăng trên Bell System Technical Journal (1948).

Nội dung bài báo của Shanon nhằm giải quyết nhu cầu về cơ sở lý thuyết của công nghệ truyền thông. Nhu cầu này nảy sinh do độ phức tạp của quá trình truyền tin trên các kênh truyền thông như các mạng lưới điện thoại, điện báo và truyền thanh. Thuật ngữ thông tin ở đây là để chỉ các thông báo được truyền đi như: tiếng nói và âm nhạc được truyền đi bằng điện thoại hoặc truyền thanh, hình ảnh được truyền đi bằng truyền hình, các dữ liệu số hoá trên các mạng máy tính.

Lý thuyết thông tin (Information theory) là lý thuyết liên quan đến các định luật toán học chi phối việc truyền, tiếp nhận và xử lý thông tin. Chính xác hơn, lý thuyết thông tin

đề cập tới các vấn đề về đo số lượng thông tin, biểu diễn thông tin (như vấn đề mã hoá) và khả năng của các hệ thống truyền thông có nhiệm vụ truyền, nhận và xử lý thông tin. Việc mã hoá có thể dùng để chuyển các tín hiệu âm thanh và hình ảnh thành tín hiệu điện, điện từ hoặc dùng để bảo mật thông tin. Lý thuyết thông tin còn được ứng dụng trong những lĩnh vực khác nhau như điều khiển học, ngôn ngữ học, tâm lý học...

Các thành phần của một hệ thống truyền thông

Một hệ thống truyền thông bao gồm các thành phần sau đây:

- Nguồn tin sản sinh ra thông tin hay thông báo sẽ được truyền đi (ví dụ một phát thanh viên).

- Vật truyền như điện thoại, micro, máy tăng âm, máy phát thanh, có nhiệm vụ chuyển thông báo thành tín hiệu điện hay điện từ.

- Các tín hiệu này sẽ được truyền đi qua kênh truyền tin như dây dẫn, khí quyển.

- Máy thu có nhiệm vụ chuyển tín hiệu trở về thông báo ban đầu như máy thu thanh, thu hình, tai nghe điện thoại.

- Nơi nhận, ví dụ như một người nghe đài, nghe điện thoại, một bạn xem truyền hình.

Hai vấn đề quan trọng phải giải quyết trong một hệ thống truyền thông là giảm các nhiễu gây ra bởi hệ thống và tăng khả năng sử dụng của kênh truyền.

Đo thông tin

Ý tưởng cơ bản của lý thuyết thông tin là thông tin có thể xử lý như một đại lượng vật lý, cũng như khối lượng hay năng lượng, nó có thể xác định và đo được bằng đại lượng toán học.

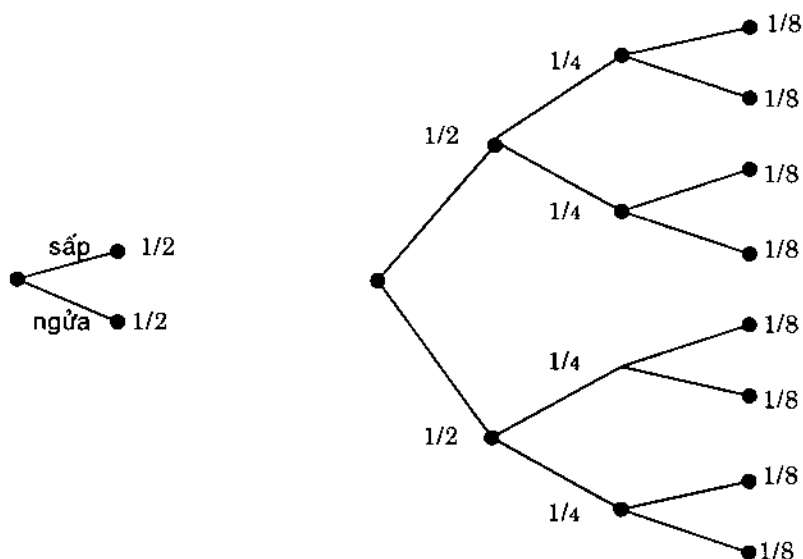
Trước khi xem xét vấn đề thông tin được đo như thế nào, ta cần làm rõ ý nghĩa của "thông tin". Thông thường các thông báo được truyền đi thường mang một ý nghĩa: chúng mô tả hay liên quan tới hiện thực hay các sự kiện có thể nhận thức được. Tuy nhiên không phải bao giờ cũng như vậy. Như trong trường hợp truyền âm nhạc chẳng hạn, ý nghĩa, nếu có sẽ tinh tế hơn trong trường hợp truyền văn bản. Trong một số trường hợp người kỹ sư phải truyền đi một chuỗi các chữ và số hoàn toàn không có nghĩa. Cũng có trường hợp, ý nghĩa không liên quan gì đến vấn đề truyền thông tin.

Xuất phát từ quan điểm truyền tin, khía cạnh ý nghĩa của thông tin là ở chỗ nhận được một thông báo riêng biệt từ một tập hợp các thông báo có thể. Cái mà phải truyền đi là thông báo riêng biệt đã được chọn từ nguồn tin. Như vậy, trong lý thuyết thông tin, thông tin là ý định lựa chọn một thông báo từ một tập hợp các thông báo có thể. Hơn nữa sự lựa chọn này xảy ra với một xác suất nào đó, một vài thông báo có thể xảy ra thường xuyên hơn các thông báo khác.

Sự lựa chọn đơn giản nhất là lựa chọn giữa hai khả năng như nhau, tức là mỗi khả năng có xác suất bằng nhau, bằng $1/2$. Đây là tình huống, ví dụ: khi gieo một đồng tiền, khả năng chọn mặt "sấp" hoặc "ngửa" là như nhau và có xác suất là $p = 1/2$. Lượng thông tin được tạo ra từ cách lựa chọn như thế được coi là một đơn vị của thông tin và đơn vị này được

gọi là bit. Trong ví dụ trên, nếu ký hiệu “sắp” là số 1, “ngủ” là số 0 thì chỉ có một cách chọn để biểu diễn thông báo là 1 hoặc 0. Số 0 và 1 là những chữ số của hệ đếm nhị phân và việc chọn giữa hai ký hiệu đó tương ứng với một đơn vị thông tin nhị phân, đó chính là bit.

Vậy theo quan điểm của lý thuyết thông tin, bit là đơn vị thông tin tương đương với kết quả của sự lựa chọn giữa hai khả năng như nhau, như giữa 1 và 0 trong hệ số nhị phân được sử dụng trong máy tính điện tử. Bit là từ viết tắt của từ “binary digit”, nghĩa là số nhị phân. Nó cũng được sử dụng làm đơn vị của bộ nhớ máy tính, tương ứng với khả năng lưu trữ kết quả của sự lựa chọn giữa hai khả năng.



Hình 4. Xác suất của những lựa chọn có khả năng như nhau

Sự lựa chọn với thông tin một bit có thể mô tả bằng sơ đồ trình bày ở hình 4 (trái). Mỗi đường trên (sấp) hoặc đường dưới (ngửa) có thể lựa chọn như nhau với xác suất bằng 1/2.

Nếu tập hợp các thông báo bao gồm N thông báo có khả năng như nhau thì số lượng thông tin, ký hiệu là I được tính bởi công thức:

$$I = \log_2 N.$$

Rõ ràng với $N = 2$ thì số lượng thông tin sẽ bằng $\log_2 N = \log_2 2 = 1$, cho nên công thức đưa ra phù hợp với đơn vị thông tin đã được lựa chọn ở trên.

Ta biết rằng nếu có N thông báo có khả năng như nhau thì xác suất để có một thông báo là $p = 1/N$, suy ra $N = 1/p$. Do đó để liên kết thông tin với xác suất, Shannon đưa ra công thức sau, đồng nhất với công thức trên:

$$I = \log_2 1/p$$

Để hiểu rõ ý nghĩa của công thức trên ta xét thêm một ví dụ. Giả sử gieo ba lần liên tiếp một đồng tiền thì với cách ký hiệu mặt "sấp" hoặc "ngửa" như trên, 8 kết quả đồng khả năng (hay thông báo) có thể biểu diễn như sau:

000, 001, 010, 100, 011, 101, 110, 111.

Xác suất của mỗi thông báo này là $p = 1/8$. Sự lựa chọn có thể xảy ra ở ba mức, mỗi mức là một bit. Bit thứ nhất tương ứng lần thứ nhất gieo đồng tiền, được lựa chọn một trong hai khả năng sấp (1) hoặc ngửa (0); bit thứ hai ứng với lần gieo thứ hai, được lựa chọn với bộ-đôi thứ nhất hoặc bộ-đôi thứ hai trong 4 khả năng; bit thứ ba ứng với lần gieo thứ ba, được lựa chọn với bộ-bốn thứ nhất hoặc bộ-bốn thứ hai trong 8 khả

năng (hình 4 (phải)). Trong trường hợp này $N = 8$ và lượng thông tin của nó là:

$$I = \log_2 N = \log_2 1/p = \log_2 8 = 3$$

Đó chính là số bit cần thiết để biểu diễn mỗi thông báo nói trên.

Entropy và thông tin

Nếu xác suất không bằng nhau, các thông báo trong tập hợp có lượng thông tin khác nhau kết hợp với chúng. Giả sử xác suất của các thông báo có thể có trong tập hợp là

$$p_1, p_2, \dots, p_N$$

thì lượng thông tin tương ứng kết hợp với chúng tương ứng với

$$\log_2(1/p_1), \log_2(1/p_2), \dots, \log_2(1/p_N).$$

Kỳ vọng toán học hay giá trị trung bình của các giá trị này, ký hiệu là H là:

$$H = p_1 \log_2(1/p_1) + p_2 \log_2(1/p_2) + \dots + p_N \log_2(1/p_N)$$

được gọi là entropy, hay thông tin trung bình của tập hợp các thông báo. Công thức trên còn có thể viết dưới dạng:

$$H = - p_1 \log_2 p_1 - p_2 \log_2 p_2 - \dots - p_N \log_2 p_N$$

Entropy là thuật ngữ mượn của nhiệt động học. Đó là đại lượng đặc trưng cho độ bất định của hệ thống. Trong lý thuyết thông tin, entropy của một tập hợp thông báo là lượng thông tin trung bình của nó. Nói cách khác lượng thông tin thu được về việc xảy ra một sự kiện nào đó được xác định là bằng độ bất định của sự kiện đó trước khi biết nó xảy ra. Điều

đó cũng có nghĩa là lượng tin càng cao khi độ bất ngờ của sự kiện càng lớn.

Chú ý rằng :

$$\log_2 k = \log_2 10 \cdot \log_{10} k = 3,32 \cdot \lg k$$

nên ta có thể tính loga cơ số 2 thông qua loga cơ số 10.

Ví dụ: Giả sử qua nhiều năm quan sát thời tiết tại một thời điểm người ta thu được kết quả sau:

Thời tiết trong ngày 15 tháng 6 (phép thử α_1):

Các kết cục của phép thử	có mưa	không mưa
Xác suất	0,4	0,6

Thời tiết trong ngày 15 tháng 11 (phép thử α_2):

Các kết cục của phép thử	có mưa	có tuyết	không mưa
Xác suất	0,65	0,15	0,2

Entropy tương ứng của hai phép thử này là:

$$H(\alpha_1) = -0,4 \log_2 0,4 - 0,6 \log_2 0,6 = 0,97$$

$$H(\alpha_2) = -0,66 \log_2 0,65 - 0,15 \log_2 0,15 - 0,2 \log_2 0,2 = 1,28$$

Vậy $H(\alpha_2) > H(\alpha_1)$, do đó tại khu vực đang xét thời tiết ngày 15 tháng 11 khó dự báo hơn thời tiết ngày 15 tháng 6.

Entropy nhận giá trị nhỏ nhất bằng 0 khi có một thông báo chắc chắn xảy ra (tức là có xác suất bằng 1) còn tất cả các thông báo khác không xảy ra (có xác suất bằng 0). Về mặt

trực giác ta cũng thấy rằng sẽ không có thông tin trong một thông báo mà ta đã biết chắc nó xảy ra. Ví dụ khi gieo một đồng tiền mà thông báo "xuất hiện sấp hoặc ngửa" thì coi như chẳng có nội dung thông tin nào. Ngược lại entropy nhận giá trị lớn nhất là $\log_2 N$, khi N thông báo trong tập hợp có khả năng như nhau (trong trường hợp này $p=1/N$ và do đó $N=1/p$).

Ví dụ: Nếu thông báo được truyền đi bao gồm các tổ hợp ngẫu nhiên của 26 chữ cái, một khoảng trống và 5 dấu chấm câu, tổng cộng là 32 ký hiệu và giả sử rằng xác suất của mỗi ký hiệu là như nhau thì entropy của nó là $H = \log_2 32 = 5$. Điều đó có nghĩa là để mã hoá mỗi ký hiệu trong 32 ký hiệu nói trên ít nhất phải cần 5 bit: 00000, 00001, 00010, ..., 11111. Hiệu quả của việc truyền và lưu trữ thông tin đòi hỏi phải rút gọn số các bit dùng để mã hoá.

Như vậy với lý thuyết thông tin, Shanon đã chứng minh được rằng thông tin có thể đo được, tức là với bản tin bất kỳ ta có thể xác định được nó chứa bao nhiêu đơn vị tin tức.

Thông tin có thể đo được. Đó là phát minh cũng có ý nghĩa về sự hiểu biết của con người đối với thế giới khách quan như ý nghĩa về khả năng đo được của năng lượng. Người ta đã chế tạo ra các máy để sản sinh và chế biến được năng lượng và giờ đây người ta cũng chế tạo ra các máy để gia công tin tức, đó là máy tính điện tử.

Những phương pháp của lý thuyết thông tin được sử dụng để mã hoá các thông tin ngữ nghĩa, đưa thông tin vào máy tính để bảo quản và thực hiện các quá trình tìm tin, truyền tin.

Một trong những nhiệm vụ chính của lý thuyết thông tin là nghiên cứu và nâng cao hiệu quả các quá trình truyền tin

theo các kênh. Các quá trình này cũng được thông tin học nghiên cứu.

6.2.2. Lý thuyết mã hoá

Lý thuyết mã hoá (Coding theory) là khoa học nghiên cứu về các hệ thống dấu hiệu, cùng với những tính chất, quy luật cơ bản của các hệ thống này và các hình thức mã hoá. Các hệ thống đó có thể là: các ngôn ngữ nhân tạo, các hệ thống đánh moóc-xơ, các hệ thống mã nhị phân dùng trong máy tính điện tử, các ký hiệu mã hoá dùng trong nhiều lĩnh vực khác nhau...

Trong kỹ thuật thông tin liên lạc, quy tắc đối chiếu mỗi bản tin với một tổ hợp nào đó các tín hiệu khác nhau được gọi là mã. Còn chính việc chuyển bản tin thành dãy các tín hiệu khác nhau đó gọi là mã hoá bản tin.

Hệ thống mã chỉ dùng hai tín hiệu sơ cấp khác nhau 1 và 0, ứng với hai trạng thái on - off của các thiết bị điện tử, gọi là hệ mã nhị phân.

Ví dụ: Trong các máy điện báo in chữ người ta dùng hệ mã nhị phân Baudot có cấu trúc như sau: ứng mỗi chữ cái với một dãy gồm 5 tín hiệu sơ cấp 1 và 0. Mỗi dãy tín hiệu là một chỉnh hợp chập 5 của hai phần tử, nên ta có $2^5 = 32$ chỉnh hợp khác nhau. Do đó mã Baudot dưới dạng đơn giản nhất có thể truyền đi được 32 chữ cái khác nhau.

Bài toán tổng quát của việc mã hoá trong kỹ thuật thông tin liên lạc được đặt ra như sau: Giả sử một bản tin được viết bằng một bản chữ cái gồm N chữ cái. Ta phải mã hoá N chữ cái đó bằng mã nhị phân. Vấn đề đặt ra là số cực đại những tín hiệu sơ cấp (tức là những chữ số 0 và 1) cần để ghi một

chữ cái là bao nhiêu. Số cực đại đó càng nhỏ thì mã càng kinh tế vì khi đó số tín hiệu sơ cấp được sử dụng để truyền đi một bản tin là ít nhất.

Bất kỳ hình thức thông tin nào cũng đều biểu diễn bằng những ngôn ngữ, những dấu hiệu nhất định. Do đó thông tin học sử dụng phương pháp của lý thuyết mã hoá để biểu diễn các thông tin ngữ nghĩa trong các hệ thống lưu trữ và tìm kiếm thông tin.

6.2.3. Điều khiển học

Điều khiển học (Cybernetics) là khoa học tổng quát về các quá trình điều khiển, xuất hiện do nhu cầu tự động hoá nền sản xuất hiện đại.

Có thể nói quá trình điều khiển gồm bốn khâu chính sau đây:

- Thông tin nhanh chóng và chính xác về những điều kiện và tình hình ảnh hưởng tới hoạt động của hệ điều khiển.
- Chính lý những tin tức đã nhận được và đề ra chương trình hoạt động tốt nhất (optiman).
- Ra lệnh thực hiện chương trình đã định.
- Kiểm tra sự thực hiện chương trình đó bằng cách tổ chức thông tin kết quả hoạt động trong từng giai đoạn. Trên cơ sở đó điều chỉnh những sai lệch xảy ra.

Theo sơ đồ đó có thể thấy bản chất của quá trình điều khiển là xử lý thông tin. Vì vậy Komogorop (một trong những nhà toán học lỗi lạc của thời đại chúng ta) đã định nghĩa: “Điều khiển học là khoa học về các phương pháp thu thập, bảo quản và chỉnh lý thông tin trong máy móc, trong cơ thể

sinh vật hoặc trong các hệ thống kết hợp máy móc với cơ thể sinh vật”.

Thuật ngữ cybernetics có xuất xứ từ thuật ngữ Hy Lạp cổ kybernetikos (nghệ thuật của người lái tàu), lần đầu tiên được sử dụng bởi nhà toán học người Mỹ vào năm 1948. Theo Norbert Wiener thì điều khiển học là khoa học về truyền thông và điều khiển trong cơ thể sinh vật, trong máy móc và trong các tổ chức. Theo quan điểm của điều khiển học, các hệ thống truyền thông và điều khiển trong các cơ thể sinh vật và trong máy móc được coi là tương tự như nhau. Để hoàn thành được nhiệm vụ đặt ra đối với cơ thể con người hoặc thiết bị máy móc, thông tin liên quan đến các kết quả của các hành động hiện tại phải có khả năng định hướng được cho hành động tương lai. Trong cơ thể con người, bộ não và hệ thống thần kinh hoạt động để kết hợp các thông tin, rồi sử dụng các thông tin đó để xác định các hành động tương lai. Cơ chế kiểm tra để tự điều chỉnh trong máy móc cũng hoạt động tương tự. Ở đây nguyên tắc thông tin phản hồi được coi là yếu tố cơ bản của tự động hoá.

Như vậy ta thấy rằng thông tin học có quan hệ mật thiết với điều khiển học. Chính thông tin học đã sử dụng phương pháp của điều khiển học để tự động hoá quá trình thu thập, xử lý, bảo quản và cung cấp thông tin.

6.2.4. Ngôn ngữ học

Ngôn ngữ học (Linguistics) là khoa học nghiên cứu về ngôn ngữ. Đó là những nghiên cứu về tiếng nói, chữ viết và ngữ pháp của một ngôn ngữ cụ thể, về mối quan hệ giữa các ngôn ngữ hay các đặc trưng phổ biến của các ngôn ngữ. Nó

cũng có thể nghiên cứu các khía cạnh xã hội và tâm lý của truyền thông.

Về mặt lý thuyết, ngôn ngữ học nghiên cứu việc xây dựng các mô hình ngôn ngữ, những vấn đề lý thuyết trong mô tả ngôn ngữ hay giải thích cấu trúc của chúng. Về mặt ứng dụng, ngôn ngữ học sử dụng các kết quả trong nghiên cứu ngôn ngữ vào việc giảng dạy ngôn ngữ, biên soạn từ điển hay chữa bệnh nói ngọng. Một thành tựu của nghiên cứu ứng dụng ngôn ngữ trong thế kỷ XX là sử dụng máy tính trong dịch thuật và nhận biết tiếng nói.

Ngôn ngữ học và thông tin học là hai khoa học có quan hệ mật thiết với nhau. Ngôn ngữ học nghiên cứu ngôn ngữ như một hệ thống truyền thông, còn thông tin học nghiên cứu các quá trình truyền tin mà trong đó ngôn ngữ được coi là công cụ truyền thông đầu tiên. Ở đây thông tin học nghiên cứu sao cho ngôn ngữ trở thành một công cụ biểu diễn và chuyển tải thông tin. Lĩnh vực mà thông tin học nghiên cứu là xử lý ngôn ngữ tự nhiên (Natural language processing - NLP) mà nội dung chủ yếu là: nhận biết tiếng nói; nhận biết lệnh; phân tích và biểu diễn nội dung; giao tiếp hệ thống. Thông tin học còn sử dụng ngôn ngữ tự nhiên để xây dựng ngôn ngữ tư liệu. Đó là ngôn ngữ dùng để mô tả nội dung tài liệu phục vụ cho việc lưu trữ và tìm kiếm thông tin.

6.2.5. Tin học

Tin học (Informatic hay Computer science) là khoa học về sự xử lý thông tin một cách hợp lý và tự động bằng cách sử dụng các thiết bị kỹ thuật hiện đại mà chủ yếu là máy tính điện tử (phần cứng) và các chương trình máy tính (phần mềm).

Đối tượng của tin học là nghiên cứu kiến trúc máy tính và khả năng sử dụng nó để tính toán, xử lý dữ liệu và điều khiển hệ thống. Tin học bao gồm các hoạt động kỹ thuật như thiết kế máy tính, phần cứng và phần mềm để xây dựng hệ thống máy tính. Nó cũng bao gồm cả các hoạt động nghiên cứu lý thuyết và toán học như phân tích và thiết kế thuật toán, phân tích và thiết kế hệ thống, đánh giá độ tin cậy và khả năng của hệ thống.

Những lĩnh vực chủ yếu mà tin học nghiên cứu bao gồm:

- Phần cứng, bao gồm: Kiến trúc máy tính và mạng máy tính.
- Phần mềm, bao gồm: Công nghệ phần mềm, ngôn ngữ lập trình, hệ điều hành, hệ thống thông tin và cơ sở dữ liệu, trí tuệ nhân tạo, đồ họa.
- Lý thuyết, bao gồm: Phương pháp tính, giải tích số, cấu trúc dữ liệu và giải thuật.

Sự phát triển của tin học

Tin học mới trở thành một khoa học độc lập từ những năm 60 của thế kỷ XX, mặc dù máy tính điện tử số đã được phát minh ra từ khoảng 20 năm trước đó. Nguồn gốc của tin học là kỹ thuật điện tử và toán học. Kỹ thuật điện tử cung cấp cơ sở cho việc thiết kế mạch điện. Ý tưởng ở đây là các xung điện khi đưa vào mạch điện có thể được kết hợp để tạo ra đầu ra tùy ý. Toán học là nguồn gốc của một trong những khái niệm cốt yếu cho sự phát triển của máy tính, đó là ý tưởng cho rằng mọi thông tin đều có thể biểu diễn bằng một dãy các chữ số 0 và 1. Đại số Boole phát triển vào thế kỷ XIX cung cấp một hình thức để thiết kế các mạch logic, lý thuyết

máy Turing (1936) của nhà toán học người Anh Alan Turing là một công cụ lý thuyết xử lý một chuỗi các số 0 và 1, rồi nguyên lý máy tính Von Neumann chính là những cơ sở toán học của kiến trúc máy tính hoạt động theo chương trình.

Những khó khăn trong việc viết chương trình bằng ngôn ngữ máy tính đã dẫn đến việc hình thành các ngôn ngữ lập trình bậc cao vào những năm 50, tạo điều kiện cho việc lập trình được nhanh chóng và dễ dàng. Việc gia tăng sử dụng máy tính trong những năm 60 đã thúc đẩy phát triển các hệ điều hành, giúp người sử dụng khai thác máy tính được thuận lợi. Trong những năm 50, các hệ thống thông tin mới chỉ bao gồm các tệp dữ liệu lưu giữ trên băng từ. Việc phát minh ra đĩa từ cho phép truy cập trực tiếp một biểu ghi tùy ý trên đĩa, đã dẫn đến việc hình thành các hệ thống thông tin được thiết kế trên khái niệm cơ sở dữ liệu trong những năm 60, 70 và phát triển các hệ quản trị cơ sở dữ liệu được dùng phổ biến như ngày nay. Từ quản lý tệp đến quản lý cơ sở dữ liệu có thể nói là một bước ngoặt phát triển của tin học. Công nghệ đồ họa bit - map ra đời đầu những năm 80, cùng với sự ra đời màn hình với độ phân giải cao và công nghệ xử lý thông tin đa phương tiện đã dẫn đến sự phát triển bùng nổ của tin học trong những năm gần đây.

Là một tiến bộ khoa học kỹ thuật mũi nhọn của thời đại, tin học cung cấp các phương pháp và công cụ cực kỳ mạnh mẽ giúp con người xử lý, khai thác và sử dụng thông tin, do đó được ứng dụng hiệu quả trong nhiều lĩnh vực hoạt động của con người như: trong nghiên cứu khoa học kỹ thuật, trong điều khiển tự động các quá trình công nghệ, trong điều tra cơ bản, trong thông tin liên lạc và đặc biệt là trong việc tổ chức quản lý sản xuất, quản lý kinh tế xã hội.

Vai trò của tin học trong các đơn vị thông tin không ngừng phát triển, với một nhịp độ cực kỳ nhanh chóng, dẫn đến việc hình thành những hệ thống thông tin tự động hoá, trong đó hình thức hiện đại nhất là các ngân hàng dữ liệu. Ngân hàng dữ liệu hoạt động nhờ một hệ thống chương trình lớn, gọi là hệ quản trị ngân hàng dữ liệu. Nó tiếp thu thông tin mang tới, lưu trữ ở bộ nhớ của máy tính điện tử, bảo đảm việc khai thác thông tin tiện lợi và nhanh chóng cho khách hàng.

Việc ứng dụng tin học trong hoạt động thông tin tư liệu đã hình thành nên một phân ngành của tin học là tin học tư liệu (informatique documentaire).

7. VAI NÉT VỀ LỊCH SỬ PHÁT TRIỂN CỦA THÔNG TIN HỌC

Thông tin học ra đời là sự kế thừa của các ngành thư viện, thư mục, lưu trữ, truyền tin liên lạc... và là sự đáp ứng các yêu cầu thực tiễn nhằm giải quyết vấn đề “bùng nổ thông tin” của xã hội.

Thông tin học phôi thai từ đầu thế kỷ XX và mới được hình thành như một lĩnh vực khoa học độc lập từ những năm 60 của thế kỷ này. Trong quá trình hình thành thông tin học có nhiều quan niệm và tên gọi khác nhau.

Việc lưu truyền thông tin qua thời gian đòi hỏi phải có các phương tiện lưu trữ, gọi chung là các tư liệu. Nguồn gốc của thông tin học là các lĩnh vực của tư liệu học, đó là khoa học về xử lý và cung cấp tư liệu, được hình thành từ đầu thế kỷ XX, song song với sự gia tăng các nghiên cứu thực nghiệm, nó đòi hỏi phải cung cấp các nguồn tin về các đề tài. Năm 1905, nhà bác học người Bỉ Pole Other lần đầu tiên sử dụng thuật ngữ

Tư liệu học (documentation), với quan niệm rằng các tài liệu mà các nhà khoa học sử dụng không chỉ là sách báo, tài liệu ghi chép, hồ sơ lưu trữ mà còn có thể là những hiện vật bảo tàng, các di chỉ văn hoá v.v... Sau này theo ý nghĩa rộng hơn, từ này chỉ các công tác sưu tập, tổ chức, lưu trữ, truy tìm và phổ biến một cách có hệ thống các thông tin và tài liệu chuyên môn, đặc biệt là các tài liệu khoa học kỹ thuật. Tư liệu học phát triển nhằm đáp ứng với sự phát triển của các ấn phẩm định kỳ và các tạp chí, được coi là những phương tiện thông dụng công bố các công trình khoa học. Trong khi sách cần phải phân loại và biên mục thì các ấn phẩm định kỳ đòi hỏi phải đánh chỉ số và tóm tắt để có thể cung cấp cho các nhà nghiên cứu các thông tin cấp một có nguồn gốc xuất bản rất khác nhau.

Trong thời gian chiến tranh thế giới thứ hai (1939-1945), các nhà khoa học tham gia chiến tranh đòi hỏi tìm nhanh chóng và chính xác các thông tin thư mục trong các tạp chí khoa học. Điều đó đã thúc đẩy họ phải đổi mới phương pháp lưu trữ và tìm kiếm thông tin truyền thống, dẫn đến việc sửa đổi các hệ thống phân loại, xây dựng các hệ thống tìm tin. Trong giai đoạn 1940 - 1950, các nhà khoa học đã phát triển kỹ thuật đánh chỉ số, chuẩn hoá các thuật ngữ mô tả thông tin, quan tâm đến việc tóm tắt tài liệu để tạo thuận lợi cho quá trình tìm tin.

Nền tảng của thông tin học dựa trên ba ngành phát triển mạnh sau chiến tranh thế giới thứ hai, đó là: mô hình lý thuyết thông tin của Shannon - Weaver, khái niệm về điều khiển học của Norbert Wiener, những tiến bộ nhanh chóng trong thiết kế và sản xuất máy tính điện tử. Những cách tân này hướng thẳng vào một lĩnh vực nghiên cứu mới, trong đó

hiều ngành có thể hoà đồng với nhau dưới một ý tưởng chung là "thông tin".

Thuật ngữ Thông tin học (Information Science) lần đầu tiên được đưa ra ở Mỹ vào đầu những năm 60 với ý nghĩa là: khoa học thông tin nghiên cứu tính chất và cơ cấu của thông tin, những phương tiện và lực lượng để thực hiện các quá trình thông tin cùng với những kỹ thuật xử lý thông tin với mục đích sử dụng thông tin thích hợp và có hiệu quả nhất. Năm 1963 lần đầu tiên Học viện công nghệ Georgia đưa môn Thông tin học vào chương trình giảng dạy của mình và nhanh chóng phát triển sang một số các trường đại học khác như một lĩnh vực nghiên cứu độc lập bên cạnh các ngành khác như thư viện học, tin học, v.v...

Là một ngành khoa học mới ra đời trong bối cảnh phát triển mạnh mẽ của cuộc cách mạng khoa học và công nghệ, những vấn đề lý luận và phương pháp luận của thông tin học không ngừng phát triển và dần hoàn thiện cùng với quá trình phát triển các ứng dụng của công nghệ thông tin trong hoạt động của mình.

Ngay từ đầu những năm 60, thông tin học đã sớm quan tâm đến việc sử dụng máy tính điện tử để xử lý và quản lý tài liệu. Nhiều nghiên cứu được triển khai như: mô hình hoá việc lưu trữ và tìm kiếm thông tin; các hình thức quan hệ giao tiếp người - máy; hiệu quả của hình thái nội dung và lĩnh hội thông tin; quá trình sản sinh, truyền và biến đổi thông tin; thiết lập các nguyên lý chung để giải thích và dự báo tác động của thông tin.

Sự ra đời của kỹ thuật số vào giữa thế kỷ XX đã tạo ra một bước ngoặt trong việc kiểm soát của con người đối với các

thông tin đã được ghi lại. Năm 1960, lần đầu tiên máy tính được dùng để số hoá tài liệu văn bản, đó là hai tạp chí tóm tắt của Mỹ Index Medicus của Thư viện Y học Quốc gia và Scientific and Technical Aerospace Reports của Cơ quan quản lý hàng không và vũ trụ Quốc gia (NASA) nhằm giảm giá thành và thời gian xuất bản hai tạp chí này. Chúng được coi là những CSDL thư mục đầu tiên và trở thành một nguồn thông tin mới bên cạnh các kho lưu trữ thông tin truyền thống của thư viện và các cơ quan lưu trữ.

Cuộc cách mạng vi xử lý vào giữa những năm 70 tạo cơ sở cho sự ra đời hàng triệu, rồi hàng chục, hàng trăm triệu máy vi tính với năng lực ngày càng cao, giá ngày càng rẻ, thâm nhập khắp mọi nơi trên thế giới, đã thực sự mở ra một chân trời mới cho mỗi cá nhân và xã hội trong việc nắm bắt những thông tin về những sự kiện và ý tưởng mới, tạo khả năng hiện thực cho việc ứng dụng máy tính trong công tác thông tin tư liệu.

Việc ứng dụng máy tính điện tử trong việc xử lý thông tin đã đem lại những hiệu quả to lớn: tập trung thông tin trong những bộ nhớ lớn, các CSDL và ngân hàng dữ liệu (NHDL), tăng nhanh tốc độ ở tất cả các công đoạn xử lý thông tin. Việc tin học hoá hoạt động thông tin - thư viện đã làm thay đổi căn bản phương thức hoạt động truyền thống từ thu thập, xử lý tài liệu đến phục vụ người đọc. Hệ thống các sản phẩm thông tin ngày càng phong phú về loại hình và số lượng, trong đó sự ra đời và phát triển của các CSDL có ý nghĩa bước ngoặt, mở ra kỷ nguyên mới cho hoạt động thông tin - thư viện. Các dịch vụ thông tin cũng trở nên hết sức đa dạng và phong phú nhằm đáp ứng nhu cầu ngày càng tăng của người dùng tin.

Cùng với việc phát triển mạnh mẽ các kỹ thuật vi xử lý, việc sử dụng ngày càng nhiều các vật mang tin điện tử như băng từ, đĩa từ (phát triển từ năm 1962), đĩa mềm (1970), đĩa quang (1980) đã loại bỏ được những khó khăn về sự quá tải của các kho chứa, nhưng đồng thời cũng đặt ra cho người làm công tác thông tin tư liệu những vấn đề về xử lý và phổ biến thông tin, dựa trên những kỹ thuật đặc biệt và dựa trên những kênh thông tin rất đa dạng.

Cuối những năm 80 sang đầu những năm 90, sự phát triển bùng nổ các mạng viễn thông truyền dữ liệu quốc gia và quốc tế trên cơ sở kỹ thuật cáp sợi quang, vệ tinh và vi ba số đã tạo khả năng nối mạng không những giữa các trung tâm thông tin, mà còn nối được đến máy vi tính của từng cá nhân. Xuất hiện viễn cảnh của các siêu “xa lộ thông tin” liên kết hàng trăm triệu người trong từng quốc gia cũng như trong phạm vi khu vực và toàn cầu, mà tiêu biểu là liên mạng thông tin toàn cầu Internet. Sự kết hợp giữa máy tính và viễn thông dẫn đến sự hình thành và phát triển các hệ thống và mạng lưới thông tin tự động hoá, cho phép các trung tâm thông tin và thư viện liên kết với nhau trên mạng máy tính để chia sẻ nguồn lực thông tin. Các dịch vụ tìm tin on-line phát triển, cho phép sử dụng và khai thác các tài nguyên thông tin không phụ thuộc vào khoảng cách địa lý giữa người dùng với cơ quan thông tin, đã mở ra khả năng và xu hướng hình thành một không gian thông tin thống nhất trên phạm vi toàn cầu.

Cho đến những năm 80 hầu hết các lĩnh vực hoạt động của con người từ sản xuất, kinh doanh đến nghiên cứu giáo dục, để hoàn thành nhiệm vụ của mình đều phải nhờ tới sự trợ giúp của các hệ thống thông tin tự động hoá, trong đó

phải kể đến các hệ thống thông tin tác nghiệp, các hệ thống thông tin hỗ trợ quyết định và hệ chuyên gia.

Như vậy trong mấy chục năm qua, hoạt động thông tin tư liệu đã có những biến đổi sâu sắc, toàn diện và phát triển dựa trên những thành tựu của công nghệ thông tin hiện đại. Lịch sử phát triển của thông tin học gắn liền với quá trình đó.

8. NGHỀ THÔNG TIN

Cùng với sự phát triển của xã hội, ngày nay hoạt động thông tin và tư liệu ngày càng phát triển, đòi hỏi phải có những người hoạt động chuyên nghiệp trong lĩnh vực thông tin và tư liệu. Từ đó trong xã hội xuất hiện một nghề mới, nghề thông tin tư liệu, mà ta thường gọi tắt là nghề thông tin.

Nghề thông tin được xác định bởi nhiệm vụ cơ bản mà nó phải hoàn thành, đó là: tìm, xử lý, sản xuất và phân phối thông tin, bao gồm cả việc tăng thêm giá trị cho các thông tin ấy, nhằm thoả mãn nhu cầu thông tin dù được thể hiện ra hay không của các nhóm người dùng tin và cung cấp cho họ những nguồn thông tin hữu ích, những thông tin này nói chung được tạo thành bởi các tài liệu dưới dạng văn bản, âm thanh hoặc hình ảnh.

Nghề thông tin bao gồm nhiều nghề, chúng ngày càng nhiều và khác nhau tùy theo yêu cầu thay đổi của xã hội và những công nghệ được áp dụng. Những nghề này có tên gọi khác nhau. Bên cạnh những người làm tư liệu truyền thống như những thủ thư, những người làm lưu trữ, ngày nay ta thấy có những người chuyên phân tích tổng hợp tin, những

người chuyên trách việc tìm tin, những người quản lý cơ sở dữ liệu, những nhà dự báo chiến lược...

Mỗi nghề trong những nghề này được đặc trưng bởi những hoạt động mà nó phải thực hiện, trong đó ngoài tri thức và kỹ thuật chuyên môn nó đòi hỏi người làm nghề phải có năng lực nhất định.

Việc xác định chính xác các năng lực này là rất cần thiết cho công tác tuyển dụng, đào tạo và nhất là để cho những người làm nghề thông tin chuyên nghiệp biết họ phải chuẩn bị những gì.

Nghề thông tin trước yêu cầu mới

Những thành tựu của cuộc cách mạng khoa học và công nghệ hiện đại trong công nghệ thông tin và truyền thông, công nghệ sinh học, xu hướng toàn cầu hoá, sự ra đời nền kinh tế điện tử, sự có mặt của Internet ở khắp mọi nơi,... đang làm thay đổi mọi lĩnh vực hoạt động của con người trong đó có nghề thông tin. Nghề thông tin đang đứng trước những thách thức sau:

- Lượng kiến thức ghi lại dưới hình thức in ấn truyền thống và bằng các phương tiện khác ngày càng gia tăng. Xuất hiện nhiều loại hình tài liệu mới: các CD-ROM, các cơ sở dữ liệu online, các nguồn thông tin trên mạng, các sách báo điện tử (e-book, e-journal), các thông tin đa phương tiện,... Khối lượng thông tin tư liệu đó tăng nhanh đến mức nếu như không có các phương tiện kỹ thuật và phương pháp xử lý thông tin hiện đại thì không thể nào kiểm soát và sử dụng nổi dòng thông tin và tư liệu khổng lồ hiện có.

- Nhờ công nghệ thông tin và truyền thông, tốc độ xử lý thông tin ngày càng nhanh, khả năng truy nhập tới các nguồn thông tin ngày càng mở rộng và nhu cầu hợp tác trong môi trường thông tin ngày càng phát triển. Công nghệ thông tin đang thực sự mở rộng bốn bức tường của các thư viện truyền thống. Ngày nay khả năng ứng dụng công nghệ thông tin mới trong việc xây dựng các hệ thống thông tin tự động hoá, các thư viện điện tử, thư viện số nhằm mở rộng và nâng cao chất lượng phục vụ và phổ biến thông tin, phát huy tiềm lực của từng đơn vị thông tin, thư viện đồng thời vươn tới sử dụng các nguồn lực của các trung tâm thông tin, các thư viện khác ở trong và ngoài nước.

- Ngày nay máy tính cá nhân đã trở thành một phần quan trọng trong cuộc sống của nhiều người và Internet đã thực sự trở thành một môi trường mới cho nhiều hoạt động kinh tế, văn hoá và xã hội. Các ứng dụng của máy vi tính và Internet ngày càng phát triển và trở nên phổ cập đang mở rộng khả năng tiếp thu tri thức và hưởng thụ văn hoá, nâng cao chất lượng dịch vụ xã hội đối với mọi người, làm thay đổi mạnh mẽ phương thức giao tiếp, lề lối làm việc của con người.

Tất cả các yếu tố trên đòi hỏi người làm thông tin ngày nay phải có kiến thức về các nguồn tin và cách tổ chức thông tin, hiểu biết về công nghệ thông tin và sử dụng công nghệ thông tin có hiệu quả, hiểu biết hơn về nhu cầu của người dùng tin, các hình thức sử dụng thông tin, sao cho có thể tổ chức, truy cập và đáp ứng tối đa yêu cầu thông tin của người sử dụng.

Yêu cầu về năng lực đối với nghề thông tin hiện nay

Năng lực có thể coi là tập hợp các khả năng cần thiết để thực hiện một hoạt động nghề nghiệp và làm chủ những cách hành xử cần thiết. Chúng bao gồm kiến thức chuyên môn, kỹ năng thực hành và cả năng khiếu nữa. Chúng có thể phân chia thành ba nhóm.

Nhóm thứ nhất bao gồm những năng lực liên quan đến các hoạt động đặc thù của công tác thông tin tư liệu. Nó bao gồm những kiến thức và kỹ năng mang tính chuyên môn và nghiệp vụ mà tất cả các nhà hoạt động thông tin tư liệu chuyên nghiệp phải nắm vững, tất nhiên với trình độ khác nhau tùy theo công việc và vị trí trách nhiệm mà họ phải đảm nhận.

Nhóm thứ hai bao gồm những năng lực liên quan đến các lĩnh vực của công nghệ thông tin và truyền thông. Đó là những kiến thức và kỹ năng liên quan đến khả năng kỹ diệu của công nghệ thông tin và truyền thông đang được ứng dụng rộng rãi trong tất cả các công đoạn của dây chuyền thông tin tư liệu.

Nhóm thứ ba liên quan đến năng lực quản lý và điều hành. Đây không chỉ là đòi hỏi đối nghề thông tin mà còn là đòi hỏi đối với nhiều nghề khác trong xã hội hiện đại.

Sau đây ta sẽ trình bày chi tiết hơn các nhóm yêu cầu trên.

1. Những kiến thức và kỹ năng về chuyên môn và nghiệp vụ

Những kiến thức và kỹ năng về chuyên môn và nghiệp vụ bao gồm:

- **Hiểu biết về người dùng tin.** Người dùng tin là đối tượng phục vụ của hoạt động thông tin, là khách hàng của mọi dịch vụ thông tin. Vì vậy người làm thông tin trước hết phải có khả năng phân tích và hiểu nhu cầu thông tin của người dùng tin, biết được tập quán thông tin của họ, đồng thời giúp họ có thói quen thông tin tốt để sử dụng tốt nhất các nguồn thông tin hiện có, đặc biệt làm quen với các phương tiện xử lý thông tin hiện đại.

- **Hiểu biết về vị trí của nghề thông tin.** Nghề thông tin có một vị trí trong xã hội và hoạt động trong một môi trường pháp lý trong mối tương quan với nhiều hoạt động khác của xã hội. Người làm thông tin phải xác định được vị trí, chức năng nhiệm vụ, hệ thống tổ chức của nghề thông tin trong xã hội; hiểu những nguyên tắc cơ bản liên quan đến chính sách thông tin và kinh tế thông tin; có hiểu biết về môi trường pháp lý của nghề thông tin, đồng thời phải biết áp dụng và quảng bá những quy định, pháp luật liên quan đến hoạt động thông tin, đặc biệt luật bản quyền và các luật liên quan đến sở hữu trí tuệ, biết áp dụng các chuẩn quốc gia và quốc tế liên quan đến hoạt động thông tin.

- **Xác định và đánh giá các nguồn thông tin.** Mọi hoạt động thông tin đều phải dựa trên các nguồn tin và tài liệu. Người làm thông tin phải biết phát hiện, đánh giá và đem lại hiệu lực cho các thông tin, các tài liệu và các nguồn của chúng. Cụ thể là: Biết sử dụng các phương tiện tra cứu thủ công cũng như hiện đại như các mục lục, các thư mục, các cơ sở dữ liệu, các máy tìm kiếm để phát hiện và tìm nguồn thông tin và tài liệu; Nắm vững các tiêu chuẩn lựa chọn tài liệu; Biết kiểm tra chất lượng của các nguồn, các tài liệu và thông tin theo những dữ liệu chỉ dẫn; Nắm vững các công cụ

cũng như các phương thức để có được các thông tin và tài liệu đó. Đó là những công việc liên quan đến khâu chọn lọc và bổ sung tài liệu.

- ***Quản lý vốn tài liệu và kho.*** Để đáp ứng yêu cầu thông tin của người dung tin, người làm thông tin phải luôn nắm vững và đánh giá được số lượng, chất lượng, cơ cấu và giá trị sử dụng vốn tài liệu của đơn vị; Xác định được những khiếm khuyết của vốn tài liệu căn cứ vào mức độ thoả mãn yêu cầu thông tin của người dung tin; Xác định và áp dụng những tiêu chuẩn để lựa chọn hay loại bỏ tài liệu; Biết tổ chức một hệ thống sắp xếp tài liệu cho phép định vị các tài liệu trong kho sao cho dễ dàng tìm kiếm tài liệu đó; Có những hiểu biết về nguyên tắc và kỹ thuật bảo quản đối với các loại tài liệu khác nhau (tài liệu truyền thống, tài liệu điện tử, tài liệu nghe nhìn...).

- ***Phân tích và trình bày thông tin.*** Người làm thông tin phải biết phân tích tài liệu rút ra những thông tin và trình bày những thông tin đó dưới một dạng thích hợp. Đó chính là những công việc liên quan đến việc xử lý hình thức và nội dung tài liệu. Cụ thể là: Xác định các dữ liệu nêu lên các đặc trưng hình thức của tài liệu và ghi chúng trên các phiếu mô tả theo các quy định và tiêu chuẩn đã được xác lập; Phân loại tài liệu theo một khung phân loại đã lựa chọn; Xác định chủ đề chính của tài liệu; Lập một bản chỉ mục (index) cho tài liệu bằng một tập hợp các từ khoá hoặc từ chuẩn nêu lên nội dung và những khái niệm mà tài liệu đề cập tới; Có đọng nội dung tài liệu bằng một bản tóm tắt. ở trình độ cao hơn người cán bộ thông tin còn phải biết biên tập các bản tin, các bài tổng quan tổng luận về một vấn đề nào đó.

- **Lưu trữ thông tin và tổ chức bộ máy tra cứu.** Đây là công việc nhằm tổ chức và cấu trúc các dữ liệu liên quan đến việc mô tả các tài liệu của đơn vị thông tin để đưa chúng vào bộ nhớ của hệ thống. ở đây người làm thông tin phải biết xây dựng các hệ thống mục lục, biên soạn các thư mục (bộ máy tra cứu thủ công) hoặc xây dựng các cơ sở dữ liệu (bộ máy tra cứu hiện đại). Đối với cơ sở dữ liệu, phải biết xây dựng cấu trúc cơ sở dữ liệu, thiết lập phiếu nhập tin, xác định các dấu hiệu tìm tin, tạo khuôn dạng trình bày cho đầu ra của dữ liệu trên màn hình hoặc trên máy in.

- **Tìm tin.** Tìm tin là thuật ngữ chung để chỉ việc tìm tài liệu hay nguồn của tài liệu, cũng như những thông tin về dữ liệu và sự kiện mà tài liệu đó cung cấp. Đó là nhiệm vụ cơ bản của hoạt động thông tin nhằm đáp ứng yêu cầu của người dùng tin. Người cán bộ thông tin phải biết xử lý các yêu cầu tin và đưa ra một chiến lược tìm tin thích hợp, xác định tất cả các nguồn tin có thể để trả lời yêu cầu tin của người dùng tin. Bên cạnh việc sử dụng thành thạo các phương tiện tra cứu thủ công, người làm thông tin ngày nay phải biết sử dụng các phương tiện tra cứu hiện đại: tìm tin trên các cơ sở dữ liệu hoặc trên mạng, đặc biệt là mạng thông tin toàn cầu Internet, biết khai thác thông tin trong các thư viện điện tử - thư viện số. Ngoài ra còn có khả năng tổng hợp, bao gói các kết quả tìm và chuyển chúng tới người sử dụng, đồng thời đánh giá mức độ phù hợp và giá trị sử dụng của các thông tin tùy theo yêu cầu của người người dùng tin.

- **Khai thác và phổ biến thông tin.** Các thông tin đã được xử lý cần được đưa vào khai thác và tạo thuận lợi cho việc sử dụng chúng thông qua việc cung cấp các sản phẩm và dịch vụ thông tin. Người cán bộ thông tin không chỉ biết phổ

biến thông tin thông qua các dịch vụ thông tin thông thường như dịch vụ thông tin hỏi đáp, dịch vụ cung cấp tài liệu, dịch vụ cung cấp thông tin theo yêu cầu (phổ biến thông tin có chọn lọc) mà còn phải biết triển khai các dịch vụ thông tin trên mạng dựa trên công nghệ thông tin hiện đại. Ngoài ra còn phải biết định hướng người dùng tin tới những sản phẩm và dịch vụ thông tin tốt nhất và giúp họ nhận được chúng dưới dạng thích hợp, phát triển sự hợp tác với các tổ chức trong và ngoài nước trong chiến lược tiếp cận và phân phối thông tin.

2. Những kiến thức và kỹ năng về công nghệ thông tin và truyền thông

Những kiến thức và kỹ năng về công nghệ thông tin và truyền thông bao gồm:

- ***Kỹ năng tin học.*** Ngoài kiến thức và kỹ năng cơ bản về tin học, sử dụng máy tính để làm các công việc văn phòng, người làm thông tin ngày nay phải biết sử dụng các phần mềm tư liệu để lưu trữ, tìm kiếm tài liệu và tạo ra các sản phẩm thông tin tư mục, biết sử dụng một phần mềm tích hợp để quản trị một thư viện điện tử. Có những hiểu biết cần thiết về các nguồn tài liệu điện tử, nắm được kỹ thuật số hoá các tài liệu, xử lý các thông tin dưới dạng âm thanh và hình ảnh, các thông tin đa phương tiện (multimedia). Ở trình độ cao hơn, người cán bộ chuyên trách cần phải biết cài đặt và bảo trì hệ thống, biết sử dụng ngôn ngữ lập trình để viết các chương trình cho các ứng dụng đơn giản.

- ***Kiến thức về truyền thông.*** Trong bối cảnh bùng nổ công nghệ thông tin và truyền thông, người cán bộ thông tin ngày nay phải có kiến thức cơ bản về mạng thông tin máy

tính, biết quản lý và khai thác một mạng cục bộ, biết sử dụng các dịch vụ tìm tin trực tuyến, các dịch vụ thông tin chủ yếu trên Internet như thư điện tử, truyền tệp, World Wide Web, ... ở trình độ cao hơn, người cán bộ chuyên trách phải có khả năng tạo lập các trang Web bằng ngôn ngữ đánh dấu, xây dựng các Website.

- **Khả năng về ngoại ngữ.** Trong môi trường giao lưu thông tin toàn cầu hiện nay, để làm tốt công việc của mình, dù ở cương vị nào, người cán bộ thông tin cũng phải phải thực hành tốt ít nhất một ngoại ngữ, đặc biệt là tiếng Anh, để sử dụng trong giao tiếp, trong lựa chọn, tìm kiếm và xử lý tài liệu, cũng như trong các giao tác với các chương trình máy tính.

3. Năng lực quản lý điều hành.

Năng lực quản lý và điều hành của người cán bộ thông tin ngày nay, nhất là những cán bộ phụ trách, thể hiện các mặt sau đây:

- **Quản lý chuyên môn và nghiệp vụ.** Có khả năng đánh giá mặt mạnh mặt yếu của các hoạt động chuyên môn nghiệp vụ, của các sản phẩm và dịch vụ thông tin và đưa ra những giải pháp để cải thiện chất lượng của chúng; Đưa ra được một hệ thống chỉ tiêu để đánh giá chất lượng của các công việc chuyên môn, các sản phẩm và dịch vụ thông tin.

- **Marketing.** Ngày nay thông tin đã trở thành hàng hoá, các hoạt động thông tin ngày càng mang tính chất dịch vụ và kinh tế. Người cán bộ thông tin phải biết thiết lập và duy trì mối quan hệ với khách hàng và đối tác để quảng bá và phát triển các sản phẩm và dịch vụ thông tin; Có khả năng phân tích yêu cầu và thị trường, đánh giá thực trạng của hoạt động

và môi trường cạnh tranh và đưa ra những giải pháp nhằm chiếm lĩnh thị trường thông tin bằng những sản phẩm và dịch vụ của mình.

- **Quản lý kế hoạch và tài chính.** Có khả năng triển khai và kiểm tra việc thực hiện một kế hoạch hoặc một dự án; Có những hiểu biết về công tác tài chính và kế toán để quản lý tốt ngân sách và đem lại những lợi ích kinh tế cho các sản phẩm và dịch vụ thông tin.

- **Quản lý cơ sở vật chất và trang thiết bị.** Ngày nay các trang thiết bị sử dụng trong hoạt động thông tin ngày càng hiện đại. Người cán bộ thông tin phải có khả năng đổi mới trang thiết bị, sắp xếp và cải thiện môi trường làm việc; lựa chọn và sử dụng có hiệu quả các thiết bị xử lý thông tin hiện đại: máy tính, máy in, máy quét, các thiết bị đọc, thiết bị mạng, các phần mềm hệ thống và phần mềm chuyên dụng,...

- **Quản lý và phát triển nguồn lực con người.** Trong hoạt động thông tin yếu tố con người bao giờ cũng giữ vị trí quan trọng hàng đầu. Do đó cần biết quản lý và phát huy năng lực chuyên môn của các cán bộ nghiệp vụ; Sử dụng đúng người đúng việc và chuyên môn hoá đối với từng vị trí công việc; Thường xuyên tổ chức bồi dưỡng chuyên môn nghiệp vụ, cập nhật kiến thức cho cán bộ bằng nhiều hình thức đào tạo linh hoạt và có hiệu quả, đồng thời khuyến khích việc tự học.

Chúng ta đang chứng kiến sự thay đổi về chất trong nghề thông tin và hoạt động thông tin đang có những đóng góp to lớn cho sự phát triển chung của xã hội. Việc nâng cao năng lực của những người làm thông tin chuyên nghiệp là đòi hỏi bức thiết hiện nay trước yêu cầu của thực tiễn và trước yêu cầu phát triển của chính nghề thông tin.

9. HOẠT ĐỘNG THÔNG TIN KHOA HỌC Ở VIỆT NAM

Lĩnh vực hoạt động thông tin ở Việt Nam được nhà nước quan tâm từ năm 1960, ngay sau khi Ủy ban Khoa học Nhà nước được thành lập. Phòng thông tin Khoa học của Ủy ban Khoa học Nhà nước thành lập ngày 18/8/1961 có thể coi là cơ quan thông tin đầu tiên của Việt Nam. Nhiệm vụ chủ yếu của phòng lúc bấy giờ là sưu tầm, nghiên cứu, chọn lọc và phổ biến những kinh nghiệm, thành tựu và tình hình phát triển khoa học và kỹ thuật trong và ngoài nước nhằm phục vụ sản xuất, nghiên cứu khoa học và giảng dạy. Phòng đã phối hợp với các ngành, xuất bản hàng tháng 13 bản tin khoa học phục vụ các ngành như cơ khí, luyện kim, nông nghiệp, y dược v.v..., biên tập các tài liệu tham khảo phục vụ lãnh đạo. Một nhiệm vụ quan trọng là bước đầu tổ chức xây dựng mạng lưới các cơ quan thông tin từ các bộ, ngành đến các viện, trường và các tỉnh thành phố, đồng thời nghiên cứu nghiệp vụ và tổ chức đào tạo cán bộ thông tin.

Hội nghị Thông tin Khoa học Kỹ thuật (KHKT) toàn quốc lần thứ nhất tổ chức tại Hà Nội ngày 4/3/1971 và Nghị quyết 89-CP của Hội đồng Chính phủ ngày 4/5/1972 về "Tăng cường công tác Thông tin KHKT" cùng với việc thành lập Viện Thông tin Khoa học Kỹ thuật Trung ương ngày 4/10/1972 có thể coi là những sự kiện lịch sử đánh dấu một bước ngoặt phát triển đối với ngành thông tin khoa học ở Việt Nam. Nhiệm vụ của Viện Thông tin KHKT Trung ương lúc bấy giờ là phục vụ thông tin cho lãnh đạo và các chương trình phát triển kinh tế và khoa học kỹ thuật, đồng thời chỉ đạo xây dựng mạng lưới thông tin quốc gia. Viện đã biên soạn các bản tin định kỳ về khoa học cơ bản, thông tin thư mục, tạp chí tóm tắt các bài báo khoa học kỹ thuật Việt Nam phục vụ tra

cứu và bản tin chọn lọc tiếng Anh " Vietnamese Scientific and Technical Abstracts" làm tài liệu trao đổi quốc tế.

Năm 1975 Viện Thông tin Khoa học Xã hội Việt Nam được thành lập trên cơ sở sát nhập Thư viện Khoa học Xã hội (thành lập năm 1968, tách ra từ Thư viện Khoa học Trung ương (1960)) và Ban Thông tin Khoa học Xã hội thuộc ủy ban Khoa học Xã hội Việt Nam (thành lập năm 1973) có nhiệm vụ quản lý, khai thác thông tin và tư liệu trên lĩnh vực khoa học xã hội và nhân văn. Cũng trong giai đoạn này, từ năm 1972 đến 1977 hơn 30 cơ quan thông tin đầu ngành được thành lập. Các cơ quan thông tin này đã có nhiều cố gắng đáp ứng những nhu cầu thông tin về khoa học - kỹ thuật, kinh tế - xã hội bằng các bản tin định kỳ, các thông tin chuyên đề, hội nghị, triển lãm... Nguồn tin được khai thác chủ yếu từ kho sách báo và tạp chí của Thư viện KHKT trung ương, Thư viện Khoa học Xã hội, Thư viện Quốc gia và Thư viện các ngành.

Sau Hội nghị Thông tin KHKT lần thứ hai (tháng 3-1977), mạng lưới Thông tin KHKT đã được mở rộng và củng cố từng bước trong toàn quốc. Công tác thông tin KHKT đã trở thành một ngành hoạt động xã hội và đã góp phần tích cực phục vụ lãnh đạo, quản lý, nghiên cứu và sản xuất. Đến năm 1985 số lượng các cơ quan thông tin đã lên tới ngót 250 đơn vị, bao quát hầu hết các ngành kinh tế, các lĩnh vực khoa học và kỹ thuật ở cả trung ương và địa phương. Đã hình thành một hệ thống ấn phẩm thông tin với hơn 200 đầu tên ấn phẩm định kỳ cùng với các loại thông tin chuyên đề tổng hợp, tin chọn lọc, thông tin thư mục, tạp chí tóm tắt, thông báo kết quả nghiên cứu, nông lịch, v.v...

Từ cuối năm 1986, bắt đầu thời kỳ đổi mới với chính sách mở cửa, cơ chế quản lý quan liêu bao cấp đang được thay thế

bằng cơ chế kinh tế thị trường gồm nhiều thành phần kinh tế theo định hướng XHCN. Công tác thông tin phải được đổi mới sao cho thông tin khoa học và công nghệ thực sự trở thành một nguồn lực phát triển kinh tế xã hội. Bước ngoặt đổi mới tổ chức là việc thành lập Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia ngày 24/9/1990, trên cơ sở hợp nhất Viện Thông tin KHKT Trung ương và Thư viện KHKT Trung ương (thành lập từ năm 1960). Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia là cơ quan đầu ngành về thông tin tư liệu khoa học và công nghệ trong mạng lưới thông tin tư liệu Việt Nam. Ngoài việc tạo lập một hệ thống sản phẩm và dịch vụ phong phú phục vụ người dùng tin góp phần thúc đẩy sự phát triển của đất nước, trung tâm còn đi đầu trong việc nghiên cứu, đào tạo, triển khai áp dụng công nghệ thông tin mới và thực hiện chức năng quản lý nhà nước đối với hệ thống thông tin quốc gia. Đến nay hệ thống này đã bao quát tất cả các ngành kinh tế quốc dân, các lĩnh vực khoa học và công nghệ từ trung ương đến các tỉnh, thành phố trong cả nước, bao gồm:

- Trung tâm Thông tin Khoa học và Công nghệ Quốc gia - cơ quan đầu mối trung tâm của hệ thống.
- 2 cơ quan thông tin - tư liệu chuyên dạng tài liệu, sáng chế và tiêu chuẩn của Bộ Khoa học và Công nghệ.
- 42 cơ quan thông tin khoa học và công nghệ của các Bộ và ngành.
- 64 cơ quan thông tin khoa học và công nghệ cấp tỉnh và thành phố trực thuộc trung ương.

- Hơn 400 tổ chức thông tin thư viện cơ sở thuộc các viện nghiên cứu, xí nghiệp, bệnh viện, trường đại học và cao đẳng.

Ngoài ra tham gia vào hệ thống chuyển giao thông tin, phổ biến tri thức còn phải kể đến mạng lưới gồm 64 thư viện tỉnh, thành phố trực thuộc trung ương, trên 480 thư viện cấp huyện và mạng lưới gần 90 kho lưu trữ ở trung ương và địa phương.

Trong những năm qua các cơ quan thông tin tư liệu đã xây dựng được vốn thông tin khoa học và công nghệ khá phong phú, trong đó có: hơn 3 triệu đầu tên sách, 6 nghìn tên tạp chí, 25 triệu mô tả sáng chế phát minh, hơn 200 nghìn tiêu chuẩn, 50 nghìn catalog công nghiệp, 4000 bộ báo cáo địa chất, 4500 báo cáo lâm nghiệp, 20 nghìn báo cáo kết quả nghiên cứu khoa học và luận án tiến sĩ, phó tiến sĩ, hàng triệu tài liệu giới thiệu thiết bị và sản phẩm.

Các cơ quan thông tin đã tạo lập được một hệ thống các sản phẩm và dịch vụ thông tin khá phong phú với gần 300 ấn phẩm thông tin định kỳ, trong đó có hàng chục ấn phẩm bằng tiếng Anh dùng để giao lưu quốc tế, hàng trăm CSDL khoa học và công nghệ được cập nhật thường xuyên, hàng trăm bản tin điện tử với nội dung khá phong phú. Hình thức phục vụ và sản phẩm thông tin ngày càng đa dạng, bước đầu đã có đóng góp tích cực cho việc chuẩn bị và ra các quyết định của lãnh đạo các cấp và đáp ứng đáng kể các nhu cầu thông tin của công tác nghiên cứu - triển khai.

Việc ứng dụng công nghệ thông tin trong hoạt động thông tin tư liệu ở nước ta được bắt đầu từ cuối những năm 80 với việc sử dụng phần mềm tư liệu CDS/ISIS của UNESCO, đã

gia công cài đặt tiếng Việt để xây dựng, quản trị và khai thác các CSDL tài liệu của mình. Từ đầu những năm 2000, nhiều Trung tâm Thông tin và Thư viện đã triển khai dự án xây dựng thư viện điện tử. Một số mạng thông tin bắt đầu được xây dựng và khai thác đã góp phần thúc đẩy việc chia sẻ nguồn lực thông tin hiện có, tiêu biểu là các mạng:

- Mạng thông tin KH&CN Việt Nam (VISTA);
- Chợ ảo về Công nghệ và Thiết bị Việt Nam;
- Mạng Thông tin Thương mại VITRANET;
- Mạng thông tin nông nghiệp nông thôn AgroViet;
- Mạng thông tin Y dược học MEDLIB, ...

Đến đầu những năm 90 ở nước ta đã hình thành một đội ngũ cán bộ thông tin chuyên nghiệp. Đội ngũ này đang được tăng cường về số lượng và từng bước nâng cao về chất lượng. Chỉ riêng lĩnh vực thông tin khoa học và công nghệ, tổng số cán bộ trong toàn mạng lưới đã có khoảng 5000 người, phân lớn đã qua các lớp đào tạo cơ sở hoặc nâng cao.

Như vậy, từ năm 1960 đến nay hoạt động thông tin khoa học ở Việt Nam đã trải qua chặng đường lịch sử 40 xây dựng và phát triển. Hoạt động thông tin khoa học đã phát triển từ một thư viện khoa học đơn lẻ đến một mạng lưới thư viện khoa học có hàng trăm đơn vị, từ một phòng thông tin khoa học đến hệ thống thông tin khoa học quốc gia với hàng nghìn cán bộ chuyên môn tham gia như ngày nay.

Bằng hoạt động thực tiễn của mình, các cơ quan này đã góp phần xứng đáng đưa thông tin khoa học thực sự trở

thành ngành hoạt động xã hội quan trọng, một nghề nghiệp hữu ích đầy triển vọng và được coi là một yếu tố của tiềm lực khoa học và công nghệ, một nguồn lực cho sự phát triển của đất nước. Từ thực tiễn đó, nhiều vấn đề lý luận và phương pháp luận của thông tin học đã được nghiên cứu, kiểm nghiệm và phát triển.

Về mặt đào tạo cán bộ, khoa Thư viện trường Đại học Văn hoá Hà Nội là cơ sở đào tạo cán bộ thư viện đầu tiên ở nước ta, đã đào tạo cán bộ thư viện bậc đại học từ năm 1961. Từ năm 1990, trước đòi hỏi của thực tiễn, khoa đã đổi mới mạnh mẽ chương trình đào tạo, đưa vào nhiều môn học mới về thông tin học, tin học tư liệu, công nghệ thông tin và trở thành khoa Thư viện - Thông tin. Hàng ngàn cán bộ thông tin - thư viện được đào tạo ở đây đã là nguồn bổ sung cán bộ cho các thư viện và các cơ quan thông tin trong cả nước. Việc đào tạo sau đại học đối với ngành thông tin - thư viện cũng đã được mở ra tại trường Đại học Văn hoá Hà Nội từ năm 1991. Trên ba trăm thạc sĩ thông tin - thư viện đã ra trường. Nơi đây cũng là một trung tâm nghiên cứu khoa học của ngành.

Trong mấy năm gần đây việc đào tạo cán bộ thông tin - thư viện còn được mở ra ở các trường Đại học Quốc gia Hà Nội, Đại học Quốc gia Thành phố Hồ Chí Minh và Đại học Dân lập Đông Đô.

Như đánh giá nêu trong Chỉ thị 95-CT ngày 4-9-1991 của Chủ tịch Hội đồng bộ trưởng (nay là Thủ tướng Chính phủ), từ sau khi có Nghị quyết 89-CP của Chính phủ năm 1972 về tăng cường công tác thông tin khoa học và kỹ thuật, ở nước ta đã hình thành Hệ thống thông tin Khoa học và Công nghệ

Quốc gia với các cơ quan thông tin từ trung ương đến địa phương, bao quát hầu hết các ngành kinh tế - xã hội, các lĩnh vực khoa học và công nghệ với vốn tư liệu ngày càng phong phú. Hoạt động thông tin khoa học và công nghệ ở nước ta đã có những đóng góp quan trọng trong việc hoạch định chủ trương, đường lối, chính sách của Đảng và Nhà nước, trong việc triển khai các chương trình, kế hoạch phát triển kinh tế - xã hội, khoa học và công nghệ và trong việc phổ biến rộng rãi kiến thức khoa học và công nghệ, góp phần thúc đẩy phong trào lao động sáng tạo của quần chúng và nâng cao dân trí.

Mới đây, ngày 31-8-2004 Chính phủ đã ban hành Nghị định số 159/2004/ND-CP về hoạt động thông tin KH&CN nhằm hướng dẫn thực hiện điều 25 và điều 45 của Luật Khoa học và Công nghệ, cũng như tăng cường hoạt động thông tin Khoa học và Công nghệ trong tình hình mới. Điều 45 Luật Khoa học và Công nghệ chỉ rõ: “Chính phủ đầu tư xây dựng hệ thống thông tin khoa học và công nghệ quốc gia, đảm bảo thông tin đầy đủ, chính xác, kịp thời về các thành tựu quan trọng trong các lĩnh vực khoa học và công nghệ ở trong nước và trên thế giới; ban hành quy chế quản lý khoa học và công nghệ”.

Nghị định 159 đã khẳng định vai trò của Nhà nước, trách nhiệm của Chính phủ, các bộ, uỷ ban nhân dân các tỉnh, thành phố trực thuộc trung ương trong việc xây dựng và phát triển hệ thống thông tin khoa học và công nghệ quốc gia hiện đại, trong việc phát triển và tích hợp các nguồn tin khoa học và công nghệ trong nước và nước ngoài, bảo đảm cho các tổ chức và cá nhân dễ dàng và bình đẳng trong việc tiếp cận thông tin khoa học và công nghệ được tạo ra bằng ngân sách nhà nước. Nghị định 159 cũng tạo khung pháp lý ban đầu cho

việc xã hội hoá hoạt động thông tin khoa học và công nghệ, khuyến khích phát triển các dịch vụ thông tin khoa học và công nghệ có thu phí, tạo lập thị trường thông tin khoa học và công nghệ ở Việt Nam.

Nghị định 159 đi vào cuộc sống sẽ tạo nên những chuyển biến tích cực trong hoạt động thông tin khoa học và công nghệ ở nước ta.

Có thể nói lịch sử phát triển thông tin học ở Việt Nam gắn liền với sự hình thành và phát triển của hoạt động thông tin khoa học ở Việt Nam, khởi đầu từ đầu những năm 60 của thế kỷ XX đến nay.

Chương 2

THÔNG TIN VÀ TIẾN BỘ XÃ HỘI

1. VAI TRÒ CỦA THÔNG TIN

Ngay từ xưa thông tin đã đóng vai trò tổ chức và phát triển xã hội. Bất cứ cộng đồng nào cũng chỉ có thể tồn tại bằng cách truyền tin, dù đó mới chỉ là tiếng nói, tín hiệu, hình ảnh hay cử chỉ.

Người ta sớm nhận thức được rằng một khi đã nắm được và truyền được thông tin thì tức là đã có thêm một thứ quyền lực, giống như một sức mạnh vật chất. Có thể nói mọi thành tạo của lịch sử đều phụ thuộc vào thực trạng của các dòng tin, cũng như phụ thuộc vào sản xuất và trao đổi hàng hoá. Ngoài ra quá trình tồn tại và phát triển của xã hội gắn chặt với sự chuyển giao thông tin giữa các thế hệ, trong đó giáo dục và đào tạo giữ vai trò trọng yếu.

Trong thời đại ngày nay, thời đại mà khoa học và công nghệ ngự trị trong đời sống tất cả mọi người thì thông tin - yếu tố tiếp sinh khí cho nó - giữ vai trò cực kỳ trọng yếu. Vai trò đó thể hiện trên các mặt sau đây:

1.1. Thông tin là nguồn lực phát triển và là nguồn tài nguyên đặc biệt của mỗi quốc gia

Hiện nay người ta thừa nhận rằng vật chất, năng lượng, thông tin và bản sắc văn hoá dân tộc là các nhân tố quyết định sự tồn tại và phát triển của mỗi quốc gia. Đặc biệt trong điều kiện cách mạng khoa học và công nghệ đang diễn ra với quy mô lớn như hiện nay, khoa học và công nghệ đang trở thành lực lượng sản xuất trực tiếp của xã hội thì thông tin khoa học và công nghệ thật sự trở thành nguồn lực quan trọng tạo nên những ưu thế kinh tế và chính trị của mỗi nước. Và nếu như tiềm lực khoa học và kỹ thuật là điều kiện để nâng cao năng suất, hiệu quả của nền sản xuất xã hội thì thông tin khoa học và công nghệ được coi là yếu tố cực kỳ quan trọng của tiềm lực khoa học kỹ thuật.

Nếu như trước đây mọi nền kinh tế đều dựa chủ yếu vào các nguồn tài nguyên thiên nhiên, lấy việc khai thác và chế biến các tài nguyên thiên nhiên là nguồn chủ yếu tạo ra các của cải và sự giàu có cho xã hội, thì từ những năm cuối của thế kỷ XX, thông tin đã được xem là một nguồn tài nguyên kinh tế, giống như các tài nguyên khác như vật chất, lao động, tiền vốn... bởi vì việc sở hữu, sử dụng và khai thác thông tin có thể đem lại hiệu quả kinh tế cho nhiều quá trình vật lý và nhận thức. Ngày nay các hoạt động liên quan đến xử lý thông tin trong công nghiệp chế tạo cũng như trong việc giải quyết các vấn đề xã hội và con người tăng lên đáng kể. Khác với các nguồn tài nguyên khác, tài nguyên thông tin có thể mở rộng phát triển không ngừng và hầu như chỉ bị hạn chế bởi thời gian và khả năng nhận thức của con người. Khả năng mở rộng này thể hiện ở các thuộc tính sau đây:

- (1) Thông tin lan truyền một cách tự nhiên;
- (2) Khi sử dụng thông tin không bao giờ bị cạn đi mà trái lại càng trở nên phong phú do được tái tạo và bổ sung thêm các nguồn thông tin mới;
- (3) Thông tin có thể chia sẻ, nhưng không mất đi trong giao dịch.

Với khả năng thay thế các nguồn tài nguyên khác, khả năng truyền với tốc độ rất cao và khả năng đem lại ưu thế cho người nắm giữ nó, thông tin đã thực sự trở thành cơ sở cho nhiều hoạt động xã hội như nghiên cứu, giáo dục, xuất bản, tiếp thị và cả hoạt động chính trị nữa. Mối quan tâm của xã hội đối với việc sử dụng có hiệu quả các nguồn tài nguyên thông tin đã mở rộng từ các lĩnh vực truyền thống như thư viện, lưu trữ sang các tổ chức, các cơ quan, các ngành. Bây giờ ở đâu người ta cũng quan tâm đến việc quản lý và khai thác các nguồn tài nguyên thông tin.

Một khía cạnh nhận thức thứ hai về vai trò của thông tin trong những năm gần đây là ngày nay ở nhiều nước thông tin đã trở thành hàng hoá. Điều đó đã thúc đẩy hình thành một bộ phận mới trong nền kinh tế quốc dân, đó là khu vực dịch vụ thông tin. Khu vực này ngày càng tạo ra nhiều sản phẩm và dịch vụ thông tin đa dạng, đóng góp một tỷ trọng đáng kể trong nền kinh tế quốc dân. Người ta thấy rằng khối lượng, chất lượng của các sản phẩm và dịch vụ thông tin đã trở thành tiêu chí đánh giá trình độ phát triển kinh tế của mỗi nước.

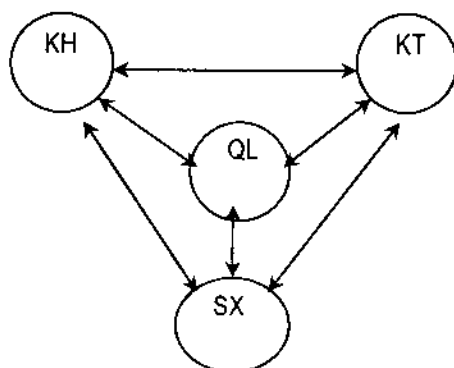
1.2. Thông tin là yếu tố quan trọng thúc đẩy sự phát triển kinh tế và sản xuất

Từ trước đến nay các hoạt động kinh tế và sản xuất đều cần đến thông tin. Các tổ chức sản xuất và kinh doanh đòi hỏi thông tin về nhu cầu đang thay đổi của khách hàng, các khuynh hướng thị trường đang tiến hoá, các vật liệu sản xuất mới đang xuất hiện, các thiết bị mới với khả năng kỹ thuật cao hơn, các công nghệ mới đang được áp dụng v.v... Chính quá trình liên hệ nhiều thông tin với nhau, dù ở xã hội nguyên sơ nhất, đã dần dần đưa đến kiến thức về những nguyên lý thông dụng trong sản xuất, giúp con người đổi mới và hoàn thiện các quy trình và phương pháp hiện hành. Đó chính là yếu tố quan trọng của sự phát triển, góp phần nâng cao sức sản xuất và tạo nên sự giàu có của xã hội. Từ giữa thế kỷ XX, nền sản xuất công nghiệp và kinh tế hàng hoá phát triển mạnh mẽ, nhu cầu thông tin và xử lý thông tin mới nảy sinh nhanh chóng và đòi hỏi được đáp ứng kịp thời, do đó vai trò của thông tin trong kinh tế ngày càng thêm quan trọng.

Trong điều kiện của cuộc cách mạng khoa học và công nghệ hiện nay, khoa học, kỹ thuật và sản xuất là các bộ phận có quan hệ khăng khít với nhau, tạo thành chu trình “Khoa học - Kỹ thuật - Sản xuất”. Trong đó mỗi bộ phận vừa là tiền đề, vừa là điều kiện thúc đẩy bộ phận kia phát triển. Thực chất của mối quan hệ hữu cơ này là trao đổi thông tin.

Phản ánh sự vận động của thông tin từ khoa học đến sản xuất và ngược lại, A.D. Urxul đã đưa ra mô hình liên hệ thông tin trong chu trình “Khoa học - Kỹ thuật - Sản xuất” thông qua tác động của quản lý, thể hiện ở hình 5.

Vì vậy thông tin có thể coi là bộ phận cấu thành chu trình “Khoa học - Kỹ thuật - Sản xuất”, góp phần rút ngắn quá trình từ nghiên cứu đến sản xuất và nâng cao hiệu quả của các lĩnh vực hoạt động này.



Hình 5. Chu trình “Khoa học - Kỹ thuật - Sản xuất”

Ngày nay sự liên hệ giữa thông tin và máy móc mang tính chất động và hai chiều. Thông tin được sử dụng để điều hành máy móc trong sản xuất. Ngược lại máy móc lưu giữ thông tin trong quá trình sản xuất (hoặc kinh doanh), chế biến chúng để tạo ra thông tin mới, đưa đến quyết định điều hành sản xuất mới.

Cùng với sự phát triển vượt bậc của khả năng lưu giữ, chế biến, tính toán và giá thành ngày càng rẻ của MTĐT và các phương tiện viễn thông, thông tin ngày càng được sử dụng hiệu quả trong sản xuất và kinh doanh. Các hoạt động đó ngày càng tạo thêm nhiều giá trị gia tăng trong các khu vực kinh tế. Đến đầu những năm 90, khu vực thông tin đóng góp hàng năm gần 1/2 tổng sản phẩm kinh tế quốc dân (GDP) trong nhiều nước phát triển.

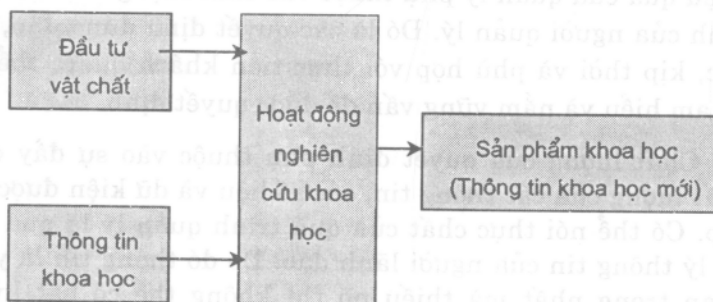
Công nghiệp hoá tại các nước tiên tiến sẽ tập trung vào phát triển mạnh dịch vụ, xuất khẩu dịch vụ, trong đó quan trọng nhất là ngành cung ứng thông tin. Điều này không có nghĩa là họ sẽ sản xuất ít hàng hoá đi mà là sản xuất ngày càng nhiều hàng hoá, nhưng với giá càng rẻ, khiến cho tỷ trọng đóng góp của hàng hoá vào GDP ngày càng nhỏ.

1.3. Thông tin giữ vai trò hàng đầu trong sự phát triển của khoa học

Vai trò của thông tin trong sự phát triển của khoa học thể hiện ngay trong quy luật phát triển của khoa học. Một trong những quy luật phát triển nội tại của khoa học là tính kế thừa và tính quốc tế của nó. Không một phát minh khoa học nào lại chỉ là sản phẩm lao động của một người, thậm chí một thế hệ. Phát minh đó là sản phẩm của nhiều người, của nhiều thế hệ. Lúc đương thời Ixaac Newton đã diễn tả tư tưởng này một cách sinh động như sau: “Nếu tôi có nhìn xa hơn người khác một phần nào, đó là vì tôi đứng trên vai những người khổng lồ”.

Tính kế thừa là yếu tố quan trọng thúc đẩy nhanh tiến bộ khoa học kỹ thuật. Người sau không làm lại việc người trước đã làm. Thế hệ sau chọn lọc, hệ thống hoá thành quả của người đi trước, phát hiện ra những quy luật mới. Quy luật mới này là sản phẩm khoa học, cũng là thông tin khoa học mới. Như vậy hoạt động nghiên cứu khoa học là một loại hoạt động đặc thù của con người nhằm thu được những thông tin khoa học mới, trên cơ sở những thông tin mà xã hội loài người đã tích lũy được. Hoạt động nghiên cứu khoa học có thể mô tả bằng mô hình ở hình 6.

Như vậy hoạt động nghiên cứu khoa học là một hệ thống tiếp nhận thông tin và tạo ra những thông tin mới khác với thông tin ban đầu.



Hình 6. Mô hình của hoạt động nghiên cứu khoa học

Có thể nói khoa học được nuôi dưỡng bằng chính khoa học. Những phát minh khoa học và cải tiến kỹ thuật sẽ chậm lại nếu cộng đồng khoa học không làm chủ được những thông tin khoa học tích lũy được theo thời gian. Điều đó giải thích nguyên nhân của sự yếu kém về khoa học kỹ thuật ở các nước thiếu nguồn thông tin tư liệu.

1.4. Thông tin là cơ sở của lãnh đạo và quản lý

Quản lý là một dạng tương tác đặc biệt của con người với môi trường xung quanh nhằm đạt được mục tiêu của tổ chức, trên cơ sở sử dụng tốt nhất các nguồn tài nguyên. Các tài nguyên ở đây bao gồm: con người, tri thức, tiền, vật chất, năng lượng, không gian, thời gian...

Quá trình quản lý có thể được xác định như một loạt các hoạt động định hướng theo mục tiêu, trong đó có các hành

động cơ bản là: xác định mục tiêu, lập kế hoạch để thực hiện mục tiêu, tổ chức và kiểm tra việc thực hiện kế hoạch đó.

Nhiệm vụ quan trọng của quản lý là ra các quyết định. Hiệu quả của quản lý phụ thuộc vào chất lượng của các quyết định của người quản lý. Đó là các quyết định đúng đắn, khoa học, kịp thời và phù hợp với thực tiễn khách quan, thể hiện sự am hiểu và nắm vững vấn đề được quyết định.

Chất lượng của quyết định phụ thuộc vào sự đầy đủ và chất lượng của các thông tin, các số liệu và dữ kiện được cung cấp. Có thể nói thực chất của quá trình quản lý là quá trình xử lý thông tin của người lãnh đạo. Do đó thông tin là yếu tố quan trọng nhất mà thiếu nó thì không thể có bất kỳ quá trình quản lý nào trong hệ thống tổ chức của xã hội.

1.5. Vai trò của thông tin trong văn hoá, giáo dục và đời sống

Con người có nhu cầu vật chất và nhu cầu tinh thần. Nhu cầu vật chất (ăn, mặc, ở, ...) dù có cao đến mấy, cũng chỉ là hữu hạn. Nhưng nhu cầu tinh thần (văn hoá, thông tin, giáo dục, giải trí, ...) bao giờ cũng là vô hạn. Sự phát triển bùng nổ của các phương tiện thông tin đại chúng, sự tiến bộ đến mức hoàn hảo của các phương tiện nghe nhìn, sự phát triển của các dịch vụ thông tin phong phú và đa dạng trên Internet,... đã ngày càng đáp ứng nhu cầu ngày càng cao về văn hoá tinh thần của con người, tạo cho con người những điều kiện tốt hơn để sáng tạo và hưởng thụ các giá trị văn hoá của dân tộc và của nhân loại.

Ngày nay máy tính cá nhân đã trở thành một phần quan trọng trong cuộc sống hàng ngày của nhiều người. Máy tính cá nhân không chỉ hỗ trợ con người trong công việc hàng

ngày, trong giao tiếp với bạn bè đồng nghiệp, mà còn đem lại cho họ những phút thư giãn như đọc những bản tin nhanh trên báo điện tử buổi sáng, nghe nhạc, xem phim buổi chiều.... Những dịch vụ web mạnh mẽ sẽ giúp con người tổ chức cuộc sống và truy cập bất cứ thông tin nào từ máy tính cá nhân ở nhà hoặc từ một máy tính xách tay có thể mang theo người mọi lúc mọi nơi. Những phần mềm ứng dụng tuyệt vời đã làm cho máy tính cá nhân không những trở thành một công cụ thiết yếu trong kinh doanh, mà còn trở thành một công cụ hỗ trợ cho sự sáng tạo của các nghệ sĩ: các nhạc sĩ sáng tác ra các bản nhạc hay, viết những tổng phổ phức tạp cho dàn nhạc, các họa sĩ sáng tác ra các tác phẩm đồ họa giá trị, các phim hoạt hình sinh động, các trang web giàu tính thẩm mỹ...

Giáo dục là hoạt động xã hội nhằm thực hiện chức năng chuyển giao thông tin giữa các thế hệ. Do đó giáo dục là nhân tố hàng đầu của sự phát triển. Các hoạt động giảng dạy, học tập, tự đào tạo ngoài quan hệ sư phạm giữa thầy và trò, luôn cần đến các kho tài liệu, các hoạt động khai thác và phổ biến tri thức của các thư viện và trung tâm thông tin. Ở các trường đại học, thư viện đóng vai trò quan trọng thứ hai sau người thầy. Các thư viện này phục vụ cho hoạt động nghiên cứu, giảng dạy, học tập của các giáo sư, giảng viên và sinh viên. Thư viện trường đại học Oxford ở Cambridge (thành lập từ thế kỷ XIV), thư viện trường đại học Harvard ở bang Massachusetts (thành lập năm 1638) còn lớn hơn nhiều thư viện quốc gia cả về tầm cỡ và chất lượng.

Ngày nay người ta có thể sử dụng máy tính cá nhân để tạo ra nhiều phương pháp học tập, phù hợp với nhiều đối tượng với nhiều trình độ khác nhau. Có nhiều phần mềm hỗ

trợ sinh viên, học sinh học tập, trình bày thông tin theo nhiều hình thức khác nhau và được cá nhân hoá dễ dàng so với phương pháp trên giấy tờ. Máy tính cá nhân làm thay đổi phương pháp học tập truyền thống, cho phép học sinh chủ động tham gia nhiều hơn, kích thích tính tò mò của học sinh ở mọi lứa tuổi, tạo điều kiện cho học sinh khám phá thông tin theo khả năng riêng của mình, làm thí nghiệm và trao đổi, học hỏi lẫn nhau.

Bước vào đầu thế kỷ 21, sự kết hợp của Internet và các công nghệ số đã tạo ra một mô hình đào tạo mới, đó là đào tạo trực tuyến, còn gọi là đào tạo điện tử hay E-learning. E-learning là một hình thức đào tạo mà trong đó các kiến thức và sự lĩnh hội được trao đổi thông qua các máy tính, Internet, các trang web hoặc từ các tổ chức mạng. Với E-learning, học viên có thể tham dự các khoá đào tạo tại mọi nơi, vào bất kỳ thời gian nào mà họ muốn, và nhất là học viên có cơ hội tiếp cận các lĩnh vực tri thức chuyên môn của thế giới với chi phí thấp theo nhiều cách với các phương tiện đa truyền thông.

Ngày nay cùng với sự phát triển của xã hội, nhu cầu thông tin phục vụ cho cuộc sống của con người ngày càng gia tăng. Mọi người sử dụng thông tin để lựa chọn sản phẩm, lựa chọn dịch vụ. Các thông tin về chính trị, xã hội và kinh tế giúp con người có định hướng đúng, làm chủ được đời sống của mình và thực hiện đầy đủ trách nhiệm và quyền hạn của người công dân.

Về mặt định tính, lý thuyết thông tin làm sáng tỏ một thuộc tính cơ bản của thông tin là: thông tin là sự đối lập với bất định và ngẫu nhiên, do đó nó phản ánh cái xác định và trật tự trong các mối quan hệ của sự vật và hiện tượng. Vì

vậy thông tin đúng đắn và chính xác bao giờ cũng đem lại trật tự và sự ổn định cho tổ chức. Ví dụ:

- Thông tin về pháp luật và hành chính đem lại trật tự, an ninh cho xã hội.
- Thông tin về kinh tế đem lại sự ổn định cho thị trường.
- Thông tin về thời tiết bảo đảm an toàn cho người sản xuất.
- Thông tin chính trị và thời sự giúp cho các thành viên trong xã hội có thái độ và định hướng xã hội đúng đắn v.v...

Ngày nay với sự phát triển của thông tin và truyền thông, khả năng giao lưu giữa các nhóm người và các thành viên trong xã hội tăng lên. Độ phức tạp của xã hội cũng tăng lên nhanh chóng. Giữ vững sự ổn định và định hướng phát triển của xã hội bằng cách xử lý đúng đắn và thông minh các luồng thông tin là vấn đề luôn đặt ra đối với các nhà quản lý và điều hành.

2. THỊ TRƯỜNG THÔNG TIN VÀ KINH TẾ THÔNG TIN

2.1. Bối cảnh mới của nền kinh tế thế giới

Sự phát triển mạnh mẽ của khoa học và công nghệ thông tin trong những thập niên gần đây đã mang lại những tác động to lớn và những biến đổi sâu sắc trong mọi lĩnh vực hoạt động kinh tế, văn hoá, xã hội ở khắp các nước trên thế giới. Chúng ta đang sống trong một thời đại mà nền kinh tế thế giới đang chuyển biến từ một nền kinh tế chủ yếu dựa vào nguồn tài nguyên hạn hẹp của thiên nhiên sang một nền kinh tế của thông tin và trí tuệ.

Trước kia sản xuất được phân bổ dựa trên lợi thế so sánh tự nhiên với hai yếu tố là các nguồn nguyên liệu và các yếu tố sản xuất như vốn và lao động. Ngày nay các ngành tăng trưởng nhanh nhất trong thập niên 1990 đều dựa vào sức mạnh của trí tuệ. Đó là các ngành vi điện tử, công nghệ sinh học, công nghệ vật liệu mới, vô tuyến viễn thông, máy tính điện tử (phần cứng, phần mềm), máy công cụ và robot, chế tạo máy bay dân dụng. Các tài nguyên thiên nhiên bị gạt ra ngoài chương trình cạnh tranh. Trong thời đại ngày nay, chỉ duy nhất có *trí thức và kỹ năng là các nguồn lực tạo ra lợi thế so sánh bền vững*. Điều đó đã nâng cao vị thế của thông tin khoa học và công nghệ.

Ở mọi ngành sản xuất, các hoạt động sản xuất có lợi nhuận cao nhất trong mỗi ngành là các hoạt động sản xuất có hàm lượng tri thức cao. Các hoạt động này đã tạo ra các sản phẩm chính xác, chất lượng cao, phục vụ cho các mục đích chuyên môn hoá cụ thể và có giá trị gia tăng. Các cơ sở kinh doanh cố gắng làm cho hàng hoá của mình “thông minh” hơn để kiếm lợi nhuận. Việc tìm tòi phát minh ra sản phẩm mới, đem lại lợi nhuận cao trở nên quan trọng. Xu thế sản xuất mới *chuyển từ khối lượng lớn sang giá trị cao* ra đời và ngày càng phát triển.

Ngày nay tri thức là nguồn duy nhất tạo ra lợi thế so sánh dài hạn, nhưng tri thức chỉ có thể áp dụng thông qua kỹ năng cá nhân. Tuy nhiên nhiều kỹ năng mới chưa được đưa vào dạy trong nhà trường, mà chúng chỉ có thể được học tập trong môi trường sản xuất. Trước sự thay đổi chóng mặt của công nghệ, người lao động phải liên tục phải học tập, trao đổi thông tin, tri thức và kỹ năng. Để sử dụng các sản phẩm thông minh, khách hàng cũng phải có một tiến trình học tập. Để tồn tại và phát triển các công ty phải xây dựng cách làm mới để thúc đẩy việc phổ biến thông tin, kỹ năng và tri thức

cho người lao động và khách hàng. *Doanh nghiệp trở thành một tổ chức học tập* là một thay đổi lớn về tư duy trong kinh doanh hiện đại.

Một đặc điểm lớn của nền kinh tế thế giới hiện nay là *xu thế toàn cầu hoá*. Theo Coitrnep (Tập chí Xã hội và Kinh tế Nga) toàn cầu hoá là giai đoạn mở rộng về chất của xu hướng tăng cường sự liên kết và phụ thuộc lẫn nhau giữa các nước. Toàn cầu hoá bao quát toàn bộ đời sống xã hội, trong đó mắt xích trung tâm là toàn cầu hoá về kinh tế, tiếp theo là văn hoá. Toàn cầu hoá về kinh tế thể hiện ở sự phân công lao động và phụ thuộc lẫn nhau về kinh tế, trong đó vai trò của các công ty đa quốc gia, các mạng thông tin liên lạc toàn cầu, các tổ chức kinh tế quốc tế ngày càng tăng; tự do thương mại, đầu tư tài chính ngày càng mở rộng; tính độc lập kinh tế của mỗi quốc gia chỉ là tương đối. Trong bối cảnh toàn cầu hoá, sự giao lưu văn hoá giữa các nước là điều tất yếu, là điều kiện để tồn tại và phát triển. Mỗi quốc gia phải biết kết hợp nhuần nhuyễn giữa giữ gìn bản sắc văn hoá dân tộc với tiếp thu tinh hoa văn hoá của thời đại.

Trong bối cảnh trên của nền kinh tế thế giới, vai trò của thông tin, nhất là thông tin khoa học và công nghệ ngày càng nâng cao. Ở nhiều nước thông tin thực sự trở thành hàng hoá và một thị trường thông tin, một nền kinh tế thông tin ra đời là một điều tất yếu.

2.2. Thị trường thông tin

Sự ra đời của thị trường thông tin

Với khả năng xử lý, lưu giữ và truyền một khối lượng lớn các dữ liệu với tốc độ cao, việc sử dụng MTĐT để xử lý thông tin đã tạo ra những thay đổi trong tổ chức hoạt động và cách thức chuyển giao thông tin. Các hoạt động này đã kích thích

việc nảy sinh những nhu cầu thông tin mới và do đó kích thích việc tạo ra các sản phẩm và dịch vụ mới có hàm lượng thông tin với chất lượng cao hơn.

Thập niên 1980 - 1990 của thế kỷ XX có thể coi là thập niên của thông tin mà sự kiện đáng chú ý là một "thị trường thông tin" hay một nền "công nghiệp thông tin" đã thực sự bắt đầu được hình thành. Hàng loạt các NHDL lưu trữ thông tin với khối lượng lớn và chất lượng cao ra đời. Nhiều tổ chức môi giới và dịch vụ thông tin được hình thành, thoả mãn nhu cầu đa dạng của người dùng tin.

Trong những năm này thị trường thông tin trực tuyến ở Tây Âu tăng hàng năm 25% và đóng góp những doanh thu đáng kể như được chỉ ra trong bảng dưới đây.

Bảng số liệu về doanh thu của hệ thống thông tin trực tuyến ở Tây Âu đối với thông tin toàn văn và thông tin thư mục trong năm 1983 và 1987 (tính theo tỷ đô la)

Năm Nước	1983		1987	
	Toàn văn	Thư mục	Toàn văn	Thư mục
Tây Đức	25,1	4,2	69	11,5
Pháp	26,5	6,3	77	15,3
Ý	13,1	1,5	40	3,4
Tây Ban Nha	5,2	0,4	14	0,6
Thụy Sĩ / Áo	62,0	2,1	113	4,2
Tổng	131,9	14,5	313	35,0

Bảng thống kê dưới đây nói lên sự phát triển của NHDL trên thế giới trong vòng 6 năm từ năm 1980 đến năm 1986.

Năm	Số NHDL	Số cơ quan sản xuất thông tin	Số cơ quan dịch vụ thông tin
1980	400	221	59
1981	600	340	93
1982	965	512	170
1983	1350	718	213
1984	1878	927	272
1985	1453	1189	362
1986	3169	1914	486

Những số liệu về thị trường thông tin trở thành chỉ tiêu đánh giá trình độ phát triển kinh tế của mỗi nước. Cần lưu ý rằng 80% bằng phát minh của thế giới được công bố trong 12 nước phát triển.

Từ những thông tin sơ cấp ban đầu, còn gọi là dữ kiện (data), người ta chế biến, lưu giữ, phân tích để tạo ra những thông tin có giá trị cao, gọi là thông tin có giá trị gia tăng (Value added information). Những thông tin có giá trị gia tăng có được qua xử lý, chế biến sẽ trở thành hàng hoá, có những giá trị cao trên thị trường. Một ngành mới phát triển mạnh mẽ trong những năm gần đây là ngành dịch vụ bán thông tin có giá trị gia tăng (value added information services). Ở Mỹ ngành này có doanh thu năm 1993 ước tính 13,6 tỷ USD, bằng 23% giá trị sản xuất máy tính và có tốc độ tăng trưởng 14-16% năm.

Việc sử dụng MTĐT để thiết kế, điều hành tự động sản xuất và robot là một bước tiến mới trong kỹ thuật sản xuất.

Đó là việc sử dụng tin học một cách thông minh. Người ta coi đó là những thông tin thông minh. Tiêu biểu là:

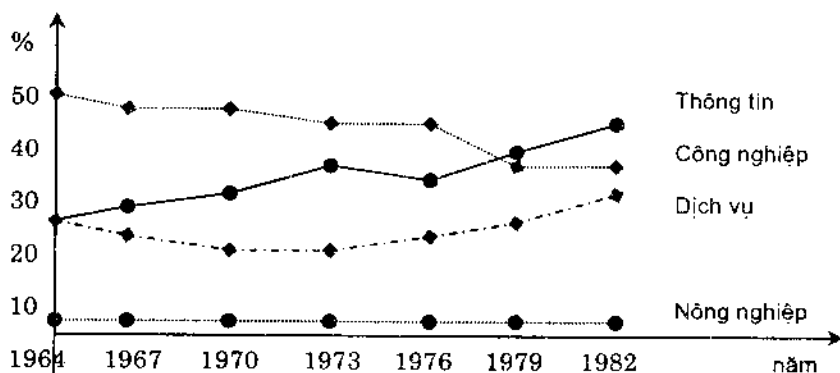
- CAD (Computer added design) dùng trong thiết kế.
- CAM (Computer added manufacturing) là chương trình điều hành sản xuất.
- CAE (Computer added engineering) là chương trình điều hành sản xuất tự động, sử dụng kết hợp với CAD.

Thị trường thế giới đối với thông tin này năm 1993 là 16,5 tỷ USD, bằng 20% thị trường sản xuất linh kiện điện tử, trong đó Mỹ chiếm 70%, châu Á 20%, châu Âu 10%.

Song song với sự phát triển của thị trường thông tin là sự thay đổi xã hội rất đáng lưu tâm. Đó là sự thay đổi trong phân công lao động của con người: những công việc có liên quan đến thông tin có chiều hướng gia tăng.

Biểu đồ cho tranh Hình 7 dưới đây cho ta hình ảnh về xu hướng phân bố lao động ở Anh từ năm 1964 đến năm 1982. Trong tổng số 21 triệu lao động ở Anh vào tháng 9 năm 1993 có tới 8,33 triệu lao động trong các ngành liên quan đến thông tin.

Xu hướng phân bố lao động ở Mỹ cũng tương tự. Vào những năm 60 ở Mỹ số người tham gia vào công việc xử lý thông tin nhiều hơn số người sản xuất ra lương thực, chế tạo ra hàng hoá và làm công việc dịch vụ. Xu hướng xã hội đó không dễ bị đảo ngược và không có lý do gì để có thể tin rằng khu vực thông tin mất vị trí của nó trong thị trường lao động. Và như vậy công nghệ thông tin sẽ có tác động trực tiếp tới một bộ phận lớn lao động ở các nước phát triển.



Hình 7. Sự thay đổi trong phân bố lao động ở Anh từ năm 1964 đến năm 1982 (tính theo tỉ lệ phần trăm)

Mặt khác giá trị của lao động trên lĩnh vực thông tin cũng cao hơn rất nhiều so với các lao động khác. Người ta không thể so sánh kết quả lao động của một giáo viên, một kỹ sư, một nhà sản xuất với sản phẩm thông tin trong hệ thống thông tin trực tuyến.

Dưới tác động của công nghệ thông tin, cơ cấu sản phẩm xã hội cũng có nhiều thay đổi. Các sản phẩm đáp ứng nhu cầu thông tin (máy thu thanh, thu hình, máy tính cá nhân, video, điện thoại truyền hình...) đang không ngừng tăng lên nhanh chóng về chủng loại, số lượng và chất lượng. Từ đó dẫn đến hình thành một thị trường thông tin.

Các thành viên của thị trường thông tin

Thị trường thông tin bao gồm nhiều thành viên, mỗi thành viên đảm nhiệm một chức năng xác định. Các thành viên này có thể phân chia theo bốn lĩnh vực sau đây:

1- *Người sản xuất ra thông tin*: thu thập thông tin trong các lĩnh vực từ các nguồn tin, xử lý, biên tập tạo ra các sản phẩm thông tin, sản xuất ra các CSDL.

Để đáp ứng yêu cầu thông tin của nhiều loại đối tượng khác nhau và để bao quát được tất cả các dạng thông tin khoa học kỹ thuật và kinh tế xã hội, nhiều cơ quan sản xuất thông tin, sản xuất các CSDL và NHDL ra đời. Đó là các cơ quan thông tin tư liệu quốc gia, các tổ chức xã hội, kinh tế và nghề nghiệp, các hội khoa học, các công ty, xí nghiệp lớn.

Ví dụ: Viện Thông tin Khoa học Kỹ thuật Pháp (INIST thuộc CNRS) là nơi sản xuất hai CSDL lớn: Pascal và Francis.

2- *Người làm dịch vụ phân phối thông tin*: tích nhập các CSDL của người sản xuất trên MTĐT và cung cấp cho người dùng tin các thông tin này nhờ một phần mềm chuyên dụng.

Các cơ quan dịch vụ thông tin thường sử dụng tin học viễn thông để nối giữa người dùng tin với các CSDL và NHDL. Đó là các tổ chức dịch vụ có trách nhiệm xử lý và khai thác thông tin bằng các phương tiện của tin học, đồng thời thương mại hoá các sản phẩm thông tin (phát triển phần mềm, kinh doanh thông tin và hướng dẫn yêu cầu).

Các cơ quan dịch vụ thông tin không chỉ bảo đảm tiếp cận thông tin của các CSDL và NHDL trong nước mà còn tiếp cận các CSDL và NHDL ở nước ngoài.

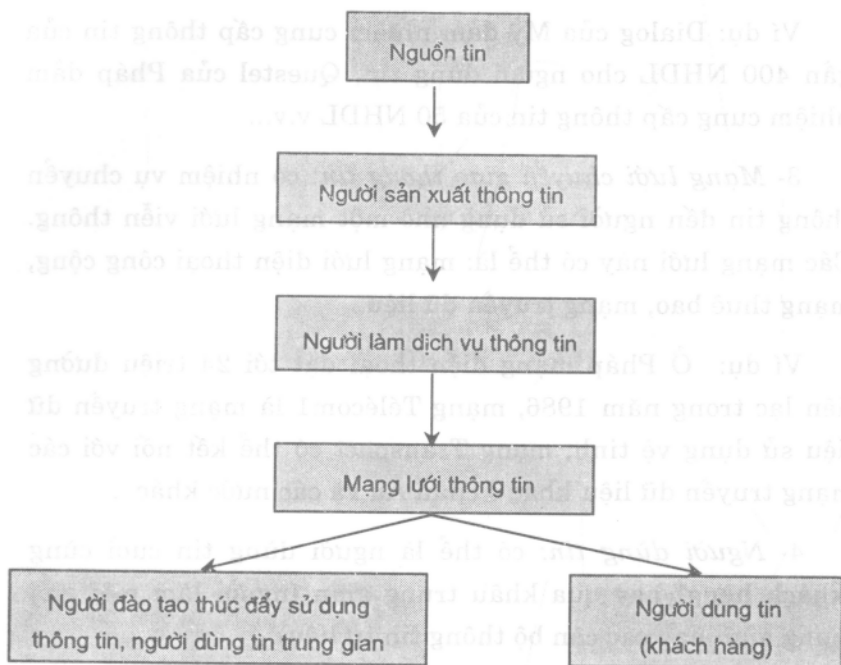
Ví dụ: Dialog của Mỹ đảm nhiệm cung cấp thông tin của gần 400 NHDL cho người dùng tin, Questel của Pháp đảm nhiệm cung cấp thông tin của 50 NHDL v.v...

3- *Mạng lưới chuyển giao thông tin*: có nhiệm vụ chuyển thông tin đến người sử dụng nhờ một mạng lưới viễn thông. Các mạng lưới này có thể là: mạng lưới điện thoại công cộng, mạng thuê bao, mạng truyền dữ liệu...

Ví dụ: Ở Pháp mạng điện thoại đạt tới 24 triệu đường liên lạc trong năm 1986, mạng Télécom1 là mạng truyền dữ liệu sử dụng vệ tinh, mạng Transpact có thể kết nối với các mạng truyền dữ liệu khác ở châu Âu và các nước khác...

4- *Người dùng tin*: có thể là người dùng tin cuối cùng (khách hàng) hay qua khâu trung gian (người làm môi giới trung gian hay các cán bộ thông tin tư liệu).

Người dùng tin ít hoặc không hỏi trực tiếp NHDL vì hai lý do. Trước hết là các ngôn ngữ hỏi thường khác nhau và khá phức tạp. Mặt khác các NHDL lại lưu trữ thông tin trên từng lĩnh vực khác nhau mà người dùng tin lại ít được hướng dẫn để tiếp cận NHDL có thể cung cấp cho câu trả lời tốt nhất. Vì vậy các trung tâm thông tin tư liệu hoặc các dịch vụ thông tin sẽ đóng vai trò trung gian giúp người dùng tin tiếp cận các thông tin cần tìm.



Hình 8. Hệ thống thông tin trực tuyến

Như vậy trong nền công nghiệp thông tin ở các nước tiên tiến hiện nay đã hình thành rõ nét sự phân công lao động giữa người sản xuất ra thông tin, người làm dịch vụ thông tin, người bảo đảm các phương tiện chuyển tải thông tin, người đào tạo, thúc đẩy sử dụng thông tin và người dùng tin. Tất cả tạo nên một hệ thống cung ứng thông tin mà người ta gọi là hệ thống thông tin trực tuyến (on-line information system). Hình 8 cho ta sơ đồ của hệ thống này.

2.3. Kinh tế thông tin

Các hoạt động của thị trường thông tin trong thập niên 80 - 90 đã kích thích làm nảy sinh những nhu cầu thông tin

mới, đòi hỏi phải tạo ra các sản phẩm và dịch vụ thông tin có chất lượng cao hơn. Đó là con đường tự nhiên dẫn đến sự hình thành một khu vực thông tin trong nền kinh tế. Các thành viên của nền kinh tế thông tin bao gồm chủ yếu các loại ngành hoạt động sau đây:

- Các hoạt động sản xuất ra thông tin: các hoạt động quản lý nhà nước (làm chính sách, kế hoạch, biện pháp quản lý kinh tế và xã hội), quản lý doanh nghiệp, nghiên cứu khoa học và công nghệ, sáng tạo trong văn hoá và nghệ thuật, hoạt động tư vấn, v.v...

Đặc điểm của các loại sản phẩm của các hoạt động này là giá trị của thông tin tạo ra không hề giảm khi nó được đem ra sử dụng, trái lại nó có thể gia tăng do được bổ sung thêm các thông tin mới. Chi phí cho việc sản xuất ra thông tin rất cao, nhưng giá trị sao chép lại rất thấp. Ví dụ, chi phí cho việc biên soạn từ điển bách khoa rất cao, nhưng công sao chép là một số tiền ai cũng có thể chịu được.

- Các hoạt động dịch vụ thông tin bao gồm: giáo dục và đào tạo, bưu điện, điện thoại, truyền tin, phát thanh truyền hình, xuất bản, quảng cáo, tiếp thị, dịch vụ thông tin từ các trung tâm thông tin tư liệu, các cơ quan công cộng, các tổ chức văn hoá, trên các mạng máy tính...

Ở quốc gia nào cũng vậy, có khoảng từ 1/5 đến 1/3 lao động làm việc trong khu vực này.

- Công nghiệp CNTT và viễn thông: sản xuất ra các sản phẩm như MTĐT, các thiết bị tin học và viễn thông, phần mềm máy tính, các dịch vụ lắp đặt và tích hợp hệ thống, v.v...

Các sản phẩm phần mềm sản xuất ngày càng nhiều, giá trị của chúng chiếm phần chủ yếu trong công nghiệp thông tin. Năm 1990 giá trị thu được từ sản xuất phần mềm ở Mỹ là 214 tỷ, đóng góp 2% GDP. Năm 1994 con số đó là 431 tỷ. Tốc độ gia tăng hàng năm của khu vực này là 12-13% năm. Hiện nay công nghệ phần mềm chiếm vị trí thứ ba trong nền kinh tế Mỹ, chỉ sau ngành công nghiệp sản xuất ô tô và điện tử.

Năm 1990 hoạt động thông tin ở Mỹ chiếm 47% lao động, ở Anh chiếm 45,8%, ở Pháp chiếm 45,1%, còn ở Đức là 40%. Nếu tính tỷ lệ đóng góp vào GDP của hoạt động thông tin thì con số cũng tương tự.

Chú ý rằng CNTT chỉ là phần rất nhỏ của kinh tế thông tin. Tuy CNTT phát triển nhanh, nhưng cũng chỉ chiếm 4,7% GDP và khoảng 10% phần đóng góp của cả khu vực kinh tế thông tin nói chung (ở Mỹ, đóng góp của phần này vào GDP trong năm 1994 là 576 tỷ USD, gồm 145 tỷ cho phần cứng và 432 tỷ cho phần mềm và dịch vụ).

CNTT chỉ có ý nghĩa khi người quản lý coi thông tin là động lực cực kỳ quan trọng để phát triển kinh tế, có khả năng tổ chức thông tin, thu thập thông tin, xử lý thông tin nhanh chóng và tạo ra ngày càng nhiều thông tin thông minh có giá trị. Đây là mấu chốt của sự phát triển.

Nền kinh tế thông tin có những đặc trưng sau:

- Nội dung thông tin bao trùm khắp các hoạt động sản xuất và kinh doanh.

- Nội dung thông tin chiếm tỷ trọng lớn trong sản phẩm hàng hoá và dịch vụ.

- Khu vực thông tin đóng góp một tỷ trọng lớn (40% hoặc hơn) phần giá trị gia tăng trong thu nhập quốc dân của một nền kinh tế phát triển.

Trong nền kinh tế thông tin công nghệ thông tin và truyền thông có vai trò động lực, thể hiện trên các mặt sau đây:

- Cuộc cách mạng công nghệ thông tin và truyền thông là nguyên nhân quan trọng của việc hình thành và phát triển kinh tế thông tin.

- Công nghệ thông tin và truyền thông phát huy vai trò của thông tin và tri thức là nguồn gốc và động lực phát triển kinh tế.

- Công nghệ thông tin và truyền thông là nền tảng của hội nhập và toàn cầu hoá.

- Bản thân công nghiệp và công nghệ thông tin và truyền thông là bộ phận chủ đạo và năng động nhất trong nền kinh tế thông tin.

Vì vậy xây dựng và phát triển hạ tầng công nghệ thông tin và truyền thông phải là ưu tiên quan trọng của mỗi quốc gia để phát triển kinh tế thông tin.

3. VÀI NÉT VỀ HIỆN TƯỢNG TIN HỌC HOÁ XÃ HỘI VÀ XÃ HỘI THÔNG TIN

3.1. Tiến trình phát triển của khoa học và công nghệ

Dưới tác động của cuộc cách mạng khoa học và công nghệ hiện đại, thế giới đang trải qua những biến đổi cực kỳ sâu

sắc. Lực lượng sản xuất của xã hội đang trong quá trình chuyển biến với những bước tiến cơ bản về chất lượng.

Sự chuyển dịch về công nghệ kéo theo sự chuyển dịch về kinh tế. Loài người đã trải qua hàng nghìn năm giai đoạn thứ nhất của nền văn minh - giai đoạn của nền sản xuất nông nghiệp thủ công. Tiếp đến giai đoạn thứ hai của nền văn minh nhân loại là giai đoạn của nền sản xuất cơ khí. Ngày nay với máy tính điện tử, với các thiết bị điều khiển tự động, các robot thông minh, loài người đang tiến tới giai đoạn thứ ba của nền văn minh, giai đoạn tin học hoá.

Trong giai đoạn của nền sản xuất nông nghiệp thủ công, con người sử dụng công nghệ thủ công. Đó là công nghệ mà nòng cốt là lao động chân tay của con người với các công cụ kỹ thuật thô sơ, bao gồm các ngành trồng trọt, chăn nuôi, chế biến nông sản, các nghề thủ công. Năng lượng và công cụ sử dụng chủ yếu là năng lượng, công cụ có sẵn trong tự nhiên. Đó là năng lượng tự nhiên của con người, sức súc vật, sức gió, sức nước v.v... Phạm vi tác dụng của công nghệ thủ công tới toàn bộ nền kinh tế xã hội của con người rất hạn hẹp, năng suất lao động rất thấp và của cải của xã hội được tạo ra rất hạn chế, vì nó bị giới hạn bởi khả năng lao động cơ bắp của con người, bởi các điều kiện sẵn có trong tự nhiên.

Trong giai đoạn của nền sản xuất cơ khí hoá, công nghệ cơ khí mà nòng cốt là các hệ thống máy móc đã nối dài bàn tay của con người trong việc chinh phục tự nhiên và xã hội. Với sự hỗ trợ của hệ thống máy móc (trong đó có máy phát lực, truyền lực và máy công tác) con người đã tạo ra năng suất lao động cao hơn nhiều so với lao động thủ công, tạo ra nhiều của cải vật chất cho xã hội, góp phần nâng cao và hoàn thiện mọi mặt của đời sống con người.

Công nghệ cơ khí cũng tạo ra nhiều cơ sở hạ tầng mới phục vụ cho đời sống con người. Đó là các cơ sở năng lượng lớn, các hệ thống giao thông phát triển rộng khắp với các phương tiện vận tải bằng máy móc, giúp con người đi lại, vận chuyển hàng hoá nhanh chóng, thuận lợi bằng cả đường bộ, đường không và đường thuỷ. Công nghệ cơ khí cũng làm biến đổi khu vực sản xuất thủ công bằng cách hiện đại hoá hoặc thay thế nó ở những công đoạn cần thiết.

Công nghệ cơ khí giải phóng con người khỏi lao động chân tay đã làm thay đổi tốc độ phát triển kinh tế và cơ cấu kinh tế của các quốc gia. Giá trị của cải gia tăng, khu vực công nghiệp ngày càng chiếm tỉ trọng cao trong cơ cấu sản phẩm của toàn xã hội.

Các thành tựu nghiên cứu về điện tử, về quá trình vận động của điện tử trong những môi trường vật chất khác nhau đã là nền tảng cho sự ra đời và phát triển nhanh chóng của kỹ thuật điện tử và vi điện tử. Nhờ đó con người giải quyết được những vấn đề phức tạp về đo lường, khuếch đại, biểu diễn và biến đổi tín hiệu. Đó là cơ sở của công nghệ thông tin, một ngành khoa học tiến nhanh như vũ bão trong mấy chục năm gần đây.

3.2. Hiện tượng tin học hoá xã hội

Sự ra đời và phát triển nhanh chóng của công nghệ thông tin, mà cốt lõi của nó là hệ thống tri thức về tin học, thông tin viễn thông và hệ thống phương tiện kỹ thuật máy tính điện tử, kỹ thuật viễn thông, là cơ sở vững chắc của quá trình tin học hóa các hoạt động của con người trong mọi lĩnh vực sản xuất, kinh tế và xã hội.

Tin học hóa - đó là quá trình tự động hoá các hoạt động điều khiển, tức là tự động hoá các vận động thông tin. Quá trình đó đang tác động mạnh mẽ và xâm nhập vào các mặt hoạt động của con người: trong khoa học kỹ thuật, trong sản xuất và tổ chức sản xuất, trong thông tin liên lạc, trong các hoạt động kinh tế như ngân hàng, tài chính, thống kê, kế hoạch, trong điều tra cơ bản, trong các hoạt động dịch vụ, trong công tác hành chính văn phòng, trong quản lý kinh tế xã hội và trong quản lý sản xuất kinh doanh v.v...

Tin học hoá xã hội, đó là sự chuyển biến về chất của lực lượng sản xuất xã hội. Sự chuyển biến này kéo theo một loạt sự chuyển biến khác về tính chất lao động sản xuất của con người, về tổ chức sản xuất và kinh tế, về cơ cấu sản phẩm, cơ cấu lực lượng sản xuất và cơ cấu tạo thành giá trị của xã hội, về quan hệ xã hội v.v...

Ngày nay cùng với nhu cầu về vật chất (ăn, mặc, ở) nhu cầu về tinh thần (thông tin, tri thức, văn hoá nghệ thuật) của con người phát triển không ngừng. Ta thấy mấy chục năm gần đây cơ cấu sản phẩm tiêu dùng của thế giới thay đổi nhanh chóng. Các sản phẩm đáp ứng nhu cầu thông tin của con người như máy thu thanh, thu hình, máy tính cá nhân, video, điện thoại truyền hình... đang không ngừng tăng lên nhanh chóng về chủng loại, số lượng và chất lượng.

Do khả năng ứng dụng rộng rãi trong đời sống kinh tế xã hội, nhu cầu của con người về thông tin và các sản phẩm về công nghệ thông tin ngày càng gia tăng. Trong công nghiệp thêm nhiều các sản phẩm là tư liệu sản xuất, công cụ lao động chứa các yếu tố xử lý thông tin và điều khiển tự động. Nên sản xuất xã hội hình thành nên một khu vực cung cấp các sản phẩm thông tin có giá trị gia tăng, các sản phẩm

phần cứng và phần mềm. Từ đó ra đời khu vực thông tin trong nền kinh tế, còn gọi là kinh tế thông tin.

Việc sử dụng thông tin thiết kế và điều hành tự động sản xuất bằng máy tính và robot là một bước tiến mới trong kỹ thuật sản xuất, dẫn đến tạo ra một phương thức sản xuất mới - phương thức sản xuất linh hoạt. Với phương thức này, sản xuất không cần hàng loạt, không cần tiêu chuẩn hoá như trước đây mà linh hoạt, dễ dàng tạo nhiều mẫu mã khác nhau với số lượng nhỏ, thoả mãn yêu cầu đa dạng của khách hàng. Với sự phát triển của công nghệ thông tin, quy mô sản xuất sẽ nhỏ lại, tính chất và công năng mặt hàng sẽ thay đổi liên tục nhằm đưa vào những đổi mới về kỹ thuật, chất lượng và mẫu mã.

Cơ cấu về nhu cầu sản phẩm và cơ cấu kinh tế thay đổi đã dẫn đến những thay đổi cơ bản trong cơ cấu tạo nên giá trị nền kinh tế quốc dân. Ở các nước có nền kinh tế phát triển hiện nay thì phần đóng góp giá trị của khu vực dịch vụ và thông tin cho nền kinh tế quốc dân ngày càng chiếm tỉ trọng lớn. Ở các nước này phần đóng góp giá trị của khu vực sản xuất nông nghiệp chỉ có 3 - 5%, của khu vực sản xuất công nghiệp 30 - 40%, của khu vực dịch vụ 20 - 30%, của khu vực thông tin 35 - 40%.

Những thay đổi trên đây trong nền sản xuất xã hội dẫn đến những thay đổi trong tính chất của lao động và cơ cấu thành phần lao động của xã hội. Ngày nay lao động điều khiển các máy móc cơ khí đã và đang được thay thế dần bằng các thiết bị tự động hoá. Kỹ thuật tin học phát triển đã làm mất đi một số lớn chỗ làm việc cũ, đồng thời cũng tạo ra nhiều chỗ làm việc mới. Nhiều nghề cũ mất đi, nhiều nghề mới ra đời. Như vậy, tin học hoá đã làm thay đổi tính chất lao

động và cơ cấu nghề nghiệp. Trong xã hội tương lai, phần chủ đạo trong lao động của con người là lao động trí tuệ: thiết kế và tổ chức các hệ thống thông tin tự động hoá, các hệ thống điều khiển và quản lý bằng máy tính, thiết kế các loại máy tính thế hệ mới, các loại thiết bị chuyên dụng, các phần mềm, các hệ chuyên gia, các robot thông minh, v.v...

Một hệ quả của quá trình tin học hoá xã hội là ngày nay tính chất, chất lượng, quy mô và khả năng của việc sản xuất và truyền thông tin thay đổi khác hẳn thời kỳ trước, thể hiện tập trung ở các hệ thống thông tin viễn thông với các CSDL và NHDL khổng lồ. Các hệ thống này cung cấp thẳng đến tận nhà những thông tin cần thiết vào lúc mà người ta yêu cầu.

Ngày nay do có khả năng với tới các thông tin một cách nhanh chóng từ mọi nơi, mọi câu hỏi đều được trả lời ngay, cho nên đã mở ra khả năng mới thực hiện sự đối thoại xã hội với quy mô lớn và chất lượng cao hơn. Thay cho các cộng đồng xã hội nhỏ, ở đó các thành viên trong xã hội đối thoại trực tiếp với nhau là mạng máy tính. Khả năng giao lưu giữa các nhóm người và các thành viên trong xã hội tăng lên. Độ phức tạp của xã hội cũng tăng lên nhanh chóng.

Một yêu cầu cấp bách hiện nay là phải xây dựng kết cấu hạ tầng cho nền kinh tế mới dưới tác động của tin học hoá. Xây dựng kết cấu hạ tầng thông tin quốc gia đang là mục tiêu của nhiều nước chuẩn bị bước vào thế kỷ XXI.

Kết cấu hạ tầng thông tin quốc gia là một hệ thống thống nhất của các mạng truyền thông, các máy tính, các CSDL và NHDL, các phương tiện điện tử dân dụng, sẵn sàng cung cấp những lượng thông tin to lớn với mọi hình thức biểu hiện tới người sử dụng ở bất kỳ đâu, vào lúc mà họ cần. Như vậy

CNTT sẽ cung cấp những cách thức mới cho việc học tập, làm việc, giao lưu,... tóm lại là cách thức hoàn toàn mới cho việc ứng xử của con người.

Kết cấu hạ tầng thông tin quốc gia là cơ sở cho việc phát triển nhanh chóng nền kinh tế thông tin và xã hội thông tin trong tương lai.

3.3. Xã hội thông tin

Sự phát triển mạnh mẽ của khoa học kỹ thuật cùng với sự trợ giúp của công nghệ thông tin tiên tiến đang đưa thế giới hiện đại chuyển từ thời đại công nghiệp sang thời đại thông tin. Sự chuyển đổi xã hội từ một xã hội căn bản dựa trên công nghiệp chế tạo sang xã hội dựa trên sản xuất và sử dụng thông tin và tri thức, ứng dụng thông tin và công nghệ truyền thông vào mọi lĩnh vực hoạt động của con người đã dẫn đến việc hình thành xã hội thông tin. Các nhà xã hội học đã phác hoạ “xã hội thông tin là xã hội mà ở đó không chỉ diễn ra quá trình chuyển dịch từ sản xuất sang dịch vụ mà còn là sự thay đổi từ một xã hội sản xuất vật phẩm sang xã hội sản xuất thông tin và tri thức”.

Xã hội thông tin là xã hội của nền kinh tế tri thức. “Nền kinh tế được dẫn dắt bởi tri thức là nền kinh tế mà việc sản sinh và khai thác tri thức có vai trò nổi trội trong quá trình tạo ra của cải” (theo Bộ Thương mại & Công nghiệp Anh – 1998)

Theo giáo sư Nick Moore, nghiên cứu viên cao cấp thuộc Viện nghiên cứu chính sách London thì các xã hội thông tin có ba đặc trưng cơ bản sau:

- Thông tin được sử dụng như một nguồn lực kinh tế. Các tổ chức sử dụng nhiều thông tin hơn để đổi mới, cải tiến, nâng cao chất lượng hàng hoá và dịch vụ, để tăng thêm hiệu quả và khả năng cạnh tranh. Hàm lượng thông tin chứa tỉ trọng lớn trong giá trị sản phẩm, trong tổng giá trị của nền kinh tế.

- Nhu cầu thông tin trong xã hội ngày càng gia tăng. Mọi người sử dụng thông tin để lựa chọn sản phẩm, dịch vụ và để làm chủ cuộc sống riêng của mình. Người ta sử dụng thông tin để thực hiện quyền hạn và trách nhiệm của người công dân. Ngoài ra các hệ thống thông tin phát triển cũng tạo ra cơ hội cho quần chúng tiếp nhận và hưởng thụ văn hoá, tiếp cận tới các cơ sở giáo dục.

- Phát triển một ngành thông tin ở ngay trong nền kinh tế. Chức năng của khu vực thông tin này là phải đáp ứng nhu cầu chung về phương tiện và dịch vụ thông tin, với cơ sở hạ tầng công nghệ là các mạng máy tính và viễn thông. Ở đây phải hình thành và phát triển một ngành công nghiệp sản sinh ra chính thông tin và các dòng thông tin được chuyển tải trên các mạng. Người ta nhận thấy rằng trong xã hội thông tin, khu vực thông tin phát triển nhanh hơn nhiều so với toàn bộ nền kinh tế.

Thuộc tính chung của xã hội thông tin là công nghệ thông tin đóng vai trò quan trọng trong việc nâng cao tốc độ và sự chính xác của các quá trình xử lý thông tin, trong việc nâng cao chất lượng của truyền thông, cũng như trong việc đổi mới cách thức quản lý tổ chức xã hội và quản lý các hoạt động kinh tế.

Việc tạo nên các xã hội thông tin riêng biệt đang diễn ra trong một quá trình thay đổi mạnh mẽ và mang tính chất quốc tế. Một phần là do các hệ thống thông tin đang phát triển hiện nay mang tính toàn cầu, mà Internet là một ví dụ. Mặt khác ngày nay tất cả các nước đều quan tâm đến việc dùng thông tin để nâng cao khả năng cạnh tranh và duy trì chỗ đứng của họ trong một thị trường toàn cầu mang tính cạnh tranh ngày càng cao.

Một thuộc tính của xã hội thông tin là có một nền công nghiệp thông tin phát triển ở trong nền kinh tế. Khi xem xét việc phát triển khu vực thông tin người ta thường phân chia khu vực này thành ba phân ngành riêng biệt: một phân ngành liên quan đến việc sáng tạo thông tin, tức là nội dung, phân ngành hai liên quan đến việc cung cấp thông tin và phân ngành ba liên quan đến việc xử lý thông tin.

Phân ngành nội dung thông tin bao gồm các tổ chức và cá nhân tham gia vào việc tạo dựng và phát triển tài sản trí tuệ. Đó là những người sáng tạo ra tri thức mới, sản xuất ra các sản phẩm và dịch vụ thông tin, đảm bảo các hoạt động nghiên cứu và triển khai (R&D).

Thông tin bắt nguồn từ nhà khoa học và kỹ thuật, các chuyên gia, tư vấn, các nhà văn và các nhà sáng tác. Các nhà xuất bản, các đài phát, các nhà phân phối và các công ty sản xuất mua sản phẩm trí tuệ của họ, xử lý chúng theo các cách khác nhau để có thể phân phối và bán chúng cho người tiêu dùng thông tin. Ngày nay với công nghệ thông tin hiện đại, ranh giới giữa các dạng thông tin bị phá vỡ, các công ty truyền thông đa phương tiện khai thác các tài liệu dưới các dạng chữ viết, âm thanh, hình ảnh,... và gộp chúng lại thành một chương trình thông tin.

Cùng với lực lượng sáng tạo ra thông tin, trong phân ngành nội dung còn có một bộ phận biên tập thông tin như: những người chuyên biên soạn các tài liệu tra cứu, các CSDL, các danh mục thống kê và các dịch vụ thông tin “tức thời” cung cấp các luồng thông tin liên tục như giá cổ phiếu, giá hàng hoá. Các nhà cung cấp thông tin này chiếm một phần quan trọng trong tổng số thu nhập của ngành nội dung thông tin, trong đó khu vực nhà nước đóng vai trò then chốt.

Phân ngành công nghiệp phân phối thông tin liên quan tới việc tạo lập và quản lý các mạng lưới phân phối và truyền thông tin. Đó là các công ty viễn thông, các công ty cung cấp các mạng lưới truyền hình cáp, các công ty truyền phát qua vệ tinh, radio và vô tuyến.

Liên kết với các tổ chức trên còn có các kênh khác để phân phối nội dung thông tin. Lĩnh vực này gồm các công ty phát hành sách, các thư viện, các hãng phát thanh truyền hình và các hãng cung cấp dịch vụ thông tin có giá trị gia tăng trên mạng.

Phân ngành công nghiệp xử lý thông tin bao gồm các nhà chế tạo phần cứng và các hãng chế tạo phần mềm. Các hãng chế tạo phần cứng thực hiện các công việc thiết kế, triển khai, sản xuất và đưa ra thị trường các máy tính, các thiết bị viễn thông và các mặt hàng tiêu dùng điện tử. Các nhà sản xuất phần mềm cung cấp cho chúng ta các hệ điều hành như UNIX, DOS hay Windows. Họ còn sản xuất các cụm chương trình ứng dụng như các hệ xử lý văn bản, bảng tính điện tử, các phần mềm thiết kế và quản lý theo yêu cầu của khách hàng.

Bảng dưới đây cho biết doanh số của ngành công nghiệp thông tin ở châu Âu và Mỹ năm 1994 (tính theo tỷ USD, theo số liệu thống kê của Ủy ban châu Âu).

Phân ngành CNTT	Liên minh châu Âu	Hoa Kỳ
Nội dung thông tin	186	255
Cung cấp thông tin	165	160
Xử lý thông tin	193	151
Tổng cộng	544	566

Gần đây với sự phát triển của CNTT, trên thị trường thông tin ngày càng xuất hiện nhiều *sản phẩm nội dung số*. Đó là những sản phẩm nội dung (văn bản, dữ liệu, hình ảnh, âm thanh...) được thể hiện dưới dạng số và được lưu giữ, lưu truyền trên môi trường điện tử (mạng máy tính, mạng Internet, mạng viễn thông, truyền thanh, truyền hình...). Những sản phẩm này được tích hợp nhiều dạng khác nhau, hoà trộn văn bản, âm thanh, hình ảnh (đa phương tiện), có thể dễ dàng lưu giữ, truy xuất, tái sản xuất, nâng cấp chỉnh sửa.

Từ đó, gần đây trên thị trường thông tin xuất hiện khái niệm *Công nghiệp Nội dung số (Digital Content Industry – DCI)*. Đó là ngành công nghiệp thiết kế, sản xuất, xuất bản, lưu trữ, phân phối và phát hành các sản phẩm nội dung số và các dịch vụ liên quan, gồm nhiều lĩnh vực như: tra cứu thông tin số liệu, giáo dục trực tuyến, thư viện số, quảng cáo tiếp thị trên Internet, v.v...

Theo số liệu của Trung tâm Thông tin Bưu điện, thuộc Tập đoàn Bưu chính Viễn thông Việt Nam thì doanh thu của

DCI toàn cầu năm 2002 đạt tới 172 tỉ USD và nửa năm 2006 đã đạt 430 tỉ USD, tốc độ tăng trưởng bình quân hàng năm là 30%. Đó là một tỉ lệ tăng rất cao.

Những thay đổi về công nghệ và sự phát triển của các ngành công nghiệp thông tin đặt ra những đòi hỏi mới đối với những người làm nghề thông tin: các cán bộ thư viện, các nhà khoa học thông tin, các nhà lưu trữ và xuất bản.

Ngày nay người làm nghề thông tin đang đứng trước hai áp lực. Thứ nhất, công nghệ thông tin đang mở rộng rất nhanh trong phạm vi công việc của họ. Thứ hai, những mong đợi của người dùng tin tăng lên cùng với thời gian và họ đòi hỏi nhiều hơn ở các dịch vụ thông tin chất lượng cao.

Những sức ép trên đây đòi hỏi một trình độ cao hơn ở các nhà chuyên môn. Họ không chỉ có kiến thức chuyên sâu về khoa học thông tin mà còn phải nắm vững những kỹ năng, kỹ thuật cần thiết để khai thác đầy đủ tiềm năng của công nghệ thông tin.

Chúng ta đang bước vào thế kỷ XXI, mọi xã hội trên toàn thế giới đều đang thay đổi. Ở các nước thuộc nhiều trình độ phát triển khác nhau, thông tin đang đóng một vai trò ngày càng quan trọng trong đời sống kinh tế xã hội, văn hoá và chính trị của họ. Chúng ta đang chứng kiến sự thay đổi rõ nét trong một nền văn hoá mà ở đó thông tin được sử dụng như một nguồn lực thúc đẩy tiến bộ xã hội và con người sử dụng thông tin như một phần trong công việc và đời sống hàng ngày của mình. Thông tin ngày càng in đậm trong đời sống văn hoá, sắp đặt cách thức làm việc, giải trí và làm phong phú thêm cuộc sống của chúng ta.

Có thể nói trong lịch sử loài người chưa có thành tựu khoa học nào lại tác động mạnh mẽ và nhanh chóng vào sự thay đổi và phát triển của xã hội như công nghệ thông tin và chúng ta cũng chỉ đang chứng kiến giai đoạn đầu của sự thay đổi đó.

4. THÔNG TIN VÀ CÁC NƯỚC ĐANG PHÁT TRIỂN

Một thực tế hiện nay là sự sản sinh ra thông tin và tình hình phát triển khoa học phân phối không đều trên thế giới. Các nước đang phát triển chỉ sản sinh ra chưa đầy 1% thông tin khoa học của thế giới. Trong khi đó phần lớn thông tin tập trung ở các nước công nghiệp phát triển. Ở đây họ giành một phần quan trọng trong thu nhập quốc dân cho hoạt động nghiên cứu, giáo dục và thông tin. Các nước này có một hệ thống cơ sở thông tin hùng mạnh với trang thiết bị hiện đại, đó là: các thư viện khổng lồ, các trung tâm thông tin tư liệu, các trung tâm phân tích tổng hợp tin. Họ có các chuyên gia và cán bộ thông tin được đào tạo; các kênh chuyển giao thông tin với nguồn tin phong phú và số lượng người dùng tin đông đảo; có chính sách thông tin quốc gia nhất quán. Ở các nước này người ta đang tổ chức lại hệ thống thông tin theo hướng hợp lý hoá và hợp tác hoá trên phạm vi quốc gia và quốc tế.

Trong khi đó ở các nước đang phát triển tình hình hoàn toàn khác. Các phương tiện giành cho việc sản sinh ra thông tin và mạng lưới thông tin còn rất thiếu thốn. Hầu hết các nước này chỉ có cơ sở hạ tầng yếu kém dựa trên các thư viện truyền thống, thiếu cán bộ, thiếu trang thiết bị hiện đại, gây nên sự chậm trễ trong việc chuyển giao và tiếp cận thông tin.

Vấn đề đặt ra đối với những nước này là phải nhanh chóng thu hẹp những lỗ hổng về thông tin gây nên do sự lạc hậu về khoa học và trang thiết bị kỹ thuật bằng cách tranh thủ tiếp nhận thông tin sẵn có ở các nước công nghiệp phát triển, hơn là nâng cao năng lực sản xuất thông tin ở địa phương, ít nhất là trong giai đoạn đầu. Điều đó đặt ra hai điều kiện là:

- Các nước đang phát triển phải nhanh chóng đào tạo cán bộ và xây dựng cơ sở hạ tầng của hệ thống thông tin quốc gia.
- Có sự tham gia của các nước công nghiệp phát triển vào hệ thống chuyển giao thông tin cho các nước đang phát triển.

Đã có những cố gắng đáng kể theo hướng này. Trong những năm gần đây chương trình thông tin khoa học kỹ thuật quốc tế UNISIST của UNESCO cùng một loạt các tổ chức liên quốc gia và quốc gia đã hoạt động nhằm tăng cường sự hợp tác thế giới trong lĩnh vực thông tin, đặc biệt giành ưu tiên cho các nước đang phát triển.

Hoạt động của UNISIST nhằm vào việc cải thiện quá trình chuyển giao thông tin theo ba hướng:

- Nâng cao khả năng sẵn có và khả năng tiếp cận thông tin khoa học kỹ thuật của các nước.
- Tăng cường liên kết giữa các hệ thống thông tin bằng cách tăng cường sử dụng các tiêu chuẩn quốc tế và công nghệ thông tin hiện đại.
- Tăng tính chọn lọc và mềm dẻo linh hoạt trong việc xử lý và phổ biến thông tin nhờ một cơ chế mới dựa vào sự đóng góp, tham gia tích cực của các cơ quan khoa học.

Tuy nhiên, vẫn còn nhiều trở ngại cho việc chuyển giao và tiếp cận thông tin ở các nước đang phát triển. Những trở ngại đó là:

- Thiếu sự quan tâm và đầu tư thích đáng của chính phủ, do không hiểu hoặc không đánh giá đúng vai trò của thông tin đối với sự phát triển của xã hội.

- Việc tiếp nhận thông tin khoa học kỹ thuật từ bên ngoài bị khống chế bởi nhiều yếu tố: hệ thống kiểm soát, hệ thống thuế quan, cước phí truyền thông từ xa quá tốn kém.

- Hàng rào ngôn ngữ: nếu khoa học là phổ dụng thì việc chuyển giao thông tin khoa học lại không phổ dụng mà nó chỉ dựa trên một số ngôn ngữ thể hiện một cách bất bình đẳng trên thế giới.

- Sự không hiểu biết: một trở ngại lớn cho sự phát triển, chuyển giao và tiếp cận thông tin là sự không hiểu biết đối với những lợi ích mà dịch vụ thông tin đem lại và không đủ khả năng sử dụng chúng.

Sự biến chuyển từ kinh tế công nghiệp sang kinh tế thông tin là hướng chủ đạo trong phát triển kinh tế thế giới hiện nay, nhưng chủ yếu xuất phát từ các nước công nghiệp phát triển nhất. Các nước đang phát triển chịu tác động to lớn và nhanh chóng có tính chất toàn cầu của sự biến chuyển đó trong khi nền kinh tế còn chủ yếu là nông nghiệp thủ công, công nghiệp còn nhỏ bé và yếu kém.

Để tiến kịp thế giới, các nước đang phát triển phải đồng thời tiến hành hai cuộc chuyển biến: từ kinh tế nông nghiệp thủ công sang kinh tế công nghiệp, từ kinh tế công nghiệp sang kinh tế thông tin. Theo cách nói của ta là công nghiệp

hoá và hiện đại hoá. Thách thức là vô cùng to lớn và nghiệt ngã.

Để thực hiện sự chuyển biến kép đó phải chọn được con đường thích hợp để phát triển nhanh chóng các yếu tố của kinh tế thông tin (hiện đại hoá), tạo nên những sức mạnh đầu tàu kéo theo sự phát triển công nghiệp theo hướng hiện đại hoá và nông nghiệp theo hướng công nghiệp hoá và hiện đại hóa.

Kinh tế thông tin là kinh tế của trí tuệ. Kinh nghiệm của một số nước chỉ rõ rằng nhân tố chủ yếu cho sự thành công là năng lực trí tuệ và sức sáng tạo của con người. Vì vậy tăng cường và hiện đại hoá hệ thống giáo dục, nghiên cứu khoa học và công nghệ, cùng với việc tạo môi trường thuận lợi khuyến khích sự phát triển không hạn chế mọi năng lực nội sinh trong sản xuất, kinh doanh, hoạt động kinh tế... là những biện pháp cấp thiết cần được ưu tiên thực hiện.

Trong hoàn cảnh nền kinh tế thế giới đang có xu hướng toàn cầu hoá thì sự giao lưu thông tin mang tính quốc tế cao độ. Bằng cuộc đấu tranh cho một trật tự thông tin mới, cho sự bình đẳng thông tin và bằng nỗ lực nội sinh, các nước đang phát triển có thể vươn lên nâng cao chất lượng của thông tin và nền kinh tế quốc dân của mình.

5. BÙNG NỔ THÔNG TIN VÀ NHỮNG BIỆN PHÁP KHẮC PHỤC

5.1. Hiện tượng bùng nổ thông tin

“Bùng nổ thông tin” là thuật ngữ đặc trưng cho sự gia tăng mạnh mẽ các sản phẩm thông tin tư liệu trên thế giới trong mấy chục năm gần đây. Khái niệm này được đưa ra

chính thức vào năm 1986 bởi Derek de la Solla Price, với ý nghĩa là sự phát triển bùng nổ của các tạp chí khoa học. Một vài con số dưới đây cho phép ta đánh giá mức độ của hiện tượng và xu thế của nó.

Từ vài vạn tên tạp chí vào đầu thế kỷ XX, đến năm 1972 đã có tới 170.000 ấn phẩm định kỳ. Trong năm 1970 trung bình mỗi ngày có 600 tài liệu khoa học được công bố (tức là khoảng 2.000.000 tài liệu mỗi năm). Con số này tăng lên từ 4 đến 5 lần vào năm 1985.

Đó là một tỷ lệ tăng rất nhanh. Một đặc trưng của hiện tượng này là độ gia tăng của nó còn tiếp tục trong những năm gần đây. Tỷ lệ này là 9,5% trong những năm 60 và đạt tới 10,5% vào năm 1971. Trái với một vài dự báo về sự bão hoà, nó còn tiếp tục gia tăng ngày càng nhanh.

Hiện tượng này gây ra chủ yếu bởi sự phát triển mạnh mẽ của khoa học kỹ thuật hiện đại và sự đổi mới công nghệ. Người ta có thể mô tả điều đó bằng hình ảnh sau. Theo tổ chức Liên hiệp Giáo dục Quốc gia Mỹ "Từ Thiên Chúa giáng sinh đến 1750 năm sau, tri thức loài người mới tăng gấp đôi. Việc tăng gấp đôi lần thứ hai được thực hiện trong vòng 150 năm sau, tức là vào năm 1900... Việc tăng gấp đôi lần thứ tư chỉ diễn ra trong vòng một thập niên sau năm 1950. Nói cách khác cứ 50 năm tri thức khoa học lại tăng lên 10 lần".

Sự phát triển nhanh của khoa học và công nghệ thể hiện ở chỗ ngày nay những thành tựu của khoa học và công nghệ ngày càng nhanh chóng đi vào đời sống. Khoảng thời gian từ lúc xuất hiện các ý tưởng khoa học, phát minh ra một công nghệ mới đến khi áp dụng rộng rãi vào sản xuất ngày càng được rút ngắn. Chẳng hạn để đưa vào áp dụng nguyên tắc

chụp ảnh phải mất hơn 100 từ khi phát minh ra nó (1727-1839); để thực hiện ý tưởng liên lạc bằng điện thoại đã mất hơn 50 năm (1820-1876); kỹ thuật vô tuyến 35 năm (1867-1902); ra đa 15 năm (1925-1940); vô tuyến truyền hình 12 năm (1922-1934); kỹ thuật bán dẫn 5 năm (1948-1953); mạch vi điện tử 3 năm (1958-1961) còn đến laze chưa đầy 2 năm.

Lực lượng những người làm khoa học cũng gia tăng nhanh chóng... Năm 1950 trên thế giới có khoảng 1 triệu nhà nghiên cứu, kỹ sư. Trong khi đó vào năm 1900 số đó chỉ có 100.000, năm 1850 là 10.000, năm 1800 là 1000⁽¹⁾.

Như vậy, số lượng các nhà khoa học, nhà nghiên cứu, nguồn chủ yếu của tri thức và thông tin khoa học không ngừng tăng nhanh. Ngày nay con số đó lên tới 10 triệu người. Nhiều nghiên cứu khẳng định rằng 9/10 các nhà khoa học trong cả chặng đường lịch sử loài người là những người đang sống cùng thời với chúng ta và 99% tri thức nhân loại là thuộc về những nhà bác học đó.

Ngoài ra cộng đồng khoa học còn được bổ sung thêm nhiều loại người dùng tin: các nhà quản lý, các giám đốc xí nghiệp, các nhà công nghệ, luật gia, các cán bộ chính trị, các nhà giáo dục... Họ không chỉ là những người dùng tin, mà còn là những người sản sinh ra những thông tin mới.

Tất cả những ai, bằng cách này hay cách khác, tham gia vào nền “công nghiệp sản xuất tri thức”, tức là sản xuất, phổ

⁽¹⁾ Các số liệu trong phần này lấy từ tài liệu tham khảo (2)

biến và tiêu thụ thông tin, đều thuộc vào những nhóm người dùng tin trên đây.

Về nguyên tắc, người ta có thể nói rằng “mọi sự chuyển giao tri thức tương đương với sự chuyển giao thông tin và ngược lại” và nền công nghiệp sản xuất tri thức khoa học, mà cơ sở của nó là truyền tri thức thông tin, tiếp tục tăng nhanh trong một thế giới xây dựng trên cơ sở của tiến bộ khoa học và công nghệ.

5.2. Hệ quả của hiện tượng bùng nổ thông tin

Sự gia tăng nhanh chóng khối lượng tri thức khoa học không thể không ảnh hưởng tới thành phần cơ cấu của kho tài liệu. Ngoài các sách báo và các ấn phẩm định kỳ xuất bản theo chu trình thương mại truyền thống, xuất hiện một loạt các tài liệu thuộc đủ các loại, không xuất bản, được phân phối ở mức độ hẹp như: các báo cáo, luận văn, tổng kết hội nghị, giáo trình, tài liệu nghiên cứu, tài liệu chuẩn bị xuất bản... Chúng tạo thành những tài liệu không công bố.

Bắt nguồn từ những cơ sở rất khác nhau như các viện nghiên cứu, các trường đại học, các hoạt động học tập nghiên cứu... những tài liệu này thường chứa những thông tin mới nhất trong lĩnh vực mà nó xem xét và chúng trở thành công cụ ưu tiên đặc biệt cho những giao lưu trực tiếp giữa các nhà khoa học. Mặc dù khó có thể biết được số lượng của những tài liệu này là bao nhiêu, nhưng người ta biết chắc rằng số lượng của chúng hiện nay tăng lên rất đáng kể.

Một hệ quả nữa của hiện tượng bùng nổ thông tin là sự rút ngắn đáng kể thời gian hữu ích của một tài liệu. Đó là tính lỗi thời của tài liệu. Trong một vài lĩnh vực, tri thức luôn

tự đổi mới với một tốc độ rất nhanh, đến nỗi người ta có thể nói rằng một quyển sách có thể trở thành vô giá trị ngay sau khi xuất bản nó. Do đó người ta phải thường xuyên bổ sung vốn tài liệu và không ngừng phải xử lý chúng, hoặc bằng thủ công hoặc bằng phương tiện tự động hoá.

Ngoài ra, bên cạnh những tài liệu văn bản in trên giấy, còn có thêm những tài liệu không ở dạng sách như: đĩa, ảnh, băng từ... Sự xuất hiện của những tài liệu loại này trong sự chuyển giao thông tin tương ứng với một yếu tố quan trọng trong xã hội ngày nay là: sự xuất hiện các phương tiện nghe nhìn. Với một tương lai rất hứa hẹn, chúng đặt ra cho những vấn đề về xử lý và phổ biến thông tin, dựa trên những kỹ thuật đặc biệt và dựa trên những kênh thông tin rất đa dạng.

5.3. Những phương hướng và biện pháp khắc phục

Để đáp ứng với dòng thông tin tư liệu gia tăng không ngừng do bùng nổ thông tin gây ra, những cơ quan thông tin được phát triển theo ba hướng:

- Mở rộng số lượng và quy mô.
- Đa dạng hoá và chuyên môn hoá.
- Tiếp thu và sử dụng các kỹ thuật và công nghệ thông tin mới.

Sự mở rộng về số lượng và quy mô dẫn đến việc hình thành những kho sách khổng lồ của những thư viện lớn. Theo số liệu tháng 1 năm 2005, Thư viện Quốc hội Mỹ (Library of Congress) sở hữu 128 triệu đầu tài liệu, trong đó có 29 triệu sách và tài liệu in khác, 12 triệu ảnh, 4,8 triệu bản đồ, 57 triệu bản viết tay, 2,7 triệu đĩa băng ghi âm, ghi hình. Các

tài liệu này xếp trên giá dài tới 530 dặm. Trong khi đó Thư viện Nhà nước Anh (The British Library) quản lý 150 triệu đầu tài liệu, xếp trên giá dài tới 625 km, và mỗi năm giá sách này lại dài thêm 12 km. Các Thư viện Quốc gia Pháp, Thư viện Nhà nước Nga cũng quản lý những khối lượng tài liệu rất lớn.

Sự đa dạng hoá và chuyên môn hoá thể hiện ở sự mở rộng chức năng, gia tăng công chúng phục vụ, những sản phẩm thông tin và phạm vi hoạt động. Nhiều tổ chức thông tin mới hình thành như những cơ sở dữ liệu và ngân hàng dữ liệu. Chúng lưu trữ thông tin ở mức độ rộng lớn chưa từng thấy.

Cuối cùng là sự đổi mới kỹ thuật. Thật vậy, sự bùng nổ thông tin tương đương với sự bùng nổ công nghệ. Đặc biệt trên ba lĩnh vực liên hệ chặt chẽ với công tác thông tin tư liệu là: tin học, viễn thông và vi xử lý.

Sự xuất hiện của máy tính điện tử trong việc xử lý thông tin tư liệu mới diễn ra trong vòng 40 năm nay. Được tạo thành bởi những thiết bị vào ra dữ liệu, vận hành với tốc độ cực nhanh, với những bộ nhớ gần như không hạn chế, một bộ tính toán không thể nhầm lẫn, máy tính điện tử đã mở ra hướng mới, đầy triển vọng cho việc xử lý thông tin.

Hệ quả của việc sử dụng máy tính điện tử trong công tác thông tin tư liệu thật là to lớn: tập trung thông tin trong những bộ nhớ lớn, những cơ sở dữ liệu và ngân hàng dữ liệu; tăng nhanh tốc độ ở tất cả các công đoạn xử lý thông tin; đảo ngược quá trình chuyển giao thông tin (không phải người dùng tin và tư liệu di chuyển mà là thông tin di chuyển); yêu

cầu thông tin có thể đặt ra từ xa, xuất phát từ một đầu cuối, nối với bộ phận trung tâm.

Sự phát triển của những bộ nhớ lớn truy đạt trực tiếp tạo cho khả năng tra cứu ngay, tại thời điểm bất kỳ những thông tin mà người dùng tin yêu cầu. Sự tiến bộ về chất trong quan hệ giữa người và máy, cùng với giá thành ngày càng hạ giúp cho việc sử dụng máy tính trong công tác thông tin tư liệu ngày càng trở nên phổ cập.

Sự kết hợp giữa máy tính và viễn thông hình thành nên ngành tin học viễn thông (télématique) là một trong những yếu tố cho sự phát triển các hệ thống và mạng lưới thông tin tự động hoá. Các hệ thống này được phát triển theo hai hướng:

- Các mạng truyền dữ liệu, là mạng thông tin được tạo thành bằng cách nối các nguồn tin với nhau sao cho các dữ liệu có thể lưu thông tự do giữa chúng.

- Các mạng thông tin máy tính, được tạo thành bởi các thiết bị đầu cuối và máy tính trung tâm gọi là máy chủ, cho phép người dùng tin ở đầu cuối có thể tiếp cận các thông tin trong các CSDL được lưu trữ trên mạng.

Giá thành cao của các thiết bị xử lý thông tin tự động hoá và phần mềm của các hệ thống này đòi hỏi phải có sự hợp tác của nhiều cơ quan. Điều đó đã tạo điều kiện cho việc điều chỉnh các phương pháp và hợp lý hoá việc phân phối nhiệm vụ và sản phẩm. Như vậy người ta đã làm chuyển hoá các hệ thống cũ, tập trung ngăn cách thành những mạng lưới tiếp nhận tư liệu thông suốt và mở ra nhiều điểm tiếp cận thông tin.

Cùng với sự phát triển mạnh mẽ của các kỹ thuật vi xử lý, việc sử dụng ngày càng nhiều các vi phim, vi phiếu đã loại bỏ được những khó khăn về sự quá tải của kho chứa và tạo điều kiện thuận lợi cho việc chuyển giao và phân phối thông tin.

Trong những năm gần đây, một vật mang tin mới là đĩa quang CD-ROM (Compact Disc - Read Only Memory) bắt đầu được sử dụng. Sự ra đời của CD-ROM dựa trên các thành tựu của công nghệ quang học - laser. Nhờ có các tính ưu việt như dung lượng nhớ lớn (khoảng 600 MB, tương đương với 300.000 trang in, 1500 đĩa mềm), độ bền vật lý cao, thao tác vận hành đơn giản, có khả năng lưu trữ văn bản, âm thanh, hình ảnh,... các đĩa quang ngày càng được sử dụng rộng rãi để lưu trữ và phổ biến thông tin.

Việc phát triển các phần mềm quản lý và khai thác CD-ROM và được ghi ngay trên đĩa làm cho việc sử dụng CD-ROM trở nên đặc biệt thuận lợi, là một bước quan trọng để hình thành thị trường và thương phẩm hoá các sản phẩm CD-ROM. Vì vậy nhiều CSDL thư mục và toàn văn đã được chuyển sang lưu trữ và khai thác trên CD-ROM. Trong những năm qua, một lượng thông tin khổng lồ, bao gồm các CSDL, các từ điển bách khoa, các sách tham khảo, các cẩm nang kỹ thuật, các chương trình phần mềm,... đã được phát hành dưới dạng CD-ROM.

Ngoài ra, việc sử dụng CD-ROM có thể phần nào thay thế hệ thống tìm tin trực tuyến, nhất là đối với các nước đang phát triển, gặp nhiều khó khăn về các phương tiện và cước phí truyền thông.

Mới đây, với sự phát triển của công nghệ thông tin đa phương tiện (multimedia), các thế hệ mới của CD-ROM ra đời, cho phép xử lý và khai thác các thông tin không chỉ ở dạng văn bản mà còn ở dạng âm thanh, hình ảnh tĩnh và động. Điều đó sẽ ảnh hưởng tới các quy trình công nghệ và xử lý thông tin truyền thống, đồng thời cũng mở rộng khả năng và chất lượng dịch vụ của các cơ quan thông tin.

Chương 3

CÁC LOẠI HÌNH TÀI LIỆU - NGUỒN TIN

1. TÀI LIỆU VÀ NHỮNG ĐẶC TRƯNG CƠ BẢN CỦA TÀI LIỆU

1.1. Tài liệu là gì?

Các thông tin chỉ có thể lưu giữ và truyền qua không gian và thời gian nếu nó được ghi lại trên các giá vật chất, với nhiều dạng và hình thức khác nhau. Chúng được gọi chung là tài liệu.

Tài liệu có thể định nghĩa là một vật thể cung cấp những chỉ dẫn và thông tin. Đó là cái giá vật chất mang tri thức của nhân loại.

Tất nhiên người ta có thể tìm kiếm thông tin qua nhiều nguồn khác nhau như: nói chuyện trao đổi, hội thảo, tham gia hội nghị, xem triển lãm, tham quan, nghe đài, xem truyền hình,... Nhưng những nguồn này bản thân nó cũng thu thập thông tin thông qua các tài liệu. Cho nên có thể nói tài liệu là yếu tố quan trọng nhất trong các hệ thống giao lưu thông tin và là bộ phận không thể tách rời trong hoạt động khoa học kỹ thuật và trong mọi lĩnh vực hoạt động của con người.

Có rất nhiều loại hình tài liệu khác nhau. Người cán bộ thông tin tư liệu cần phải biết những đặc trưng của chúng,

phân biệt được chúng thuộc loại hình tài liệu nào để có thể xử lý và sử dụng những tài liệu đó một cách thích hợp.

1.2. Những đặc trưng của một tài liệu

Mỗi tài liệu có hai đặc trưng chủ yếu: đặc trưng về mặt vật chất và đặc trưng về mặt tri thức.

a. Đặc trưng về mặt vật chất: thể hiện ở chất liệu và bản chất của các tín hiệu sử dụng, kích thước, trọng lượng, cách trình bày, phương tiện sản sinh, khả năng tra cứu trực tiếp hay thông qua một thiết bị, trạng thái thời gian....

Chất liệu là cái giá vật chất của tài liệu. Các chất liệu truyền thống thường là: đá, gạch, gỗ, sừng, kim loại, tre nứa, da thuộc, vải và phổ biến nhất là giấy. Giấy được phát minh ở Trung Quốc từ thế kỷ thứ 2 trước Công nguyên, được truyền bá vào các nước Ả Rập từ thế kỷ thứ 7 và được sử dụng ở Ý và Tây Ban Nha từ thế kỷ thứ 12. Những tiến bộ về khoa học và công nghệ đã tạo ra các chất liệu mới như chất dẻo (đĩa hát), chất liệu từ tính (băng âm thanh, băng hình, đĩa từ dùng trong máy tính), chất liệu hoá học (phim, giấy ảnh), đĩa quang v.v...

Bản chất của tài liệu thể hiện ở hình thức thể hiện thông tin. Căn cứ vào hình thức thể hiện thông tin người ta phân biệt tài liệu văn bản và tài liệu không văn bản.

Trong các tài liệu văn bản các thông tin được trình bày dưới dạng một bài viết mà người ta có thể đọc được. Ta có thể kể đến: các sách, các ấn phẩm định kỳ, các chuyên khảo, các văn bản về luật và hành chính, các catalogue kỹ thuật và thương mại.

Các tài liệu không văn bản có thể chứa một phần văn bản, nhưng ở đây chủ yếu các thông tin được trình bày dưới các dạng khác mà người ta có thể nghe, nhìn, sờ mó được. Người ta phân biệt các loại tài liệu sau:

- Các tài liệu dạng đồ biểu: bản vẽ, sơ đồ, bản đồ, bảng biểu, biển quảng cáo, tranh ảnh...

- Các tài liệu âm thanh: đĩa hát, băng nhạc, đĩa quang CD.

- Các tài liệu dạng nghe nhìn, bao gồm cả âm thanh và hình ảnh: phim, băng và đĩa video.

- Các tài liệu có bản chất vật chất: các tượng đài, các mẫu vật, các sản phẩm mẫu, các ma-két (maquettes), các giáo cụ trực quan...

- Các tài liệu từ dùng trong tin học: các phần mềm hệ thống và chuyên dụng, các chương trình tính toán và quản lý, các tệp dữ liệu...

Căn cứ vào cách thức sản sinh ra tài liệu người ta phân biệt tài liệu dạng nguyên khai và tài liệu sản sinh bởi con người. Các tài liệu nguyên khai là những vật thể tìm thấy trong tự nhiên như các mẫu đá, mẫu quặng, các vân thạch, các hoá thạch. Trong các tài liệu chế tạo bởi con người ta phải kể đến: các vật khảo cổ, các sản phẩm mẫu, các mẫu hàng, các sản phẩm trí tuệ như các sản phẩm mỹ nghệ, các tác phẩm văn học nghệ thuật, các phát minh sáng chế, các công trình nghiên cứu khoa học kỹ thuật...

Các kỹ thuật chủ yếu để sản sinh ra tài liệu là trạm khắc, in ấn, sao chụp. Kỹ thuật sao chụp hiện đại cho phép nhân bản tài liệu một cách nhanh chóng và tiện lợi. Kỹ thuật sản xuất các tài liệu vi phim, vi phiếu đem lại nhiều lợi ích cho

việc bảo quản và sử dụng tài liệu, cũng như trong giao lưu thông tin. Việc sử dụng công nghệ thông tin hiện đại trong thế kỷ XX, với kỹ thuật điện tử và kỹ thuật số để ghi các tín hiệu văn bản, âm thanh, hình ảnh đã tạo ra những tài liệu mới có khả năng lưu trữ thông tin đa dạng, với khối lượng lớn, cho phép truy cập thông tin một cách thuận lợi và có thể truyền đi xa. Thông tin ở đây được ghi dưới dạng tương đồng hay dạng số. Phương tiện để ghi thông tin với công nghệ mới là các băng từ, đĩa từ, đĩa quang.

b. Đặc trưng về mặt tri thức: thể hiện ở nội dung chủ đề, giá trị sử dụng, đối tượng công chúng, mức độ xử lý biên tập, mức độ phổ biến của tài liệu...

Nội dung của tài liệu được đánh giá theo nhiều tiêu chí khác nhau: chủ đề của tài liệu; cách trình bày tài liệu có hệ thống hay không, giành cho công chúng rộng rãi hay các đối tượng chuyên môn; xuất xứ của vấn đề và những nội dung mới; tính cập nhật của thông tin; tài liệu có chứa một phần hay chủ yếu là các số liệu.

Tất cả các tiêu chí trên chỉ mang tính chất tương đối: một tài liệu có thể không có một thông tin mới nào, nhưng nó có thể được trình bày sáng sủa hơn và dễ hiểu hơn đối với người sử dụng.

Tùy theo mức độ xử lý nội dung, người ta phân biệt tài liệu cấp một, tài liệu cấp hai và tài liệu cấp ba.

Tài liệu cấp một là tài liệu gốc được biên tập bởi tác giả. Tác giả của tài liệu có thể là một người, một nhóm người hay một tổ chức. Có nhiều loại tài liệu khác nhau. Về mặt hình thức, người ta phân biệt các loại sau:

- Các sách chuyên khảo, sách giáo khoa.
- Các báo cáo hội nghị, luận án, tài liệu phát minh sáng chế.
- Các ấn phẩm định kỳ (tạp chí, tập san...).
- Các thông báo, các ấn phẩm thư mục, tóm tắt.
- Các tiêu chuẩn, các catalog kỹ thuật và thương mại.
- Các tập bản đồ, đồ biểu.
- Phim ảnh, băng nhạc, băng hình, đĩa hát, đĩa quang.
- Các sổ tay, ghi chép, bản thảo.

Phân loại trên đây tạo điều kiện thuận lợi cho việc xử lý, sắp xếp và tìm kiếm tài liệu.

Tài liệu cấp hai là tài liệu giúp ta tra cứu tài liệu cấp một. Bản thân nó không thể tồn tại nếu không có tài liệu cấp một. Đó là các mô tả của tài liệu cấp một: các thư mục, mục lục, tạp chí tóm tắt... Ta sẽ trình bày chi tiết các loại tài liệu này ở mục sau.

Tài liệu cấp ba là những tài liệu được biên tập từ tài liệu cấp một và cấp hai. Nó tập hợp, đúc kết, biên tập những thông tin gốc dưới dạng phù hợp với yêu cầu của một lớp đối tượng người dùng tin nhất định. Đó là các tổng kết, tổng quan, tổng luận, báo cáo về một hiện trạng...

Tùy theo mức độ phổ biến người ta phân biệt tài liệu công bố và tài liệu không công bố.

Tài liệu công bố là những tài liệu có bán rộng rãi trên thị trường, ai cũng có thể mua được ở các hiệu sách hoặc ở các cơ quan xuất bản.

Tài liệu không công bố là những tài liệu không có bán trên thị trường và mức độ phổ biến của nó rất hẹp. Chúng tạo thành một loại tài liệu ngầm. Một số là những bản chép tay hay đánh máy, một số là những bản chụp lại hay in lại với số lượng rất hạn chế. Đó là những ghi chép nhật ký khoa học, những báo cáo kết quả nghiên cứu, những luận văn khoa học, những giáo trình giảng dạy, tài liệu hội nghị. Bắt nguồn từ những cơ sở rất khác nhau như các viện nghiên cứu, các trường đại học, các hoạt động học tập nghiên cứu v.v... những tài liệu này thường chứa những thông tin mới nhất trong lĩnh vực mà nó xem xét, những thông tin có giá trị cao, tuy chưa được thẩm định và kết luận một cách chính xác. Mặc dù không biết số lượng của những tài liệu này là bao nhiêu, nhưng người ta biết chắc rằng số lượng của chúng hiện nay gia tăng rất đáng kể. Việc tiếp cận các tài liệu này là rất cần thiết, nhưng đòi hỏi nhiều công sức và thời gian.

2. CÁC TÀI LIỆU TRA CỨU

Tài liệu tra cứu là những công cụ đầu tiên của việc tìm tài liệu. Các tài liệu này hoặc cho ta những chỉ dẫn hoặc chỉ rõ nguồn thông tin được cung cấp, hoặc xác định nội dung và giới hạn của các vấn đề đặt ra. Có nhiều loại tài liệu tra cứu tương ứng với các yêu cầu khác nhau.

Đặc điểm chung của các tài liệu tra cứu: Đó là các tài liệu cấp hai, được biên tập từ tài liệu gốc. Nó không chứa những tri thức mới mà chỉ sắp xếp lại các tri thức đã có.

Người ta phân biệt các loại tài liệu tra cứu sau:

- Các thư mục và mục lục, nó chỉ dẫn đến các tài liệu.

- Các bách khoa toàn thư và các từ điển, nó chỉ dẫn đến các ý tưởng hay các từ chuyên môn.

- Các bản danh mục tra cứu, nó chỉ dẫn đến các tên, các địa chỉ và các thông tin thường thức.

Ngoài ra còn có danh mục của các tài liệu tra cứu như: danh mục các tài liệu tra cứu trong năm, thư mục của các thư mục, chỉ dẫn về các bách khoa toàn thư hiện có v.v...

Thư mục

Thư mục là bản liệt kê đầy đủ hay có chọn lọc các tài liệu về một chủ đề nào đó hay liệt kê định kỳ các tài liệu mới. Trên đó ghi các thông tin về các đặc trưng bên ngoài của tài liệu như: tên tác giả, tên tài liệu, nguồn gốc và dạng của tài liệu, ngôn ngữ, các yếu tố xuất bản v.v... Có nhiều loại thư mục: thư mục quốc gia, thư mục chuyên ngành, thư mục các ấn phẩm định kỳ...

Thư mục quốc gia thường do thư viện quốc gia biên soạn, đó là danh sách các ấn phẩm, văn bản hay không văn bản, được xuất bản trong một nước mà thư viện quốc gia nhận được theo luật lưu chiếu. Thư mục quốc gia được xuất bản theo định kỳ, một tuần, một tháng hay ba tháng một lần. Ở các nước có số lượng xuất bản phẩm lớn, thư mục quốc gia được biên tập và lưu trữ trong máy tính điện tử dưới dạng các cơ sở dữ liệu thư mục.

Mục đích của thư mục quốc gia là cho phép người dùng tin thường xuyên nắm bắt được nguồn tài liệu xuất bản của một quốc gia, đồng thời phục vụ cho việc tìm tin hồi cố.

Thư mục chuyên ngành trình bày các tài liệu về một chủ đề xác định. Chúng tập hợp tất cả các loại tài liệu như: sách,

tạp chí, luận án, bằng phát minh sáng chế, tài liệu hội nghị... liên quan đến một chủ đề, được xuất bản ở nhiều nước, bằng nhiều ngôn ngữ khác nhau.

Thư mục các bài báo của các ấn phẩm định kỳ nói chung là thư mục chuyên ngành. Chúng tập hợp các bài báo đăng trong các ấn phẩm định kỳ, có nội dung đề cập đến một chủ đề xác định.

Để phục vụ cho việc lưu trữ thông tin và tìm tin, nội dung của các thư mục này có thể có cả bản tóm tắt, chú giải và tập các từ khoá mô tả nội dung tài liệu. Chúng có thể sắp xếp theo vần chữ cái, theo hệ thống phân loại hay theo thứ tự nhập của tài liệu. Khi đó ta có thể tiếp cận các tài liệu theo tác giả, theo chủ đề, theo các yếu tố xuất bản, theo từ khoá hay theo loại hình tài liệu.

Mục lục

Nhờ các chỉ dẫn thư mục người ta có thể mô tả tài liệu trên một phiếu theo một quy tắc nhất định, gọi là phiếu mục lục. Phiếu mục lục xác định vị trí của tài liệu trong kho. Mỗi tài liệu trong đơn vị thông tin đều phải được mô tả trong phiếu mục lục của nó.

Mục lục là tập hợp các phiếu mục lục của tất cả các tài liệu có trong một đơn vị thông tin, được trình bày theo một quy tắc nhất định.

Mục lục có thể được sắp xếp theo vần chữ cái, theo hệ thống của khung phân loại hay theo thứ tự nhập của tài liệu.

Hệ thống mục lục giúp cho việc quản lý vốn tài liệu và xác định vị trí của tài liệu trong kho. Đó là công cụ tra cứu quan trọng bậc nhất trong các thư viện truyền thống.

Các phiếu mục lục có thể tập hợp và in trong một ấn phẩm mục lục để phân phối cho các đơn vị thông tin và tạo thuận lợi cho việc tìm tài liệu của người dùng tin.

Trong các thư viện truyền thống có nhiều loại mục lục, có mục lục tác giả, mục lục chủ đề hệ thống, mục lục chủ đề chữ cái, mục lục xếp kho v.v...

Mục lục của các ấn phẩm định kỳ thường được thực hiện trên các phiếu đặc biệt, gọi là Kardex, cho phép theo dõi các số liên tiếp của ấn phẩm. Các đơn vị thông tin thường có thư mục riêng cho loại tài liệu này và chúng thường được quản lý một cách tự động.

Mục lục liên hợp là tập các mục lục của nhiều thư viện. Các mục lục liên hợp thường được tổ chức theo chủ đề hay theo loại hình tài liệu và thường được xuất bản dưới dạng một tuyển tập với sự tham gia của các đơn vị thành viên. Ví dụ The National Union Catalog là mục lục liên hợp về sách của Thư viện Quốc hội Mỹ và các thư viện của Hoa Kỳ.

Các catalog thương mại và công nghiệp chính là mục lục về các sản phẩm của một đơn vị sản xuất. Chúng thường mô tả tóm tắt về các đặc tính của sản phẩm và cung cấp các thông tin cần thiết về sản phẩm đó.

Ngày nay với việc phát triển ứng dụng tin học trong hoạt động thông tin - thư viện, mục lục có thể trình bày dưới dạng đọc được bằng máy và từ đó hình thành khái niệm mục lục đọc được bằng máy (machine readable catalog). Mục lục đọc được bằng máy lưu trữ trong các CSDL thư mục. Từ các CSDL này người ta có thể in ra các loại phiếu mục lục khác nhau một cách tự động, nhờ các format in mà người sử dụng có thể tạo ra.

Từ điển

Từ điển là tài liệu dùng để tra cứu các thuật ngữ, các tên gọi. Có nhiều loại từ điển khác nhau.

Từ điển của một ngôn ngữ là bộ sưu tập các từ của một ngôn ngữ, được sắp xếp theo vần chữ cái và được giải thích bằng chính ngôn ngữ ấy. Ví dụ: Oxford English Dictionary là bộ từ điển nổi tiếng của Anh gồm 20 tập (năm 1989), bộ từ điển Le Grand Robert nổi tiếng của Pháp gồm 9 tập đều là từ điển của một ngôn ngữ.

Từ điển song ngữ hay đa ngôn ngữ là bộ sưu tập các từ của một ngôn ngữ và được dịch sang một hay nhiều ngôn ngữ khác.

Từ điển tiểu sử trình bày vắn tắt cuộc đời và sự nghiệp của các nhân vật nổi tiếng. Ví dụ: International world who's who, Who's who in data processing, Who's who in Africa...

Từ điển tiểu sử dùng để kiểm tra tên của các nhân vật và xác định xem đó có thể là nguồn thông tin về một vấn đề nào đó hay không.

Bách khoa toàn thư

Bách khoa toàn thư trình bày dưới dạng các bài viết ngắn hiện trạng tri thức của nhân loại về tất cả các lĩnh vực (Bách khoa toàn thư phổ dụng) hoặc về một chủ đề nào đó (Bách khoa toàn thư chuyên ngành).

Trong bách khoa toàn thư các chủ đề được biên tập và sắp xếp theo vần chữ cái. Tuy nhiên cũng có bách khoa toàn thư sắp xếp các chủ đề theo hệ thống của một khung phân loại.

Bách khoa toàn thư khác từ điển ở chỗ nó không chứa tất cả các từ hay khái niệm của một ngôn ngữ trên tất cả các lĩnh vực hay một chủ đề cho trước, mà là sự chọn lọc các đề mục thiết yếu và được tổ chức biên tập một cách kỷ lưỡng bởi các chuyên gia đầu ngành. Người ta sử dụng bách khoa toàn thư để tra cứu nội dung của các thuật ngữ cũng như các vấn đề cần nghiên cứu.

Còn có các cẩm nang kỹ thuật tập hợp các dữ liệu về một lĩnh vực khoa học như Chemical engineer' s handbook hay Standard handbook for civil engineers.

Ngày nay với sự phát triển của công nghệ CD-ROM, nhiều bách khoa toàn thư hàng chục tập có thể lưu trữ trong một đĩa CD-ROM, với thông tin được trình bày rất sinh động dưới dạng chữ viết, âm thanh, hình ảnh tĩnh và động và khả năng truy cập rất thuận lợi và linh hoạt.

Ví dụ: Từ điển bách khoa Encarta CD của hãng Microsoft là một bách khoa toàn thư lưu trữ thông tin đa phương tiện, với 30.000 bài mục, có bài mục dài tới 3000 từ, được trình bày thẩm mỹ và truy cập rất thuận lợi. Encarta CD chứa hàng nghìn tranh ảnh màu và những chuỗi âm thanh nếu được nối lại sẽ dài tới chín tiếng đồng hồ. Từ điển bách khoa Britannica của Anh được xuất bản lần thứ nhất năm 1768, lần xuất bản gần đây nhất gồm 32 tập lớn, do bốn ngàn học giả thuộc hơn một trăm quốc gia biên soạn, cũng đã được ghi trên ba đĩa CD. Britannica CD 99 có 73000 bài mục, trong đó có hơn 700 bài mục lớn giới thiệu có hệ thống từng chủ đề lớn. Mỗi bài mục lớn thực sự là tập sách 4, 5 chục trang với những thông tin súc tích, hệ thống và rất có giá trị. Ngoài ra Britannica CD còn có 1200 bản đồ và 8500 hình ảnh minh hoạ cùng với 1,5

triệu kết nối siêu văn bản (hypertext links) cho phép ta truy cập thông tin một cách nhanh chóng và thuận tiện.

Các danh mục

Các danh mục là các tài liệu tra cứu cung cấp các chỉ dẫn về các cá nhân, các tổ chức, các thông tin thường thức, các tài liệu trong một lĩnh vực cho trước, các sách mới xuất bản...

Người ta phân biệt các loại sách tra cứu sau:

- Danh mục các sách đã xuất bản, cho biết tên các sách đang có bán trên thị trường. Ví dụ: Les livres disponibles en 1977 tập hợp 220000 tên sách bằng tiếng Pháp xuất bản trong năm 1977 của 43 nước.

- Danh mục các sách sắp xuất bản, cho phép biết những sách sẽ xuất bản trên một chủ đề nào đó. Ví dụ: British books in print hay The reference catalog of current literature.

- Danh mục các cơ quan, đôi khi còn gọi là niên giám (annuaire), cung cấp địa chỉ và những thông tin nhanh về các cơ quan quốc gia hay quốc tế đang làm việc trên một lĩnh vực. Ví dụ: Répertoire mondial des institutions en sciences sociales/World index of social sciences institutions.

- Danh mục các nhân vật, các thành viên của các hiệp hội hay các tổ chức nghề nghiệp, cung cấp tên, địa chỉ của các cá nhân đang hoạt động trong một ngành nào đó, cùng với chuyên môn và chức năng của họ. Nhưng khác với từ điển tiểu sử, nó không giới thiệu cuộc đời và sự nghiệp của họ.

- Danh mục các dự án, danh mục các công trình nghiên cứu đang triển khai hay đã hoàn thành, thường do trung tâm quốc gia, các cơ quan tài trợ xuất bản.

- Danh mục các luận án đã bảo vệ hay đang tiến hành nghiên cứu, do các viện nghiên cứu hay các trường đại học xuất bản. Cũng có danh mục các luận án trên một lĩnh vực.

- Danh mục các hợp đồng về kinh tế hay nghiên cứu thuộc các lĩnh vực khác nhau của các đơn vị trong một quốc gia.

3. TÀI LIỆU KHOA HỌC KỸ THUẬT VÀ QUY LUẬT PHÁT TRIỂN CỦA NÓ

3.1. Tài liệu khoa học kỹ thuật và vai trò của nó

Kết quả của những hoạt động nghiên cứu khoa học kỹ thuật được phản ánh bằng những tài liệu dưới những hình thức khác nhau (báo cáo, biên bản, nhật ký khoa học, sách, ấn phẩm định kỳ, vật triển lãm, sản phẩm mẫu...) được gọi chung là các tài liệu khoa học kỹ thuật. Có thể nói tài liệu khoa học kỹ thuật là đối tượng vật chất ghi cố định các thông tin khoa học kỹ thuật để lưu truyền trong không gian và thời gian.

Tài liệu khoa học kỹ thuật là bộ phận không thể tách rời trong hoạt động khoa học kỹ thuật, là thành phần quan trọng nhất trong hệ thống giao lưu khoa học.

Vai trò đó thể hiện trên các mặt sau đây:

- Tài liệu khoa học kỹ thuật là công cụ ghi cố định các kết quả của các công trình nghiên cứu và triển khai, ghi lại các sáng chế phát minh, các kinh nghiệm sản xuất tiên tiến, các công nghệ mới được áp dụng, các kết quả thí nghiệm.

- Tài liệu khoa học kỹ thuật là phương tiện tốt nhất để lưu trữ và phổ biến thông tin khoa học kỹ thuật theo không gian và thời gian.

- Tài liệu khoa học kỹ thuật là phương tiện chính để xác nhận quyền sở hữu các kết quả nghiên cứu khoa học, xác nhận quyền sáng chế phát minh và do đó là phương tiện để khuyến khích sáng tạo khoa học.

- Số lượng và chất lượng của tài liệu khoa học kỹ thuật là chỉ tiêu tổng hợp đánh giá thành quả hoạt động sáng tạo khoa học kỹ thuật và đánh giá trình độ khoa học của một đất nước.

3.2. Quy luật phát triển của tài liệu khoa học kỹ thuật

Sự phát triển của tài liệu khoa học kỹ thuật chịu ảnh hưởng trực tiếp của cuộc cách mạng khoa học, kỹ thuật và công nghệ đang diễn ra mạnh mẽ trên thế giới trong mấy chục năm gần đây, với những đặc điểm là:

- Khoa học đã chuyển từ khoa học nhỏ thiên về lý thuyết (trong các trường đại học, các phòng thí nghiệm do một số nhà khoa học tiến hành) sang khoa học lớn, khoa học kiểu công nghiệp, gắn liền với sản xuất.

Đặc điểm trên thể hiện ở các mặt sau:

Thiết bị nghiên cứu ngày càng tinh vi, phức tạp, đắt tiền. Chi phí cho nghiên cứu khoa học tăng lên rất nhanh, chiếm tỷ lệ đáng kể trong tổng thu nhập quốc dân (5% hoặc hơn). Đội ngũ những người tham gia vào các công việc sáng tạo khoa học gia tăng nhanh chóng. Ở các nước tư bản phát triển

nhà nước tham gia đầu tư vào các công trình nghiên cứu khoa học ngày càng nhiều.

Tình hình đó ảnh hưởng trực tiếp đến hệ thống giao lưu khoa học. Các tài liệu khoa học kỹ thuật gia tăng nhanh chóng theo quy luật hàm số mũ.

- Song song với việc phân hoá theo chuyên môn, đang hình thành xu hướng liên kết thống nhất các khoa học lại.

Quá trình phân hoá các khoa học là hệ quả tất yếu của sự phát triển khoa học. Tri thức của loài người càng mở rộng, càng đi sâu thì người ta càng thấy rõ sự khác nhau giữa những việc trước kia gần nhau. Hơn nữa số lượng tri thức tích lũy ngày một nhiều thì một người dù thông minh đến đâu cũng không nắm giữ được hết.

Nhưng bên cạnh quá trình phân hoá, còn có quá trình thống nhất các khoa học. Khi sự hiểu biết trong từng lĩnh vực chuyên môn đạt tới một trình độ sâu sắc nào đó thì người ta phát hiện ra những mối liên hệ giữa những hiện tượng tưởng như rời rạc, do đó nảy sinh những quan niệm tổng quát, liên kết một số ngành chuyên môn tách biệt thành những khoa học mới. Hiện tượng đó thường gọi là sự xâm nhập lẫn nhau giữa những ngành khoa học khác nhau. Ngày nay thật khó xác định ranh giới giữa vật lý và hoá học, giữa hoá học và sinh học...

Đặc điểm trên đây đã ảnh hưởng trực tiếp đến cơ cấu của tài liệu khoa học kỹ thuật, tạo nên quy luật tập trung và tản mạn thông tin của tài liệu khoa học kỹ thuật. Đồng thời gây nên hiện tượng khủng hoảng thông tin, thể hiện ở các mâu thuẫn giữa khối lượng tài liệu tăng nhanh và khả năng xử lý

có hạn của con người, giữa nhu cầu thông tin và khả năng thu nhận được những thông tin phù hợp.

Các quy luật phát triển của tài liệu khoa học kỹ thuật ngày nay thể hiện ở các mặt sau đây:

a. Quy luật về sự gia tăng tài liệu

Cùng với sự phát triển mạnh mẽ của khoa học và công nghệ, đội ngũ những người làm khoa học tăng lên nhanh chóng. Người ta dự đoán rằng số cán bộ khoa học kỹ thuật ngày nay chiếm tới 90% tổng số các nhà khoa học của nhân loại có từ trước đến nay. Hệ quả tất yếu là các sản phẩm của họ, các tài liệu khoa học kỹ thuật cũng tăng lên nhanh chóng. Người ta thấy rằng cứ trong vòng từ 12 đến 15 năm, số lượng tài liệu khoa học kỹ thuật lại tăng lên gấp đôi.

Ngày nay, mỗi ngày trên thế giới có khoảng 2000 trang tạp chí khoa học kỹ thuật được in ra, mỗi năm có tới 5 triệu bài báo được đăng trên các tạp chí khoa học xuất bản định kỳ. Người ta cũng dự báo trong vòng 20 năm tới không có dấu hiệu nào phát triển chậm lại.

Do tăng nhanh về số lượng nên chất lượng của tài liệu khoa học kỹ thuật bị giảm sút. Để khắc phục điều đó các cơ quan thông tin phải tăng cường khâu chọn lọc và xử lý tài liệu, xây dựng các hệ thống thông tin tự động hoá và mở rộng mạng lưới thông tin quốc gia và quốc tế.

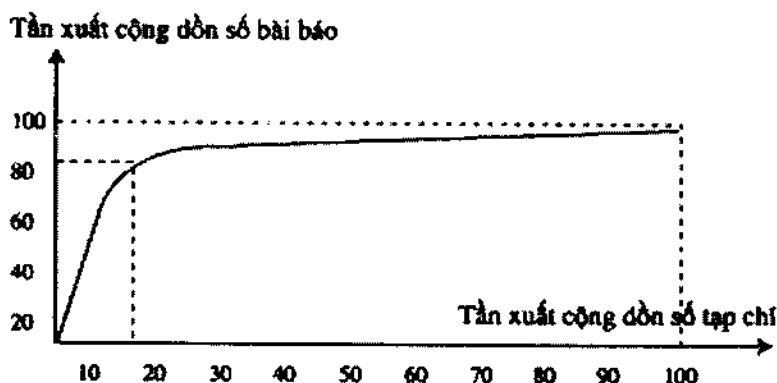
b. Quy luật về sự tập trung và tản mạn thông tin

Quy luật này hình thành do đặc điểm phân hoá và tổng hợp khoa học của sự phát triển khoa học hiện đại. Quá trình phân hoá theo chuyên môn dẫn đến hình thành các tài liệu

theo chuyên môn hẹp. Quá trình liên kết các khoa học hình thành các tài liệu khoa học liên ngành.

Năm 1930, qua thống kê nhà thư viện học người Anh Bradford thấy rằng: Nếu sắp xếp số tạp chí khoa học theo thứ tự giảm dần số bài báo về một chuyên ngành nào đó thì trong danh sách nhận được ta có thể tìm thấy các “tạp chí hạt nhân”. Số tạp chí này không lớn, chỉ chiếm khoảng 10 - 15% số tạp chí, nhưng chứa đựng tới 90% số bài báo liên quan đến ngành đó.

Mối quan hệ đó có thể biểu diễn bằng biểu đồ sau:



Hình 9. Biểu đồ xác định tạp chí hạt nhân

Nhìn vào biểu đồ ta thấy chỉ có khoảng 10 - 15% số tạp chí là có hiệu suất cao, chúng chứa đựng hầu hết các bài báo về một chuyên ngành nào đó.

Nhờ xác định các tạp chí hạt nhân, người ta xây dựng được danh mục các tạp chí cần mua một cách hợp lý và tránh được lãng phí. Qua thực tế ta thấy tỉ lệ sau đây là thích hợp:

- Từ 10 đến 15% tạp chí chuyên ngành.
- Từ 25 đến 30% tạp chí liên quan.

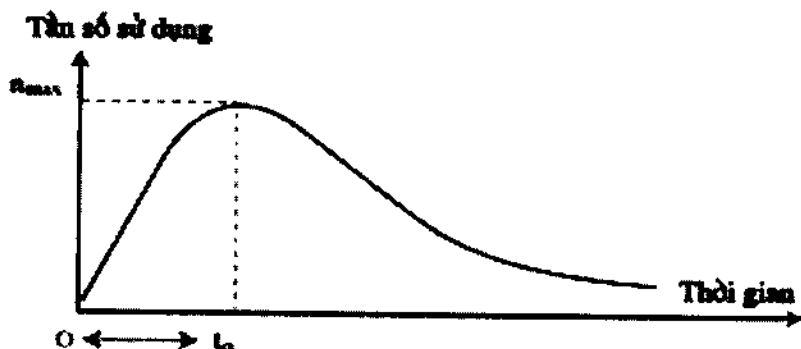
c. Quy luật về thời gian hữu ích và tính lỗi thời của tài liệu

Thời gian hữu ích hay tuổi thọ của tài liệu khoa học kỹ thuật phụ thuộc vào lĩnh vực tri thức và giá trị nội dung của tài liệu, cũng phụ thuộc vào tính thời sự và khả năng tương thích của nó đối với tình trạng của tri thức và nhu cầu của người dùng tin.

Chẳng hạn một tài liệu về triết học vẫn còn giữ nguyên giá trị trong hàng thế kỷ, nhưng một bản vẽ về một cỗ máy thì chỉ có giá trị trong một vài năm. Một số tài liệu có tuổi thọ rất ngắn, chúng thường mất ngay giá trị sau lần xuất bản mới. Chẳng hạn như các niên giám, các tiêu chuẩn, các bảng tra cứu. Tuổi thọ của các tạp chí cũng phụ thuộc vào thực tế đương thời. Nói chung một quyển sách có tuổi thọ dài hơn những bài báo của các ấn phẩm định kỳ.

Bằng cách thống kê người ta có thể tính được tần số sử dụng, tần số trích dẫn của một tài liệu riêng biệt hay một loại tài liệu nào đó. Người ta gọi “nửa đời của tài liệu” là thời gian từ lúc công bố đến lúc nó được sử dụng nhiều nhất, sau đó giá trị sử dụng giảm dần. Tuổi thọ của tài liệu được tính từ lúc công bố đến lúc lỗi thời, không được sử dụng nữa.

Sự phát triển nhanh của khoa học kỹ thuật và công nghệ làm cho tài liệu khoa học kỹ thuật nhanh chóng trở nên lỗi thời. Nhưng sự lỗi thời của tài liệu phụ thuộc vào lĩnh vực của tri thức và nội dung của tài liệu. Khoa học nào càng ổn định thì giá trị sử dụng của nó càng lâu hơn.



Hình 10. Nửa đời của tài liệu: $0 - t_0$

Nói chung các tài liệu được sản sinh theo một yêu cầu nào đó thường giảm dần giá trị sử dụng cùng với sự phát triển của yêu cầu này cho đến khi trở nên lỗi thời. Điều này khác hẳn với các tài liệu nguyên khai, các vật khảo cổ hay các tài liệu có giá trị lịch sử.

Ta cũng không nên nhầm lẫn giữa tần số sử dụng với giá trị sử dụng của tài liệu. Vì có thể có những tài liệu ít được sử dụng, đôi khi khá cũ, nhưng lại có lợi ích đáng kể đối với một vài đối tượng hay mục đích sử dụng nào đó.

4. NGUỒN THÔNG TIN ĐIỆN TỬ – THÔNG TIN SỐ

Việc ứng dụng công nghệ thông tin trong hoạt động thông tin tư liệu, đặc biệt việc sử dụng kỹ thuật số để biểu diễn thông tin đã dẫn đến việc ra đời một nguồn thông tin mới, đó là nguồn thông tin điện tử (Electronic Information Resources) hay còn gọi là nguồn thông tin số (Digital Information Resources). Và từ đó cũng xuất hiện khái niệm thông tin điện tử hay thông tin số.

Có thể coi thông tin điện tử là tất cả các thông tin được xử lý, lưu trữ và truy cập trên máy tính hay trên mạng máy tính. Thông tin điện tử được trình bày và lưu trữ trên các vật mang tin điện tử. Đó là các băng từ, đĩa từ, đĩa quang. Chúng tạo thành nguồn tài liệu điện tử.

Các cơ sở dữ liệu do đơn vị thông tin xây dựng là thông tin điện tử quan trọng có thể truy cập trực tiếp trên máy tính. Nhưng nguồn thông tin điện tử dồi dào nhất là nguồn thông tin trực tuyến được truy cập trên máy tính thông qua mạng Internet hoặc các mạng máy tính khác.

Thông tin điện tử bao gồm:

- Các cơ sở dữ liệu chuyên ngành, đa ngành lưu trữ trên các đĩa từ, băng từ, đĩa quang CD-ROM.

- Các cơ sở dữ liệu trực tuyến do các cơ quan thông tin xây dựng, muốn sử dụng phải đăng ký với một server để được quyền truy cập.

- Bản tin điện tử, do một cơ quan phát hành, thường đặt trong trang chủ của đơn vị mình trên mạng Intranet, Internet.

- Báo và tạp chí điện tử, được ấn hành trên mạng Internet.

- Các Website trên Internet, chứa thông tin về cơ quan hành chính sự nghiệp, các doanh nghiệp, các công ty, các trường đại học, ... Việc truy cập thường là miễn phí.

- Các sơ sở tri thức, chứa hệ thống các luật dùng để xác định và thay đổi các mối liên hệ giữa các sự kiện đầu vào.

Các yêu cầu đối với một tài liệu điện tử là:

- Thiết kế trình bày đẹp, thẩm mỹ.

- Dễ đọc, dễ xem. Có thể đọc tuần tự, có thể điều khiển di chuyển dễ dàng trên tài liệu thông qua cấu trúc logic và các mối liên kết mở.

- Bảo đảm tính toàn vẹn của biểu đồ, minh hoạ, công thức toán học.

- Có khả năng tìm kiếm toàn văn bên trong tài liệu.

- Có khả năng in toàn bộ hoặc một phần của tài liệu.

Thông tin điện tử có những đặc trưng nổi bật sau đây:

- Có mật độ thông tin cao. Chỉ cần từ 4 đến 5 đĩa CD-ROM có thể lưu trữ toàn bộ nội dung trong một năm của bộ *Chemical Abstract* gồm 100 tập, mỗi tập 2000 trang.

- Thông tin luôn mới nhờ có khả năng cập nhật nhanh, thường xuyên và kịp thời.

- Thông tin có thể lưu trữ ở nhiều dạng khác nhau: văn bản, đồ thị, âm thanh, hình ảnh tĩnh và động.

- Có khả năng truy cập theo nhiều dấu hiệu khác nhau: tác giả, nhan đề, từ khoá, năm xuất bản, ...

- Có khả năng truy cập từ xa, không giới hạn về không gian, thời gian.

- Cùng một thời điểm có thể nhiều người truy cập.

- Tạo khả năng cho người dùng tin có thể tiếp cận với tác giả thông qua kênh thông tin phản hồi.

Bên cạnh những ưu điểm trên đây, thông tin điện tử cũng có những nhược điểm và hạn chế:

- Tính ổn định không cao, độ bền vững không đồng nhất: có thông tin tồn tại lâu dài như các thông tin ghi trên CD-ROM, nhưng cũng có những thông tin có vòng đời rất ngắn như các bản tin, bài báo trên mạng Internet.

- Dễ bị vi phạm bản quyền, do bị sao chép, sửa đổi, thậm chí bị huỷ hoại do những vi phạm cố tình.

Với sự phát triển của công nghệ thông tin và truyền thông, các nguồn tài liệu điện tử phát triển rất nhanh. Chỉ riêng tạp chí điện tử năm 1991 mới có 110 tên tạp chí điện tử, thì năm 1995 con số đó đã là 700, năm 1998 là khoảng 5000, năm 2002 số tên tạp chí điện tử đã lên tới trên 6000. Người ta dự báo trong 5 năm tới có thể tất cả các tạp chí khoa học sẽ được xuất bản dưới dạng điện tử. Ngày nay thông tin điện tử giữ vai trò quan trọng trong giao lưu thông tin và trong mọi lĩnh vực hoạt động của con người.

Sự ra đời của nguồn thông tin điện tử đã làm thay đổi về chất hoạt động giao lưu thông tin nói chung và hoạt động thông tin – thư viện nói riêng. Các thư viện điện tử ra đời với phần cốt lõi là các kho thông tin số hoá. Xuất bản điện tử ra đời và phát triển đã làm thay đổi ngành xuất bản từ nhiều năm nay. Ngày nay hầu hết các tài liệu khoa học và công nghệ được xuất bản và phổ biến dưới dạng điện tử. Theo số liệu năm 2002, trên thế giới 61% thông tin khoa học và công nghệ, 41% thông tin y học được xuất bản dưới dạng số.

Sự phát triển bùng nổ của nguồn thông tin số hoá trong những năm gần đây đã dẫn đến hình thành khái niệm “*nội dung số*”, tiếng anh là *E-content* hay *Digital Content*.

Nội dung số là thuật ngữ dùng để chỉ các thông tin số hoá. Nó có thể ở dạng trang Web cũng như thông tin chứa trong các

tệp dữ liệu ở dạng văn bản, âm thanh, hình ảnh hoặc dạng tích hợp hoà trộn văn bản, âm thanh, hình ảnh (đa phương tiện).

Nội dung số là một khái niệm rộng, bao quát nhiều lĩnh vực và phát triển liên tục, bao gồm phần mềm nội dung, trò chơi điện tử, học tập trực tuyến (e-learning), xuất bản điện tử, âm nhạc, chương trình phát thanh truyền hình, quảng cáo tiếp thị trên Internet, v.v...

Có cả một ngành công nghiệp cho các sản phẩm nội dung số. Đó là *công nghiệp nội dung số (Digital Content Industry – DCI)*. DCI là một khái niệm rất mới và trên thế giới người ta vẫn chưa tìm được một định nghĩa thống nhất.

Theo quan điểm của Bộ Bưu chính – Viễn thông: Công nghiệp nội dung số 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, thư viện và bảo tàng số, phát triển nội dung cho mạng băng rộng, quảng cáo tiếp thị trên Internet, v.v...

5. XÂY DỰNG VỐN TÀI LIỆU - CHỌN LỌC VÀ BỔ SUNG

Chọn lọc là lựa chọn các tài liệu mà đơn vị thông tin cần phải có. Bổ sung là phương thức cho phép nhận được các tài liệu đó. Hai công việc này đặt ngay ở đầu dây chuyền tư liệu, cho phép tạo lập và duy trì vốn tư liệu của đơn vị thông tin, nhằm đáp ứng nhu cầu của người dùng tin. Chọn lọc là công việc đòi hỏi phải có tri thức rộng và phải do người có năng lực và am hiểu người dùng tin thực hiện. Còn bổ sung là một công việc có tính chất nghiệp vụ, nó đòi hỏi phải có phương pháp và tổ chức tốt.

Việc chọn lọc và bổ sung bao gồm các yếu tố sau:

- Chính sách và kế hoạch bổ sung.
- Tiếp cận các nguồn tài liệu.
- Hình thức và phương thức bổ sung.

4.1. Chính sách bổ sung

Ngày nay số lượng các tài liệu xuất bản gia tăng rất nhiều. Không một thư viện nào, thậm chí là thư viện lớn nhất cũng không thể thu thập được mọi tài liệu; vả lại trong nhiều lĩnh vực chuyên môn hẹp, người ta cần có những bộ sưu tập tài liệu riêng. Vì vậy việc bổ sung tài liệu cho một đơn vị thông tin không thể làm một cách tùy tiện, mà phải thực hiện theo một chính sách xác định. Đó là một tập hợp các nguyên tắc và tiêu chuẩn cho việc lựa chọn hay loại bỏ tài liệu.

Một chính sách như thế không thể thiếu được nhằm:

- Xác lập các tiêu chuẩn để lựa chọn và thanh lọc tài liệu.
- Xác định nguồn tài liệu.
- Điều chỉnh các phương thức tiếp nhận tài liệu.
- Sử dụng có hiệu quả nguồn kinh phí mà đơn vị thông tin được cấp.
- Duy trì sự hợp tác về tài liệu giữa các đơn vị thông tin, tránh lãng phí.

Chính sách thông tin được xây dựng trên cơ sở các yếu tố sau đây:

- Khả năng ngân sách và tiềm năng của đơn vị thông tin bao gồm: tổng số vốn, số lượng và chất lượng đội ngũ cán bộ,

các phương tiện kỹ thuật vì người ta không chỉ tiếp nhận nguồn tài liệu mà còn phải xử lý chúng.

- Chức năng chuyên môn của đơn vị thông tin, tức là lĩnh vực chuyên môn mà đơn vị thông tin quan tâm. Cần phải xác định lợi ích và tính thích ứng của các tài liệu theo đối tượng người dùng tin.

- Những mục tiêu thường xuyên và mục tiêu ưu tiên của đơn vị thông tin.

- Vị trí và tính chất phục vụ của đơn vị thông tin và đối tượng người dùng tin.

- Mối quan hệ với các đơn vị thông tin khác.

Nội dung của chính sách bổ sung bao gồm:

- Xác định các loại hình tài liệu, tỉ lệ giữa các loại hình tài liệu đó theo yêu cầu của người dùng tin.

- Xác định các nội dung và chủ đề của tài liệu. Các tài liệu phải bao quát các lĩnh vực mà đơn vị thông tin quan tâm (diện của vốn tư liệu).

- Sự thích ứng của các tài liệu với lợi ích của người dùng tin.

- Ngôn ngữ của các ấn phẩm.

- Mức độ phổ cập của các tài liệu thông thường, tài liệu quý hiếm, tài liệu không công bố.

- Niên hạn của tài liệu.

- Các tài liệu tặng biếu và trao đổi.

- Các tiêu chuẩn cho sự thanh lọc tài liệu đã lỗi thời.

4.2. Cách tiếp cận các nguồn tài liệu

Việc tiếp cận các nguồn tài liệu được thực hiện nhờ nhiều nguồn thông tin bổ trợ khác nhau: các cá nhân, các tổ chức hay các tài liệu. Yếu tố thứ ba này chỉ khác ở chỗ các tài liệu cần tìm là tài liệu công bố hay không công bố.

- Các chuyên gia, các nhà khoa học thể hiện một nguồn thông tin quan trọng mà các đơn vị thông tin cần phải duy trì mối quan hệ thường xuyên. Người ta gọi đó là những đồng nghiệp vô hình của đơn vị thông tin. Trong từng lĩnh vực chuyên môn, họ luôn là những người sản sinh ra những tài liệu, đề cập tới những tri thức mới nhất mà người ta quan tâm. Người ta có thể tìm đến các nhà chuyên môn bằng con đường trực tiếp như: tiếp xúc thường xuyên với các tác giả, các cơ quan nghiên cứu, các cán bộ thông tin chuyên ngành, gặp gỡ ở các hội nghị quốc gia hay quốc tế. Người ta cũng có thể tiến hành một cách gián tiếp, thông qua các mục lục chuyên ngành.

- Việc tiếp cận các tài liệu công bố có thể thực hiện bằng cách sử dụng các nguồn và các tài liệu chỉ dẫn khác nhau như: các dịch vụ làm tóm tắt và đánh chỉ số, thư mục quốc gia và thư mục chuyên ngành, thư mục in ở trang cuối các tài liệu cấp một, mục lục và các tủ phiếu của các đơn vị thông tin khác, các danh mục thông báo sách sắp xuất bản, các tài liệu giới thiệu các sách đã xuất bản, các ấn phẩm thương mại và quảng cáo, danh mục các ấn phẩm định kỳ, các xuất bản phẩm chính thức, các tạp chí hàng năm, các danh mục chọn lọc liên quan tới các nhà xuất bản hay cơ quan xuất bản lớn, các báo cáo tổng kết... Trong số này thư mục quốc gia được xây dựng dựa trên các tài liệu bổ sung của thư viện quốc gia

là có vị trí đặc biệt. Thư mục quốc gia của Anh (British National Bibliography), được xây dựng từ năm 1950 ở Bảo tàng Anh là một ví dụ: nó được xuất bản hàng tuần, tích lũy đều đặn các tài liệu mới xuất bản và trở thành một công cụ để tìm tài liệu theo chủ đề cũng như cho sự lựa chọn tài liệu.

Hiệp hội các thư viện quốc tế (IFLA) đã lập một chương trình nhằm thúc đẩy sự gia tăng các công cụ thư mục như thế. Chương trình này gọi là chương trình Kiểm soát thư mục toàn thể và MARC quốc tế (Universal Bibliographic Control và International MARC) nhằm động viên thư viện quốc gia các nước xuất bản các thư mục quốc gia theo một khổ mẫu chuẩn và nhập chúng vào máy tính ở những nơi có thể. Nhờ đó người ta có thể tiếp cận các nguồn tài liệu bằng các hệ thống thông tin trực tuyến.

- Việc tiếp cận các tài liệu không công bố đặt ra những khó khăn riêng, đòi hỏi phải có sự cố gắng đặc biệt. Người ta có thể tìm những tài liệu này bằng hai cách:

Trực tiếp bằng tiếp xúc cá nhân với các tác giả, các cơ quan xuất bản tài liệu này, các cá nhân thường được thông báo những thông tin mới nhất trong từng lĩnh vực.

Gián tiếp bằng cách tham khảo các thư mục của các ấn phẩm, các bản luận văn, tra cứu các mục lục chỉ dẫn của các cơ quan quốc gia và quốc tế như UNESCO, FAO....

Việc thăm dò và bổ sung những tài liệu này đòi hỏi sự cố gắng lớn của đơn vị thông tin. Giá trị của loại tài liệu này thể hiện ở thời gian và công sức dùng để tìm kiếm các tài liệu đó.

4.3. Các phương thức và thủ tục bổ sung

Phương thức bổ sung

Có hai phương thức bổ sung: bổ sung phải trả tiền và bổ sung không mất tiền.

- Việc bổ sung phải trả tiền hay mua tài liệu được thực hiện bằng hai cách:

Trực tiếp ở các nơi sản xuất tài liệu: tác giả, nhà xuất bản. Phương thức này nhanh, nhưng đòi hỏi những công việc quản lý ngân sách và dự trù đặt hàng.

Gián tiếp thông qua các cơ quan phát hành sách báo. Các cơ quan này lo toàn bộ các khâu kỹ thuật và tài chính. Phương thức này thường áp dụng đối với việc đặt mua các tài liệu nước ngoài, các tài liệu mua với khối lượng lớn và thường xuyên.

- Việc bổ sung không mất tiền có thể thực hiện bằng nhiều cách:

Bằng trao đổi tài liệu giữa cơ quan thông tin này và cơ quan thông tin khác. Lợi ích của phương thức này là tránh phải chi tiền mặt và có khả năng trao đổi tất cả các loại hình tài liệu. Điều bất lợi là các tài liệu nhận được không phải bao giờ cũng đáp ứng với yêu cầu và giá trị tài liệu trao đổi.

Bằng các tài liệu tặng biếu ở các dạng khác nhau: các tài liệu biếu đột xuất hay thường xuyên của các sử quán, các cơ quan thương mại, các tác giả. Lưu chiếu cũng là dạng tài liệu biếu được quy định bằng luật đối với các nhà xuất bản. Họ phải cung cấp một số bản ấn phẩm cho thư viện quốc gia hay cơ quan lưu trữ quốc gia.

Những khó khăn đặt ra cho công tác bổ sung thường là:

- Khó khăn về tài chính: thiếu tiền mặt, thiếu ngoại tệ.
- Khó khăn về trang thiết bị: diện tích kho, các thiết bị lưu trữ và bảo quản một số loại tài liệu.
- Khó khăn về cơ chế, về chính sách thông tin, về khả năng có quan hệ với các nguồn thông tin trong nước và ngoài nước.

Các thủ tục bổ sung

Các thủ tục bổ sung được tiến hành theo một kế hoạch tỉ mỉ, gồm các bước sau:

- Xác định yêu cầu bổ sung và làm dự trù (đơn đặt hàng). Trước khi gửi đơn đặt hàng đi cần kiểm tra lại các chi tiết như: tên sách, tác giả, nhà xuất bản, năm xuất bản, số ký hiệu ISBN, ISSN. Cũng cần kiểm tra xem trong kho đã có tài liệu đó chưa để tránh sự trùng lặp.

- Lập bản kê khai các tài liệu cần mua thành nhiều bản. Tiến hành tất cả các thủ tục hành chính cho việc bổ sung.

- Khi nhận tài liệu, cần kiểm tra xem tài liệu nhận được có đúng với đơn đặt hàng hay không, tình trạng của tài liệu có còn tốt hay không.

- Ghi các tài liệu vào “sổ nhập”, theo thứ đến của tài liệu. Mỗi loại tài liệu phải có một sổ nhập tương ứng.

- Cám ơn các cơ quan, cá nhân đã gửi biếu tài liệu để xây dựng quan hệ lâu dài.

- Chuyển tài liệu đến bộ phận xử lý hình thức, xử lý nội dung để nhập vào hệ thống lưu trữ và khai thác thông tin của đơn vị thông tin.

Chương 4

XỬ LÝ TÀI LIỆU VÀ CÁC SẢN PHẨM THÔNG TIN

1. MÔ TẢ THƯ MỤC

1.1. Khái niệm chung về mô tả thư mục

Mô tả thư mục vừa là một công đoạn, vừa là một sản phẩm.

Với tư cách là một sản phẩm, người ta gọi đó là một chỉ dẫn thư mục hay một tra cứu thư mục. Nó bao gồm một tập hợp các chỉ dẫn nhằm cung cấp cho ta một mô tả duy nhất và chính xác của tài liệu và được xem như một vật mang tin.

Với tư cách là một công đoạn, người ta gọi đó là công tác biên mục (cataloging). Đó là bước đầu tiên của việc xử lý tài liệu, nhờ đó những chỉ dẫn được rút ra và trình bày theo một quy tắc chặt chẽ.

Mô tả thư mục bao gồm các công việc:

- Khảo sát tài liệu để xác định một số dữ liệu nêu lên những đặc trưng hình thức của tài liệu (tác giả, nhan đề, các yếu tố xuất bản, số trang...).

- Ghi các dữ liệu này trên một vật mang tin nhất định (phiếu, tờ nhập tin) theo các quy định và tiêu chuẩn được xác lập trên phạm vi quốc tế để khai thác sau này.

Những dữ liệu tìm kiếm này là những đặc điểm vốn có của mỗi tài liệu. Nó cho phép mô tả tài liệu đó, nhận biết tài liệu này một cách chính xác, không nhầm lẫn và như vậy có thể dễ dàng tìm lại nó khi có yêu cầu.

Nói chung việc biên mục được thực hiện bởi một trung tâm quốc gia như thư viện quốc gia (biên mục tập trung) hay được phân công cho một số đơn vị thông tin đảm nhận (hợp tác biên mục) và được thực hiện ngay sau khi bổ sung và nhập tài liệu vào đơn vị thông tin.

Mục đích của mô tả thư mục cung cấp cho tài liệu một mô tả duy nhất, không mơ hồ, nó giúp cho ta cùng một lúc có thể xác định được tài liệu, sắp xếp chúng, đưa chúng vào các bộ phiếu và tìm kiếm các tài liệu đó.

Ta cần phân biệt sự khác nhau giữa mô tả thư mục và chỉ dẫn thư mục. Mô tả thư mục là tập hợp tất cả các chỉ dẫn cần thiết để mô tả tài liệu. Còn chỉ dẫn thư mục là một số những chỉ dẫn xác định, không cần phải tất cả, trình bày trong các bộ phiếu thủ công hay tự động hoá, giúp cho việc tra cứu của người dùng tin.

1.2. Các vùng dữ liệu

Mô tả thư mục bao gồm các vùng dữ liệu. Đó là những tập con của các dữ liệu ứng với các loại chỉ dẫn riêng biệt, mà mỗi yếu tố của nó mô tả một khía cạnh của tài liệu. Các vùng dữ liệu được sắp xếp theo một trật tự logic.

Các vùng dữ liệu này hơi khác nhau đối với các loại hình tài liệu khác nhau, đặc biệt giữa các sách chuyên khảo và các ấn phẩm hàng loạt. Đối với các tài liệu không văn bản, việc mô tả chúng có những nét riêng, nhưng về cơ bản các dữ liệu cũng được chia thành các vùng và được sắp xếp theo một trật tự nhất định.

Một vài vùng dữ liệu không thể thiếu được trong tất cả các tình huống và luôn có mặt trong các chỉ dẫn thư mục còn một số vùng khác có thể lựa chọn và không bắt buộc. Các vùng có thể bao gồm chỉ một yếu tố dữ liệu hay nhiều yếu tố có quan hệ với nhau.

Với các tài liệu văn bản người ta xác định khoảng 15 vùng dữ liệu. Chẳng hạn, các chỉ dẫn thư mục của các sách chuyên khảo và các ấn phẩm định kỳ, theo quy tắc mô tả ISBD, bao gồm các vùng sau:

Các sách chuyên khảo

- Tên sách và tác giả
- Xuất bản
- Địa chỉ
- Đặc trưng số lượng
- Tàng thư
- Phụ chú
- ISBN, bìa và giá

Các ấn phẩm định kỳ

- Tên tạp chí và tác giả
- Xuất bản
- Số thứ tự
- Địa chỉ
- Đặc trưng số lượng
- Tàng thư
- Phụ chú
- ISSN, khoá tiêu đề và giá

Chỉ dẫn thư mục của một tài liệu có thể được bổ sung các yếu tố khác, sau khi mô tả nội dung, như các chỉ số phân loại, các bản đánh chỉ số nội dung, các bản tóm tắt nếu cần thiết.

Các mô tả thư mục còn được dùng trong hệ thống tư liệu khi không có tài liệu cấp một. Nó thay thế cho tài liệu trong một số công đoạn của dây truyền tư liệu. Nó cho một biểu diễn đúng đắn về tài liệu và cung cấp cho người dùng tin các chỉ dẫn để có thể lựa chọn và tìm tài liệu cấp một.

1.3. Phương pháp mô tả thư mục

Việc mô tả thư mục được tiến hành theo các bước sau:

- Làm quen với tài liệu.
- Xác định loại hình tài liệu và các quy tắc có thể áp dụng trong trường hợp đặc biệt.
- Xác định mức độ thư mục cần phải xử lý.
- Xác định các dữ liệu cần thiết, theo thứ tự các vùng mà các chuẩn và format đã quy định.
- Ghi lại các dữ liệu này theo các chuẩn và format đã quy định.
- Kiểm tra tính đúng đắn của mô tả và tính tương hợp theo các chuẩn đã quy định.
- Chuyển các mô tả cho bộ phận in và các bước tiếp theo.

Các dữ liệu thư mục được rút ra chủ yếu từ tài liệu gốc và cả nguồn phụ bên ngoài khi cần thiết. Trang tiêu đề nói chung chứa phần lớn các yếu tố của mô tả thư mục. Tên sách và tác giả cần được chọn theo trang tiêu đề chứ không phải theo trang bìa vì trang bìa thường ít đầy đủ hơn. Ngoài ra các

bản mục lục và minh hoạ cũng được dùng để bổ sung và chính xác hoá các dữ liệu do trang tiêu đề cung cấp, chẳng hạn như số trang và các minh hoạ. Khi thiếu thông tin hoàn toàn người ta có thể tìm những dữ liệu thư mục từ nguồn bên ngoài. Chẳng hạn tên thực của tác giả, giá của tài liệu, niên hạn của tác phẩm có thể tìm trong các thư mục, các mục lục, các danh mục, các thông báo của nhà xuất bản...

Không thể lập những dữ liệu mà chưa xác định được tính xác thực của nó.

1.4. Quy tắc mô tả thư mục theo tiêu chuẩn quốc tế ISBD

Tiêu chuẩn hoá kỹ thuật xử lý thông tin

Ngày nay nhu cầu thông tin trên thế giới ngày càng cao và bức thiết, cùng với sự phát triển các kỹ thuật xử lý thông tin đã thúc đẩy sự hợp tác và phối hợp giữa các tổ chức thông tin. Từ đó đòi hỏi phải thiết lập sự tương hợp giữa các hệ thống thông tin, tức là phải tiêu chuẩn hoá các kỹ thuật xử lý thông tin.

Tiêu chuẩn hoá giữ vai trò quan trọng trong việc nâng cao hiệu quả của hoạt động khai thác và chuyển giao thông tin, cũng như mở rộng sự hợp tác giữa các đơn vị thông tin trên phạm vi quốc gia và quốc tế.

Tiêu chuẩn hoá là một hoạt động tập thể mà ở đó người ta lập nên các tiêu chuẩn về kỹ thuật, về phương pháp, về chất lượng mà mọi người phải tuân theo như một chuẩn mực chung.

Tiêu chuẩn là những đặc trưng kỹ thuật chứa đựng một tập hợp các điều kiện phải thực hiện. Chúng có thể do các hội, các tổ chức quốc gia và quốc tế biên soạn. Việc biên soạn các tiêu chuẩn thường dựa vào kinh nghiệm sau đó sẽ được điều chỉnh để đạt mức tối đa. Tiêu chuẩn có thể được chấp nhận ở nhiều mức khác nhau: tiêu chuẩn địa phương, tiêu chuẩn quốc gia, tiêu chuẩn quốc tế.

Tiêu chuẩn giữ vai trò không thể thiếu trong hoạt động khoa học kỹ thuật, trong sản xuất kinh doanh cũng như trong đời sống xã hội. Nhờ chuẩn hoá mà người ta có thể thúc đẩy sản xuất, mở rộng thị trường và các sản phẩm có thể trao đổi và tương hợp lẫn nhau.

Chuẩn hoá trong lĩnh vực hoạt động thông tin tư liệu cũng có vai trò cơ bản như trong các lĩnh vực khác, đặc biệt quan trọng khi nhu cầu hợp tác về sản phẩm và dịch vụ giữa các đơn vị thông tin ngày càng gia tăng. Cần chuẩn hoá về các trang thiết bị được sử dụng, về sản phẩm thông tin cũng như về các công cụ và phương pháp xử lý nội dung, hình thức tài liệu. Nó cho phép đơn giản và hợp lý hoá các quy trình kỹ thuật và phương pháp, thống nhất sản phẩm, giảm chi phí giá thành và thúc đẩy sự trao đổi hợp tác giữa các đơn vị thông tin.

Trên thế giới người ta đã lập ra cơ quan tiêu chuẩn quốc tế ISO (International Standardization Organization). Đó là một tổ chức quốc tế được thành lập năm 1946 để biên soạn các tiêu chuẩn cho tất cả các lĩnh vực. ISO gồm nhiều ủy ban. Chẳng hạn, ủy ban ISO/TC46 là ủy ban kỹ thuật của ISO có nhiệm vụ biên soạn các tiêu chuẩn cho lĩnh vực thông tin tư liệu. Nó bao gồm 40 cơ quan thành viên ở khắp nơi trên thế giới.

Nhu cầu tiếp cận một cách thuận tiện, phổ dụng các thông tin thư mục và phát triển sự hợp tác giữa các đơn vị thông tin đã dẫn đến phải chuẩn hoá công tác biên mục.

Quy tắc mô tả thư mục quốc tế ISBD

Năm 1960, với sự cố gắng của tổ chức tiêu chuẩn quốc tế ISO/TC46 và sự hỗ trợ của Hiệp hội các thư viện quốc tế IFLA, quy tắc mô tả thư mục theo tiêu chuẩn quốc tế ISBD (International Standard Bibliographic Description) được biên soạn.

Đầu tiên là các quy tắc mô tả thư mục dành cho sách (Monographies) - ISBD (M), rồi đến các quy tắc dành cho các ấn phẩm định kỳ (Serials) - ISBD (S), sau đó mở rộng ra cho các loại hình tài liệu chuyên dạng khác. ISBD(G) là quy tắc mô tả thư mục dùng cho các tài liệu nói chung. Các quy tắc mô tả các tài liệu chuyên dạng sau đây đều được xây dựng dựa trên ISBD(G):

ISBD (NBM): mô tả tài liệu không ở dạng sách (Non - Book Materials).

ISBD (CM): mô tả tài liệu dạng đồ biểu (Cartographic Materials).

ISBD (PM): mô tả ấn phẩm âm nhạc (Printed Music).

ISBD (A): mô tả sách cổ (Antiquarian).

ISBD (CP): mô tả các bài trích (Component Parts).

ISBD (ER): mô tả các nguồn tài liệu điện tử (Electronic Resources).

ISBD là một tập hợp các quy tắc trình bày các dữ liệu thư mục theo một quy định chặt chẽ, cùng với các dấu hiệu dùng

để xác định chúng. Nó phân chia các dữ liệu thư mục thành từng vùng và vùng con với những thuyết minh đủ chính xác về các vùng đó. Ngoài ra ISBD còn đưa vào một hệ thống các dấu phân cách để báo hiệu chỗ bắt đầu hoặc kết thúc một vùng, vùng con.

Các vùng mô tả của ISBD

Các dữ liệu thư mục trong ISBD được phân thành 8 vùng mô tả:

1. Vùng nhan đề và thông tin về trách nhiệm.
2. Vùng thông tin về lần xuất bản và trách nhiệm liên quan đến lần xuất bản.
3. Vùng thông tin đặc thù (dành cho ấn phẩm định kỳ và tài liệu chuyên dạng).
4. Vùng địa chỉ xuất bản (Nơi xuất bản, nhà xuất bản, năm xuất bản).
5. Vùng mô tả vật lý hay đặc trưng số lượng (số trang, minh hoạ, khổ cỡ).
6. Vùng từng thư (sách bộ).
7. Vùng phụ chú.
8. Vùng chỉ số ISBN, ISSN và điều kiện có được tài liệu.

Trong các vùng trên có những vùng được dùng thường xuyên, nhưng cũng có nhiều vùng ít được sử dụng. Mỗi vùng lại chứa một số yếu tố dữ liệu, còn gọi là vùng con. Ví dụ vùng thông tin về xuất bản bao gồm các yếu tố: nơi xuất bản, nhà xuất bản, năm xuất bản.

Hệ thống ký hiệu

ISBD sử dụng các dấu ký hiệu để:

- Chỉ rõ sự bắt đầu của mỗi vùng.
- Phân cách các yếu tố trong một vùng.
- Xác định các yếu tố đặc thù bởi dấu ký hiệu đứng trước chúng.

Ví dụ trong vùng mô tả vật lý: minh hoạ bao giờ cũng đặt sau dấu hai chấm (:), khổ sách bao giờ cũng đặt sau dấu chấm phẩy (;) v.v...

Ví dụ: Dưới đây là một mô tả thư mục theo ISBD:

The vision and dream / Marguerite Hargrove ;
photographs by Harry and Claude Frauca. – Cammeray,
N.S.W. : Horwitz, 1980. – 304 p. – : ill. ; 18 cm. – (Bucks
books)

ISBN 0-7255-0934-1 : \$15.95.

Trong mô tả trên:

Nhan đề:	The vision and dream
Tác giả thứ nhất:	/ Marguerite Hargrove
Tác giả thứ hai:	; photographs by Harry and Claude Frauch
Các yếu tố xuất bản:	. – Cammeray, N.S.W. : Horwitz, 1980
Mô tả vật lý:	. – 304 p. – : ill ; 18 cm.
Tùng thư:	- (Bucks books)

Chỉ số ISBN, giá: ISBN 0-7255-0934-1 : \$15.95.

Chức năng của ISBD là đặc trưng hoá các yếu tố mô tả dùng để xác định tài liệu, gán một thứ tự cho các yếu tố ấy và quy định các dấu ký hiệu.

Mục đích của ISBD là tạo thuận lợi cho việc trao đổi quốc tế các ấn phẩm thông tin thư mục, cho phép các mô tả này không bị hàng rào ngôn ngữ hạn chế.

Quy tắc mô tả thư mục ISBD có những hạn chế nhất định, bởi vì nó chỉ cho những quy tắc chung. Vì vậy người ta phải lập ra các ủy ban ISBD để nghiên cứu và tổ chức việc áp dụng nó trong các thư viện hay các trung tâm thông tin.

1.5. Quy tắc biên mục Anh-Mỹ AACR2

Trong các nước nói tiếng Anh, người ta biên soạn ra quy tắc biên mục mục Anh - Mỹ AACR (Anglo - American Cataloguing Rules). AACR được xuất bản lần đầu vào năm 1967. AACR được biên soạn dựa trên cơ sở của ISBD nhưng chi tiết và sâu sắc hơn, đi sâu xét nhiều trường hợp tỉ mỉ, khắc phục được những khó khăn cụ thể. Năm 1978, với sự hợp tác của các thư viện quốc gia của Anh, Mỹ và Canada, quy tắc biên mục AACR2 được xuất bản và nhanh chóng được áp dụng. Năm 1998, AACR2 được in lại có sửa đổi và có tên viết tắt là AACR2R (*Anglo - American Cataloguing Rules .Second Edition 199. Revision*).

Ngoài các quy định về hình thức mô tả và tiêu đề, AACR2 đưa ra các quy tắc mô tả đối với các tư liệu không phải dạng sách báo. Quy tắc này đặc biệt coi trọng thông tin trong trang tiêu đề, ưu tiên lấy thông tin trong trang tiêu đề để đưa vào mô tả, chú trọng lập tiêu đề theo tên người (tác giả chính,

phụ) hơn là thể loại ấn phẩm. Quy tắc này cũng phân biệt hai hình thức trình bày mô tả: theo tác giả và theo nhan đề.

Về bố cục, khác với ISBD và các quy tắc biên mục trước đó, AACR2 trình bày các quy tắc mô tả trước, các quy tắc lựa chọn tiêu đề sau. Điều đó phù hợp với thực tiễn công tác biên mục. Cơ cấu nội dung của AACR2 gồm hai phần chính:

Phần một, là các quy tắc dành cho việc *mô tả tài liệu*, gồm 13 chương, trong đó có một chương trình bày các quy tắc chung cho mô tả dựa theo tinh thần của ISBD(G), các chương còn lại dành cho các loại tài liệu riêng như: sách, bản đồ, bản thảo, nhạc phẩm, tài liệu ghi âm, tài liệu ghi hình, tệp máy tính, xuất bản phẩm nhiều kỳ, v.v...

Phần hai, là các quy tắc về *thiết lập các điểm truy nhập (tiêu đề tên, nhan đề đồng nhất)* để thêm vào cho phần mô tả và *các tham chiếu cần thiết* để dẫn đến các điểm truy nhập đó. Phần này gồm 6 chương, giới thiệu các vấn đề về: Lựa chọn điểm truy cập, tiêu đề tác giả cá nhân, tên địa danh, tiêu đề tác giả tập thể, nhan đề thống nhất và tham chiếu. Các tiêu đề này được dùng để tìm tài liệu trong hệ thống mục lục và trong dịch vụ tìm tin.

AACR2 cung cấp các quy tắc mô tả chung cho tất cả các loại hình tài liệu bao gồm: các quy tắc về lựa chọn nguồn thông tin chính để mô tả tài liệu; các quy tắc về sử dụng hệ các ký hiệu mô tả; các quy tắc đối với các vùng dữ liệu.

Nguồn thông tin chính để mô tả tài liệu phụ thuộc vào loại tài liệu. Ví dụ: đối với sách là trang tiêu đề; đối với tệp máy tính là màn hình nhan đề; đối với băng ghi hình là nhãn; đối với bản đồ là chính tài liệu, v.v...

Hệ thống ký hiệu và các vùng dữ liệu của AACR2 cũng tương tự như ISBD, nhưng quy định chi tiết hơn.

AACR2 có 8 vùng dữ liệu:

Vùng 1: Nhan đề và xác minh về trách nhiệm (nhan đề chính, nhan đề song song, thông tin về nhan đề, xác minh đầu tiên về trách nhiệm, các xác minh tiếp theo về trách nhiệm)

Vùng 2: Vùng lần xuất bản (lần xuất bản, trách nhiệm liên quan đến lần xuất bản)

Vùng 3: Vùng đặc biệt (dành cho ấn phẩm nhiều kỳ, tệp máy tính, bản đồ đồ biểu, tác phẩm âm nhạc)

Vùng 4: Vùng xuất bản (Nơi xuất bản, nhà xuất bản, năm xuất bản)

Vùng 5: Vùng mô tả vật lý (Quy mô của tài liệu, các chi tiết vật chất khác, kích thước, các tài liệu đi kèm)

Vùng 6: Vùng từng thư (Tên từng thư, thông tin trách nhiệm liên quan đến từng thư,...)

Vùng 7: Vùng ghi chú (các ghi chú về nhan đề, về trách nhiệm, về xuất bản, về vật chất,... cung cấp thêm thông tin cho phần mô tả)

Vùng 8: Vùng số tiêu chuẩn (ISBN, ISSN, điều kiện có được tài liệu)

Dưới đây là lược đồ chung cho mô tả ấn phẩm trên phiếu mục lục thủ công theo AACR2:

Số hiệu Tiêu đề chính

Nhan đề chính = nhan đề song song : thông tin về
nhan đề / xác minh đầu tiên về tác giả ; xác minh tiếp theo
về tác giả ; ... – Lần xuất bản / xác minh liên quan đến lần
xuất bản. – Nơi xuất bản : Nhà xuất bản, năm xuất bản.

Số trang : minh hoạ ; khổ cỡ + tài liệu kèm. – (tùng
thư)

Ghi chú

Ghi chú ...

ISBN : giá tài liệu

1. Tiêu đề. 2. Tiêu đề... I. Tiểu dẫn. II. Tiểu dẫn,...

AACR2 đặt nền tảng cho sự hợp tác biên mục trên phạm
vi quốc gia và quốc tế. Do cung cấp một khổ mẫu mô tả chuẩn
mục thống nhất cho tất cả các loại hình tài liệu, nên quy tắc
này đã tạo khả năng tích hợp các mô tả của các loại hình tài
liệu khác nhau trong cùng một thư mục. Ngoài ra người dùng
tin cũng dễ dàng tìm liếm tài liệu hơn do được cung cấp
những tiêu đề truy cập nhiều hơn.

Hiện nay, AACR2 không chỉ được áp dụng ở các nước
Anh, Mỹ, Canada mà còn được áp dụng rộng rãi ở nhiều nước
khác trên thế giới. Với những ưu điểm vốn có về sự đầy đủ,
tính toàn diện và khả năng bao quát tất cả các loại hình tài
liệu, quy tắc này hiện được coi như là sự lựa chọn số một cho
công tác biên mục trong thời đại mà số lượng và các loại hình
tài liệu trở nên đa dạng và phong phú hơn bao giờ hết.

1.6. Khổ mẫu MARC và UNIMARC

Việc ứng dụng tin học trong xử lý thông tin càng thúc đẩy yêu cầu phải chuẩn hoá công tác biên mục. Tiêu chuẩn ISBD chưa hoàn toàn thích nghi với cách xử lý tin học đối với các mô tả thư mục. Các phương pháp xử lý tin học đòi hỏi dữ liệu phải được cấu trúc hoá. Để máy tính có thể nhận biết được các dữ liệu thư mục, các chỉ dẫn không những phải tuân theo những quy tắc chặt chẽ, mà còn phải được trình bày một cách chính xác theo một khổ mẫu thống nhất.

Khổ mẫu hay format là hình thức trình bày dữ liệu trên biểu ghi được cấu trúc hoá. Các dữ liệu trên biểu ghi được sắp xếp thành các trường, trường con, kết hợp với các mã số và các chỉ thị để điều hành sự sắp xếp này sao cho có thể nhận biết, trình bày và tìm kiếm dữ liệu bằng máy tính.

Khổ mẫu biên mục đọc được bằng máy - MARC

Năm 1966 lần đầu tiên Thư viện Quốc hội Mỹ xuất bản tập quy tắc của khổ mẫu MARC. MARC là từ viết tắt của cụm từ tiếng Anh: Machine Readable Cataloging có nghĩa là biên mục đọc được bằng máy. Khổ mẫu MARC là một mô tả có cấu trúc, dành riêng cho các dữ liệu thư mục được đưa vào máy tính điện tử. Nó là khổ mẫu cho phép máy tính lưu giữ và truy xuất thông tin.

Khổ mẫu MARC do Thư viện Quốc hội Mỹ xây dựng đầu tiên và sử dụng nên còn gọi là USMARC. Sau khi được chỉnh lý vào năm 1968, format MARC là cơ sở cho sự ra đời của một loạt các format quốc gia như: CAN.MARC của Canada, UK.MARC của Anh, AUS.MARC của Úc, INTERMARC của Pháp, IBERMARC của Tây Ban Nha. Mỗi hệ thống biên mục

sử dụng format MARC đều có hướng dẫn riêng về hệ thống phân cách và hình thức trình bày các biểu ghi của mình.

Năm 1997 USMARC của Thư viện Quốc hội Mỹ kết hợp với CANMARC của Thư viện Quốc gia Canada tạo thành MARC21 và trở thành format chuẩn được nhiều phần mềm quản trị thư viện sử dụng.

Sau đây ta sẽ tìm hiểu cấu trúc của khổ mẫu MARC21.

Cấu trúc của khổ mẫu MARC là cấu trúc biểu ghi, trong đó các dữ liệu thư mục được sắp xếp trong các trường, có độ dài xác định, được mã hoá và trình bày theo một quy định chặt chẽ. Khổ mẫu MARC sử dụng các chữ số, chữ cái, các ký hiệu ngắn gọn đặt ngay trong biểu ghi thư mục để đánh dấu và nhận biết các loại thông tin khác nhau trong mỗi biểu ghi.

Mỗi biểu ghi của khổ mẫu MARC bao gồm các trường (fields). Ngoài các trường dành cho các yếu tố mô tả thư mục theo AACR2 như: nhan đề, thông tin về trách nhiệm, thông tin về xuất bản, từng thư, đặc trưng số lượng, phụ chú, tóm tắt, v.v... còn có các trường dành cho đề mục chủ đề, ký hiệu phân loại của Thư viện Quốc hội Mỹ, ký hiệu phân loại thập phân Dewey,... Các trường này lại có thể chia nhỏ thành các trường con (subfields). Vì tên của trường thường khá dài nên trong biểu ghi MARC mỗi trường được biểu diễn bằng một nhãn gồm 3 chữ số.

Để tiện trình bày, người ta tập hợp các nhãn trường thành từng nhóm. Các trường có nhãn bắt đầu bằng số "0" thuộc nhóm trường "0XX", các trường có nhãn bắt đầu bằng số "1" thuộc nhóm trường "1XX", các trường có nhãn bắt đầu bằng số "2" thuộc nhóm trường "2XX", v.v...

Theo bản MARC21 đầy đủ cho dữ liệu thư mục, MARC21 có 10 khối trường chính:

- 0XX: Trường điều khiển, các chỉ số nhận dạng và phân loại
- 1XX: Tiêu đề mô tả chính (tên cá nhân, tên tập thể)
- 2XX: Nhan đề và thông tin liên quan đến nhan đề.
- 3XX: Mô tả vật lý
- 4XX: Thông tin về từng thư
- 5XX: Phụ chú
- 6XX: Các tiêu đề mô tả theo chủ đề
- 7XX: Các tiêu đề mô tả bổ sung khác chủ đề, từng thư
- 8XX: Tiêu đề mô tả bổ sung về từng thư
- 9XX: Thông tin nội bộ

MARC21 có khoảng trên 200 trường (không kể trường con). Qua thống kê người ta thấy rằng chỉ có khoảng 10% số nhãn được sử dụng thường xuyên trong các biểu ghi MARC, 90% còn lại ít khi được sử dụng.

Các nhãn trường thường được sử dụng là:

- 010 Số kiểm tra của Thư viện Quốc hội Mỹ
- 020 Chỉ số sách quốc tế (ISBN)
- 022 Chỉ số ấn phẩm định kỳ quốc tế (ISSN)
- 043 Mã vùng địa lý

050 Chỉ số phân loại của Thư viện Quốc hội Mỹ (LCC)

082 Chỉ số phân loại thập phân Dewey (DDC)

100 Tiêu đề chính, tên cá nhân (tác giả)

245 Nhan đề và thông tin về trách nhiệm

250 Thông tin về lần xuất bản

260 Thông tin về xuất bản, phát hành

300 Mô tả vật lý (đặc trưng số lượng)

440 Thông tin về từng thư

520 Chú giải hay tóm tắt

650 Điểm truy nhập chủ đề, đề mục chủ đề

700 Điểm truy nhập bổ sung, tên cá nhân

Danh sách tất cả các nhãn trường được in trong 2 tập *MARC 21 Format for Bibliographic Data* do Văn phòng Phát triển mạng lưới và tiêu chuẩn MAC của Thư viện Quốc hội Mỹ xuất bản.

Để hỗ trợ cho công tác biên mục Văn phòng này còn cho xuất bản các tài liệu bổ trợ sau:

- MARC21 format for authority data.
- MARC21 format for classification data.
- MARC21 format for holdings data.
- MARC21 format for community information.
- MARC21 code list for geographic areas.
- MARC code list for countries.

- MARC code list for languages.
- MARC code list for organizations.

Các tài liệu trên phối hợp với nhau để tạo ra một biểu ghi thư mục hoàn chỉnh, kiểm soát tính nhất quán của các tiêu đề mô tả, các thuật ngữ chủ đề, hỗ trợ phân loại chính xác, tạo tham chiếu qua lại giữa các tiêu đề và chủ đề, giữa các ký hiệu hay mục lục phân loại, theo dõi và thông báo vốn tài liệu hiện có gắn liền với công tác bổ sung và lưu thông.

Ngoài nhãn trường, trong biểu ghi MARC một số trường còn được xác định bằng các chỉ thị (indicators). Chỉ thị được mã hoá bằng hai chữ số (từ 0 đến 9) đi theo sau nhãn trường. Có trường chỉ dùng chữ số thứ nhất hoặc thứ hai, có trường dùng cả hai. Trong các trường không cần chỉ thị người ta bỏ trống.

Ví dụ:

245	14	\$a The emperor's new clothes / \$c adapted from Hans Christian Andersen and illustrated by Janet Stevens
-----	----	---

Trong ví dụ trên, 245 là nhãn trường nhan đề và thông tin về trách nhiệm, 14 là chỉ thị. Chỉ thị thứ nhất "1" có nghĩa là trong phiếu mục lục phải in từ "Title" vào trước nhan đề. Còn chỉ thị thứ hai "4" có nghĩa là bỏ qua 4 ký tự đầu của nhan đề (T, h, e và dấu trống). Khi đó nhan đề "The emperor's new clothes" sẽ được ghi ra là "emperor's new clothes".

Trong biểu ghi MARC, trường con được nhận biết bởi mã trường con (subfield codes), đó là một ký tự bằng chữ in thường, đặt sau dấu ngăn cách \$ (delimiter).

Ví dụ:

300	\$a 675p.	: \$b ill.	; \$c 24cm
-----	-----------	------------	------------

Trong đó 300 là nhãn trường mô tả vật lý của cuốn sách, bao gồm các trường con \$a (Số trang), \$b (Thông tin minh họa), \$c (Khổ, cỡ). Trường này không cần chỉ thị.

Ví dụ:

260	\$a Newyork :	\$bChelsea House,	\$c 1986
-----	---------------	-------------------	----------

Trong đó 260 là nhãn trường xuất bản, bao gồm các trường con: \$a (Nơi xuất bản), \$b (Nhà xuất bản), \$c (Năm xuất bản).

Các nhãn trường, chỉ thị và mã trường con là các dấu hiệu để nhận biết và điều khiển cách bố trí các trường và trường con, do các chương trình quản trị CSDL quy ước khi xây dựng. Khi một biểu ghi thư mục đã được đánh dấu một cách chính xác và được lưu trữ dưới dạng một tệp dữ liệu trên máy tính thì chương trình quản trị CSDL sẽ đánh dấu và tạo khuôn dạng cho các thông tin này để in ra thành một bản thư mục, một phiếu mục lục hay hiển thị trên màn hình. Các chương trình này còn cung cấp công cụ tìm kiếm các thông tin thoả mãn yêu cầu đặt ra, dựa trên các điểm truy nhập nằm trong các trường của biểu ghi MARC.

Ngoài các trường dữ liệu thư mục là phần chính của biểu ghi thư mục, mỗi biểu ghi MARC còn có các thành phần cố định sau, xuất hiện ở đầu mỗi biểu ghi:

a. *Đầu biểu (Leader)*. Đầu biểu gồm 24 ký tự đầu tiên của biểu ghi, ứng với 24 vị trí, mỗi vị trí này được gán cho một ý nghĩa xác định và thể hiện bằng một mã (chữ in thường, chữ

số hoặc khoảng trống) cho ta biết các thông tin về trạng thái và các thuộc tính của biểu ghi như: độ dài biểu ghi, loại hình tài liệu, cấp thư mục, mức độ mã hoá, quy tắc mô tả được sử dụng (ISBD, AACR),... Nhiều thông tin trong đầu biểu ghi là để dành cho máy tính sử dụng để nhận dạng biểu ghi.

Ví dụ: Giả sử một biểu ghi MARC có đầu biểu:

01401cam#2200265#a#4500

thì ý nghĩa cơ bản của nó là:

Vị trí	Ký tự	Ý nghĩa ký tự của vị trí
0- 4	01041	Độ dài biểu ghi: 1041 ký tự
5	c	Trạng thái biểu ghi: đã sửa chữa
6	a	Dạng tài liệu: văn bản in
7	m	Cấp thư mục: sách chuyên khảo
....		
18	a	Quy tắc mô tả được sử dụng: AACR2
19	#	Không yêu cầu các biểu ghi liên kết
20-23	4500	Sơ đồ các mục trong thư mục, luôn là 4500

b. *Danh mục (Directory)*. Tiếp theo đầu biểu là một loạt các tiểu dẫn (bằng các chữ số) có độ dài xác định gọi là danh mục. Danh mục này cho biết các nhãn trường có trong biểu ghi, độ dài của trường và vị trí bắt đầu của trường trong biểu ghi. Danh mục được tạo ra bởi máy tính từ biểu ghi thư mục, dựa trên các thông tin đã nhập. Nó không phải là phần MARC hiển thị cho người biên mục hay người sử dụng mục lục. Nó chỉ được sử dụng bởi người lập trình và máy tính để

thực hiện việc trao đổi các biểu ghi thư mục giữa các hệ thống.

Ví dụ: Đoạn tiểu dẫn sau đây trong danh mục

... 245003600354250001200390260003700402 ...

có nghĩa là: trong biểu ghi có trường với nhãn 245, có độ dài là 36 ký tự và bắt đầu từ vị trí thứ 354. Tiếp theo là trường có nhãn 250, có độ dài 12 ký tự và bắt đầu ở vị trí thứ 390 ($36 + 354 = 390$). Tiếp theo nữa là trường có nhãn 260, có độ dài 12 ký tự, bắt đầu từ vị trí 402 ($402 = 12 + 390$). Như vậy đoạn tiểu dẫn trên trong danh mục có cấu trúc và ý nghĩa như sau:

Nhãn	Độ dài	Bắt đầu từ vị trí
....		
245	0036	00354
250	0012	00390
260	0037	00402
....		

Tóm lại mỗi biểu ghi MARC phải bao gồm các thành phần cơ bản sau:

a. Đầu biểu gồm 24 ký tự.

b. Một danh mục các trường dữ liệu mà với mỗi trường dữ liệu phải bao gồm một nhãn với 3 chữ số, độ dài của trường dữ liệu và vị trí của ký tự đầu tiên.

c. Các dữ liệu tương ứng với các trường trong danh mục, chứa các dữ liệu thư mục cần xử lý với độ dài thay đổi.

Dưới đây là ví dụ về một biểu ghi MARC:

LEADER 00718nam//2200217/a/4500

001 ///93154367

003 DLC

005 19951019113246.0

008 940827s1993///at/ac///b///001/0/eng//

010 Sa///93154367

020 \$a1863731695 (pbk.) :\$c\$22.95

043 \$au-at---

050 00 \$aHQ1391.A8\$bS28 1993

082 00 \$a320.994082\$220

100 \$aSawer, Marian,\$d1946-

245 12 \$aA woman's place :\$bwoman and politics in
 Australia /\$cMarian Sawyer and Marian Simms

250 \$a2nd ed.

260 \$aSt. Leonards, N.S.W. :bAllen & Unwin,\$c1993

300 \$sxiii, 345 p. :\$bill., ports. ;\$c22 cm.

500 \$aPrevious ed.: Sydney : Allen & Unwin, 1984.

500 \$aIncludes index.

504 \$aBibliography: p. 309-328.

650 0 \$aWomen in politics\$zAustralia.

650 0 \$aWomen legislators\$zAustralia.

700 1 \$aSimms, Marian.

Trong biểu ghi trên, 001 là nhãn trường số kiểm soát, 003 là nhận dạng số kiểm soát, 005 là ngày và thời gian thực hiện giao tác gần nhất, 008 là các thông tin chung được mã hoá về biểu ghi.

Trong biểu ghi MARC, các trường cách nhau bởi dấu phân cách trường. Trong MARC21, dấu phân cách trường ký hiệu là ^. Dấu phân cách trường đặt ở cuối mỗi trường. Người ta dùng ký hiệu \ để đánh dấu kết thúc mỗi biểu ghi.

Cấu trúc biểu ghi của khổ mẫu MARC21 là cấu trúc điển hình của khổ mẫu biên mục hiện đại, với ba thành phần cơ bản: đầu biểu, danh mục các trường, các dữ liệu tương ứng với các trường trong danh mục. Hệ thống chỉ định nội dung với các nhãn trường, chỉ thị và mã trường con, các ký tự phân cách được quy định rất chi tiết và cụ thể, tạo thuận lợi cho máy tính nhận biết, sắp xếp và tìm kiếm dữ liệu. MARC21 phù hợp với tiêu chuẩn biên mục hiện đại AACR2 nên tạo khả năng biên mục linh hoạt cho nhiều loại hình tài liệu trong thư viện. MARC21 được xây dựng dựa trên chuẩn ISO.2709 nên việc trao đổi thông tin được thực hiện dễ dàng.

Cấu trúc biểu ghi của khổ mẫu MARC tạo ra nhiều khả năng cho máy tính lựa chọn và sắp xếp các dữ liệu thư mục:

- Cho phép người sử dụng dễ dàng truy cập tới các biểu ghi.
- In ra các thông báo sách mới, các ấn phẩm thư mục, các mục lục dưới dạng thức khác nhau, các nhãn trên gáy sách.
- Trao đổi dữ liệu thư mục với các thư viện khác trong nước và trên thế giới.

Khổ mẫu MARC có ý nghĩa quan trọng trong biên mục tự động. Thực chất của biên mục tự động là sử dụng một phần mềm tư liệu hoặc sử dụng phân hệ biên mục của phần mềm quản trị thư viện để tạo lập các biểu ghi cho một CSDL thư mục và tạo ra các mục lục thích hợp.

Trong biên mục tự động việc tạo lập biểu ghi thường là xử lý tiến máy và nhập dữ liệu, do con người thực hiện. Người ta nhập dữ liệu qua các khổ mẫu hiển thị trên màn hình theo kiểu xử lý văn bản. Còn việc tổ chức và sắp xếp biểu ghi và biên soạn các mục lục thì do máy tính thực hiện. Máy tính có thể in các phiếu mục lục và chế bản cho các ấn phẩm thư mục. Đó là sản phẩm đầu ra của CSDL thư mục.

Khổ mẫu trao đổi UNIMARC

Năm 1977, dựa theo tiêu chuẩn ISO-2709, Hiệp hội các Thư viện Quốc tế (IFLA) đã phổ biến format UNIMARC. Mục đích đầu tiên của UNIMARC là tạo thuận lợi cho sự trao đổi quốc tế các dữ liệu thư mục đọc được bằng máy giữa các trung tâm thư mục quốc gia. UNIMARC cũng có thể được sử dụng như một mô hình để phát triển các format trao đổi mới.

UNIMARC xử lý tất cả các loại hình tài liệu như: sách, ấn phẩm định kỳ, tài liệu đồ biểu, âm nhạc, phim ảnh, tài liệu nghe nhìn...

UNIMARC là một áp dụng riêng của tiêu chuẩn ISO 2709-1981. Đó là chuẩn quốc tế dùng để xác định cấu trúc của các biểu ghi chứa các dữ liệu thư mục. Cấu trúc của biểu ghi UNIMARC về cơ bản cũng giống như cấu trúc của biểu ghi MARC, tức là gồm ba phần:

a. Đầu biểu dài 24 ký tự.

b. Danh mục các trường dữ liệu.

c. Các dữ liệu tương ứng với các trường trong danh mục.

Các vùng dữ liệu của UNIMARC được phân thành 10 khối như sau:

0xx- Khối thông tin nhận dạng, dùng để nhận dạng tài liệu và nhận dạng biểu ghi, bao gồm các trường: mã số biểu ghi, mã số thư mục quốc gia, mã số lưu chiếu quốc gia, mã số xuất bản, chỉ số ISBN, ISSN,...

1xx- Khối thông tin mã hoá, dùng 3 ký tự để chỉ mã nước xuất bản hay mã ngôn ngữ của tài liệu, ví dụ: Vie (Việt Nam), Eng (Anh); mã hoá các loại hình tài liệu như: sách, tạp chí, tài liệu đồ biểu, tài liệu nghe nhìn, hình ảnh, đồ thị, âm thanh...

2xx- Khối thông tin mô tả thư mục: bao gồm các trường, trường con. Về cơ bản giống như ISBD, trừ phụ chú. Ví dụ: nhan đề, thông tin về trách nhiệm, thông tin về lần xuất bản, loại hình tài liệu, đặc trưng số liệu (collation), từng thư (collection - sách bộ),...

3xx- Khối phụ chú: phụ chú về mô tả thư mục, về nhan đề và tác giả, về nguồn gốc tài liệu và về bản dịch.

4xx- Khối liên kết các chỉ dẫn: bộ sưu tập, xuất bản gốc, đồng xuất bản, nhan đề bản dịch.

5xx- Khối nhan đề kết hợp: nhan đề chính, nhan đề sắp xếp, nhan đề song song, nhan đề ngoài bìa, nhan đề đầu trang,...

6xx- Khối chủ đề phân tích: tác giả, nhan đề, nơi xuất bản, ngày tháng, mã vùng địa lý, ký hiệu phân loại, từ khoá,...

7xx- Khối thông tin về trách nhiệm nội dung: tác giả chính, đồng tác giả, tác giả tập thể,...

8xx- Khối thông tin về các dữ liệu quốc tế: mã nước, cơ quan biên mục.

9xx- Khối thông tin nội bộ: ký hiệu kho, số lượng tài liệu trong kho.

Năm 1984 khổ mẫu trao đổi chung CCF (Common Communication Format) đã được biên soạn nhằm đáp ứng việc trao đổi dữ liệu thư mục giữa các cơ quan sử dụng format MARC. Sản phẩm từ các nguồn khác nhau dần trở nên tương hợp giúp các thư viện có thể trao đổi biểu ghi cho nhau. Việc cho mượn giữa các thư viện cũng thuận lợi hơn nhờ hệ thống mục lục liên hợp đã bảo đảm tính nhất quán trong biên mục. Ngoài ra quá trình nối mạng cũng dễ dàng hơn. Người ta có thể truy nhập các CSDL khác nhau qua việc sử dụng giao thức chuẩn.

1.7. Chỉ số ISBN và ISSN

Người ta còn xây dựng những mã số để xác định một cách chính xác một tài liệu. Đó là chỉ số sách quốc tế ISBN và chỉ số tạp chí quốc tế ISSN.

ISBN (International Standard Book Number) là một chỉ số gán cho mỗi quyển sách trong khuôn khổ của một hệ thống thông tin quốc tế. Nó bao gồm một tập hợp 10 chữ số được

chia thành bốn nhóm, cách nhau bằng dấu gạch ngang, mà ba nhóm đầu có độ dài thay đổi, đó là các chỉ số nhằm:

- Xác định khu vực (chỉ số khu vực).
- Xác định cơ quan xuất bản (chỉ số xuất bản).
- Xác định tên sách (chỉ số tên sách).

Số cuối cùng là số kiểm tra, nó cho phép tự động kiểm tra tính hiệu lực của ISBN.

Ví dụ: Với chỉ số ISBN 2 - 7081 - 0324 - 5 thì trong đó:

2: chỉ vùng nói tiếng Pháp

7081: chỉ nhà xuất bản "Les éditions d'organisation"

0324: chỉ tác phẩm của G.Ven Slipe "Conception et gestion des systèmes documentaires"

5: số kiểm tra.

Hệ thống chỉ số sách quốc tế ISBN được đưa ra ở Rayaune-Uni vào năm 1967. Lợi ích của ISBN trong việc mua bán và trao đổi sách thật là rõ ràng. Bằng cách gán cho mỗi quyển sách một chỉ số, nó cho phép mỗi quyển sách được xác định một cách đơn giản và rõ ràng trên phạm vi toàn thế giới. Chỉ số còn cho ta biết ngôn ngữ sử dụng và nhà xuất bản. Ngoài ra sử dụng chỉ số này trong MTĐT làm cho đơn giản hoá và tăng tốc độ xử lý trong nhiều công đoạn của dây chuyền tư liệu.

ISSN (International Standard Serials Number) là chỉ số xác định ấn phẩm định kỳ.

ISSN bao gồm 8 chữ số, trong đó chữ số cuối cùng là số kiểm tra và được chia thành hai nhóm, phân cách bởi dấu gạch ngang.

Ví dụ: ISSN 0002 - 8231 xác định tạp chí "Journal of the American society for information science".

Chỉ số ISSN xác định tên tạp chí một cách duy nhất. Vai trò của nó đối với ấn phẩm định kỳ cũng giống như vai trò của chỉ số ISBN đối với sách.

2. MÔ TẢ NỘI DUNG TÀI LIỆU

2.1. Khái niệm chung về mô tả nội dung tài liệu

Người ta gọi mô tả nội dung tài liệu là một tập hợp các công đoạn, ở đó người ta mô tả nội dung tài liệu cùng với những sản phẩm của chúng. Đó là phân loại, đánh chỉ số, tóm tắt và phân tích tài liệu.

Cùng một tài liệu có thể là đối tượng của nhiều bước mô tả nội dung với những mức độ khác nhau. Chẳng hạn, xác lập cho nó một chỉ số phân loại, một bản chỉ mục với gần một chục thuật ngữ hoặc một bản tóm tắt dài khoảng 100 từ. Những công đoạn này nối tiếp nhau và phụ thuộc lẫn nhau, được thực hiện theo một quy trình đòi hỏi phải có tri thức và kinh nghiệm.

Mục đích của mô tả nội dung là:

- Nắm bắt được nội dung tài liệu để thông báo cho người dùng tin.
- Tiến hành, khi cần thiết việc lựa chọn để duy trì hay loại bỏ tài liệu, xác định cách thức và mức độ xử lý tài liệu.

- Giúp cho việc sắp xếp, lưu trữ thông tin và tìm kiếm tài liệu.

Việc mô tả nội dung tài liệu có thể được thực hiện ở nhiều mức độ khác nhau. Mô tả nội dung càng sâu sắc thì giá trị sử dụng càng cao. Đối với một tài liệu thường có ba mức mô tả chính sau đây, theo mức độ sâu sắc tăng dần, đó là:

- Phân loại: là gán cho tài liệu một chỉ số của khung phân loại.
- Đánh chỉ số: là gán cho tài liệu một số từ chuẩn để nêu lên nội dung, chủ đề của tài liệu.
- Tóm tắt: là cô đọng tài liệu bằng một bài viết ngắn.

Để mô tả nội dung tài liệu người ta dùng ngôn ngữ tư liệu. Đó là ngôn ngữ nhân tạo, trong đó mỗi thuật ngữ có một ý nghĩa duy nhất, tạo điều kiện thuận lợi cho việc lưu trữ và tìm kiếm thông tin.

2.2. Ngôn ngữ tự nhiên và ngôn ngữ tư liệu

Ngôn ngữ là một hệ thống tín hiệu có bản chất bất kỳ, thực hiện chức năng nhận thức và giao tiếp trong quá trình hoạt động của con người. Ngôn ngữ có thể là ngôn ngữ tự nhiên hoặc ngôn ngữ nhân tạo.

Ngôn ngữ tự nhiên

Ngôn ngữ tự nhiên hay ngôn ngữ nói có những nét đặc trưng riêng khiến người ta không thể dùng nó trong việc xử lý thông tin. Những đặc trưng đó là:

- Ngôn ngữ tự nhiên có rất nhiều, có đến hàng trăm thứ tiếng khác nhau, mỗi thứ tiếng lại có hàng chục nghìn từ. Nhiều từ thừa và vô ích khi áp dụng vào công tác tư liệu.

- Khi giao tiếp bằng ngôn ngữ tự nhiên, ý nghĩa các từ thường thay đổi theo ngữ cảnh. Có nhiều từ đồng âm khác nghĩa lại có nhiều từ đồng nghĩa khác âm.

- Ngôn ngữ tự nhiên có nhiều loại từ mà chúng có giá trị thông tin khác nhau. Người ta thấy có danh từ, tính từ, động từ, trạng từ, liên từ, giới từ. Có những từ chỉ làm rõ nghĩa cho từ khác. Có từ dùng rất nhiều nhưng lại có giá trị thông tin rất ít. Nhiều từ quá rộng, quá chung.

- Trong ngôn ngữ tự nhiên người ta phải dùng các quy tắc ngữ pháp để kết hợp các từ với nhau nhằm diễn tả một ý tưởng, một sự việc. Các quy tắc cú pháp làm thay đổi các từ trong cách viết và cả trong ý nghĩa của chúng.

Tóm lại ngôn ngữ tự nhiên có những đặc điểm mà ta không thể sử dụng nó trong việc xử lý thông tin. Tuy nhiên ta phải quan tâm đến các đặc điểm đó để khắc phục tính mơ hồ của chúng trong việc xây dựng ngôn ngữ tư liệu.

Ngôn ngữ tư liệu

Để khắc phục những nhược điểm của ngôn ngữ tự nhiên trong việc diễn đạt thông tin, người ta phải xây dựng ngôn ngữ tư liệu sao cho thoả mãn ba yêu cầu sau:

- Quan hệ ngữ nghĩa là một - một: một thuật ngữ diễn tả một sự vật.

- Quan hệ cú pháp là nhất quán: chỉ có một cách biểu đạt.

- Tránh tính chủ quan của người sử dụng: không phụ thuộc vào ngữ cảnh.

Ngôn ngữ tư liệu là ngôn ngữ nhân tạo chuyên dụng, dùng để diễn đạt nội dung, ý nghĩa cơ bản của một tài liệu, phục vụ cho việc lưu trữ và tìm kiếm thông tin.

Các ngôn ngữ tư liệu được xây dựng trên hai yếu tố cơ bản sau:

- *Vốn từ vựng của ngôn ngữ.* Đó là các từ chuẩn dùng để mô tả thông tin. Từ chuẩn là những thuật ngữ được rút ra từ ngôn ngữ tự nhiên, thu gọn dưới một dạng ngữ pháp duy nhất, sau khi đã loại bỏ đi các từ đồng nghĩa.

- *Các yếu tố cú pháp.* Đó có thể là cách trình bày hoặc cách sử dụng các thuật ngữ, có thể là các ký hiệu để nối chúng với nhau, cho phép diễn tả các thông tin.

Có nhiều loại ngôn ngữ tư liệu. Người ta phân biệt chúng theo nhiều nét đặc trưng như: nguyên tắc xây dựng, diện phủ của ngôn ngữ, các loại từ sử dụng, quan hệ giữa các từ... Tuy nhiên, căn cứ vào cấu trúc, người ta phân biệt hai loại ngôn ngữ tư liệu chính:

Loại thứ nhất gồm các ngôn ngữ có cấu trúc cấp bậc một cách hệ thống, gọi là ngôn ngữ kết hợp trước. Các hệ thống phân loại theo chủ đề, các khung đề mục... là các ngôn ngữ thuộc loại này.

Loại thứ hai gồm các ngôn ngữ có cấu trúc tổ hợp, còn gọi là ngôn ngữ kết hợp sau. Các từ điển từ chuẩn (Thesaurus), danh mục các từ chuẩn, danh mục các từ khoá... là các ngôn ngữ thuộc loại này.

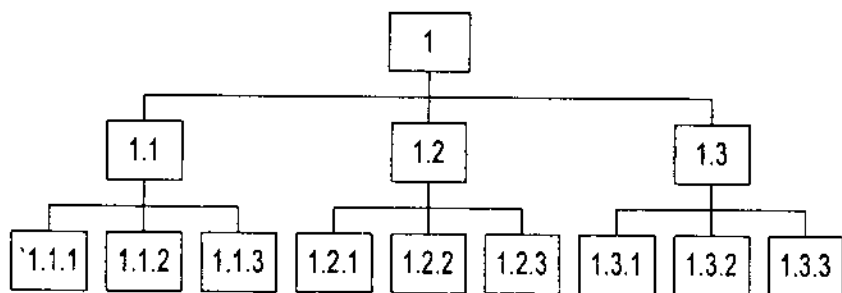
Đó là hai loại ngôn ngữ đối lập nhau về nguyên tắc xây dựng cũng như cách sử dụng.

2.3. Các hệ thống phân loại

Các hệ thống phân loại, còn gọi là khung phân loại, được xây dựng dựa trên quan hệ cấp bậc giữa các thuật ngữ. Quan hệ này dựa trên giả thiết cho rằng các đề mục hay chủ đề có thể chia thành các chủ đề nhỏ, chính xác hơn. Quá trình này được lặp lại nhiều lần, cho đến khi hình thành một cấu trúc cấp bậc của các thuật ngữ, bao trùm lên một lĩnh vực của tri thức mà hệ thống thông tin xem xét.

Do có cấu trúc cấp bậc nên trong khung phân loại người ta gán cho mỗi thuật ngữ một chỉ số. Chỉ số này có thể bằng chữ số, bằng chữ cái hay kết hợp giữa chữ số và chữ cái. Các chỉ số xác định vị trí của thuật ngữ trong khung phân loại và được dùng để thay thế cho nó trong nhiều công đoạn khác nhau vì nó ngắn gọn và dễ sử dụng.

Hình vẽ dưới đây cho ta hình ảnh trực quan về cấu trúc cấp bậc của các khung phân loại:



Hình 11. Cấu trúc cây của khung phân loại

Ngoài ra mỗi khung phân loại còn sử dụng một hệ thống các ký hiệu để làm rõ thêm ý nghĩa của các chỉ số, hoặc dùng để ghép các chỉ số với nhau nhằm mở rộng khả năng diễn tả thông tin.

Ta thấy các khung phân loại là ngôn ngữ chính xác, nhưng cũng nhắc, khả năng sử dụng hạn chế vì phụ thuộc vào vốn từ vựng của nó.

Các khung phân loại được dùng để tổ chức các bộ phiếu mục lục và để sắp xếp các tài liệu trên giá trong kho. Nó rất thích hợp với các hoạt động thư viện truyền thống.

Hệ thống phân loại thập phân Dewey (DDC)

Hệ thống phân loại thập phân của Dewey là khung phân loại đầu tiên của thư viện, do nhà thư viện học người Mỹ là Melvil Dewey (1851-1931) xây dựng, được xuất bản lần đầu vào năm 1876. Lúc đầu chỉ có 1000 đề mục. Ở lần xuất bản thứ 18 vào năm 1971 đã bao gồm 18.000 đề mục. Phân loại của Dewey là phân loại thập phân, trong đó toàn bộ tri thức của nhân loại được chia ra thành 10 lớp chính, đó là:

000 Kiến thức chung	500 Khoa học thuần túy
100 Triết học	600 Kỹ thuật
200 Tôn giáo	700 Nghệ thuật
300 Khoa học xã hội	800 Văn học
400 Ngôn ngữ	900 Lịch sử và địa lý

Mỗi lớp trên lại chia thành 10 lớp, thể hiện những lớp con của chủ đề. Chẳng hạn, lớp 500 là khoa học thuần túy lại chia thành 10 lớp sau:

500	Khoa học thuần túy
510	Toán học
520	Thiên văn học và khoa học liên quan
530	Vật lý
540	Hoá học
550	Địa chất học
560	Cổ sinh vật học
570	Nhân loại học và sinh học
580	Thực vật học
590	Động vật học

Lớp 510 là toán học, lại chia một lần nữa thành các lớp toán học khác nhau, cũng theo nguyên tắc thập phân nếu cần thiết:

500	Khoa học thuần túy
510	Toán học
511	Số học
512	Đại số
513	Hình học
513.1	Hình học phẳng
513.2	Hình học các đường cong
513.3	Hình học các khối

Cứ như thế, các lớp lại chia thành các lớp con, trong cố gắng nhằm liên kết các chủ đề có liên quan với nhau. Các chỉ số chi tiết giúp cho định vị các lớp đó một cách dễ dàng trong khung phân loại. Phân loại của Dewey do Thư viện Quốc hội Mỹ công bố và được sử dụng rộng rãi ở nhiều nước trên thế giới.

Hệ thống phân loại thập phân của Dewey là một trong nhiều đóng góp của Dewey cho lý luận và phương pháp luận của thư viện học. Với hệ thống phân loại Dewey, lần đầu tiên sự sắp xếp một cách hệ thống các sách trên giá được phản ánh bằng một lược đồ được xây dựng trên cơ sở lý thuyết của phân loại học. Việc sử dụng hệ thống số thập phân làm ký hiệu phân loại đã thể hiện một cách rõ ràng và đơn giản tính phân cấp một cách hệ thống của các chủ đề.

Khung phân loại DDC không ngừng phát triển. DDC xuất bản lần thứ 20 vào năm 1989, gồm 4 tập. Trong lần xuất bản thứ 21 vào năm 1996, ngoài dạng ấn phẩm DDC21 còn được phát hành dưới dạng CD-ROM chạy trên môi trường Windows. DDC đã được dịch ra trên 30 ngôn ngữ, trong đó có Ả Rập, Đức, Pháp, Nga, Trung Quốc, Ý, Hiện nay có trên 200.000 thư viện của 135 quốc gia sử dụng DDC.

Hệ thống phân loại thập phân bách khoa (UDC)

Hệ thống phân loại thập phân bách khoa do các luật sư người Bỉ là Pierre Otlet và Henry Lafontaine biên soạn dựa trên cơ sở phân loại của Dewey, được xuất bản lần đầu vào năm 1905. Nó được đưa ra trong khuôn khổ hoạt động của Liên đoàn Tư liệu Quốc tế (FID) và được sử dụng nhiều nhất ở châu Âu, rất thích hợp với các thư viện khoa học kỹ thuật.

Phân loại thập phân bách khoa giữ lại hầu hết các lớp của phân loại Dewey, riêng lớp 4 - Ngôn ngữ đưa vào lớp 8 - Văn học, còn lớp 4 hiện nay vẫn để trống. Phân loại thập phân bách khoa xây dựng một hệ thống chỉ số khá chi tiết, gồm ba loại chỉ số:

- *Chỉ số chính* gán cho tất cả các lớp của bảng chính, đó là các lớp:

- 0 Phần chung
- 1 Triết học
- 2 Tôn giáo
- 3 Các khoa học xã hội
- 4 (để trống)
- 5 Toán học, các khoa học tự nhiên
- 6 Các khoa học ứng dụng, y học, kỹ thuật
- 7 Nghệ thuật
- 8 Văn học
- 9 Lịch sử, địa lý

Các lớp này lại chia thành các lớp nhỏ theo nguyên tắc thập phân và nguyên tắc đi từ chung đến riêng, từ lớn đến nhỏ.

- *Chỉ số xác định chung* cho ta những đặc trưng về loại hình tài liệu, về địa điểm, ngôn ngữ, thời gian, quan điểm và nhân vật. Các chỉ số này gồm các số độc lập và có thể kết hợp với bất cứ phần nào của bảng chính.

Ví dụ:

622	Ngành mỏ
(031)	Từ điển, bách khoa toàn thư
622(031)	Từ điển, bách khoa toàn thư về mỏ
(597)	Việt Nam
622(597)	Ngành mỏ ở Việt Nam
= 20	Tài liệu viết bằng tiếng Anh
622=20	Tài liệu về ngành mỏ viết bằng tiếng Anh

- *Các chỉ số xác định riêng*, còn gọi là chỉ số phân tích để phản ánh rõ hơn bản chất đối tượng được xem xét. Các chỉ số này chỉ dùng trong từng phần cụ thể. Ví dụ: chỉ số 62 phản ánh thiết bị máy móc thì 62-59 là chỉ các bộ phận giảm tốc.

Ngoài ra UDC còn dùng các dấu ký hiệu. Các ký hiệu này dùng để phân biệt các loại chỉ số khác nhau và có thể dùng để ghép nối các chỉ số với nhau, tạo thành những chỉ số phức hợp, phản ánh những khái niệm mới tổng hợp mà từng chỉ số không phản ánh được.

Ví dụ:

Dấu : dùng để chỉ phép tuyển

Dấu + dùng để chỉ phép hội

Dấu / dùng để chỉ sự mở rộng

Áp dụng các ký hiệu trên, chẳng hạn:

(438) Ba Lan

(439) Hungarie

622(439+438)	Ngành mỏ ở Ba Lan hoặc Hungarie
669.1	Sắt, thép
546.22	Lưu huỳnh
669.1:546.22	Lưu huỳnh trong sắt, thép
669.1(439)"1975/1980"	Sắt, thép ở Hungarie từ năm 1975 đến 1980.

Mặc dù UDC cho phép miêu tả các khái niệm phức tạp với độ chính xác khá cao, nhưng nó vẫn là ngôn ngữ tư liệu có cấu trúc cấp bậc như tất cả các khung phân loại khác. Nó không cho phép kết hợp một cách linh hoạt và thuận tiện các khái niệm với nhau.

Khung phân loại là loại ngôn ngữ tư liệu cổ nhất. Ngoài phân loại Dewey và phân loại thập phân bách khoa, ta có thể kể một số khung phân loại khá nổi tiếng khác như:

- Phân loại của Thư viện Quốc hội Mỹ.
- Phân loại hai chấm của Ranganathan.

Phân loại của Thư viện Quốc hội Mỹ (LC)

Phân loại của Thư viện Quốc hội Mỹ (Library of Congress Classification - LC) do Herbert Putnam, một chuyên viên của thư viện, đề xuất từ đầu thế kỷ XX, được xây dựng trên quan điểm thực dụng, dựa vào cách sắp xếp các sách trên giá theo chủ đề. Trong LC tập hợp các tri thức được chia thành 21 nhóm, mỗi nhóm được ký hiệu bằng một chữ cái.

Các lớp chính của LC là:

A	Tác phẩm chung	M	Âm nhạc
B	Triết học-Tôn giáo	N	Nghệ thuật
C	Khoa học bổ trợ của lịch sử	P	Ngôn ngữ và văn học
D	Lịch sử thế giới	Q	Khoa học
E,F	Lịch sử Hoa Kỳ	R	Y học
G	Địa lý, nhân loại học	S	Nông nghiệp
H	Khoa học xã hội	T	Kỹ thuật
I	Khoa học chính trị	U	Khoa học quân sự
K	Luật	V	Khoa học hàng hải
L	Giáo dục	Z	Thư mục học, thư viện học

Các ký hiệu của LC là một tổ hợp bao gồm cả chữ cái và chữ số.

Ví dụ:

Q	Khoa học
QC	Vật lý
QC1	Ấn phẩm định kỳ về vật lý
QC12	Ấn phẩm chuyên ngành về vật lý
QC5	Từ điển, bách khoa toàn thư về vật lý
QC555	Tóm tắt và chỉ mục về vật lý
QC68	Lịch sử vật lý
QA	Toán học
QA101	Số học

QA152 Đại số

QA445 Hình học

Phân loại hai chấm của Ranganathan (CC)

Năm 1933 nhà thư viện học người Ấn Độ là Shilvali Ramanrita Ranganathan đưa ra hệ thống phân loại hai chấm (Colon Classification-CC). Đó là một hệ thống phân loại theo diện, ở đó các khái niệm và các đối tượng được sắp xếp theo 5 diện cơ bản: nhân vật (personality), vật chất (matter), năng lượng (energy), không gian (space), thời gian (time).

CC sử dụng một hệ thống ký hiệu khá phức tạp, nhưng đó là một ngôn ngữ được biên soạn công phu, nó cho phép mô tả đầy đủ các thông tin. CC ít được sử dụng so với các hệ thống phân loại phổ dụng khác. Nhưng nguyên tắc phân loại theo diện mà nó sử dụng có thể coi là một bước tiến trong lý thuyết phân loại và được áp dụng rộng rãi trong nhiều ngôn ngữ tư liệu, như từ điển từ chuẩn.

Cấu trúc toán học của các hệ thống phân loại

Khi phân loại một tập hợp các đối tượng nào đó, người ta thường căn cứ vào các dấu hiệu của các phần tử tùy theo yêu cầu đặt ra. Phân loại là sắp xếp các sự vật thành từng lớp ổn định, có hệ thống từ đầu đến cuối, dựa trên những dấu hiệu bản chất của đối tượng. Về phương diện toán học, phân loại một tập hợp đối tượng là xác định các lớp tương đương theo một đặc tính hay một quan hệ nào đó.

Giả sử X là tập hợp các đối tượng nào đó. Một quan hệ tương đương R xác định trên X các lớp tương đương $X_1, X_2, \dots, X_n$ sao cho:

$$1) X_i \cap X_j = \emptyset \quad \text{với } i \neq j$$

$$2) X_1 \cup X_2 \cup \dots \cup X_n = X$$

Tính chất trên đây của các lớp tương đương bảo đảm thực hiện hai nguyên tắc của phân loại là:

1- Mỗi đối tượng đều thuộc một lớp, tức là được phân loại.

2- Không có đối tượng nào thuộc hai lớp khác nhau, tức là thuộc hai loại khác nhau.

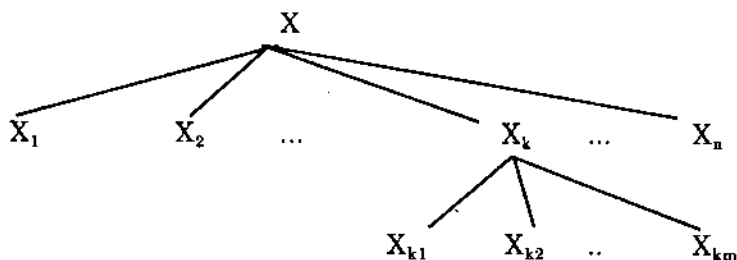
Đối với một lớp X_k nào đó, một quan hệ tương đương R_1 lại xác định ra các lớp tương đương $X_{k1}, X_{k2}, \dots, X_{km}$ sao cho:

$$1) X_{ki} \cap X_{kj} = \emptyset \quad \text{với } i \neq j$$

$$2) X_{k1} \cup X_{k2} \cup \dots \cup X_{km} = X_k$$

Quá trình trên cứ tiếp tục như thế, tập hợp các đối tượng được phân loại thành các lớp nhỏ ngày càng chi tiết hơn. Tính riêng của các lớp ngày càng cao.

Đồ thị sau đây biểu diễn cấu trúc cây của phân loại:



Hình 12. Cấu trúc cây của phân loại

Các hệ thống phân loại như: hệ thống phân loại của Dewey (DC), hệ thống phân loại thập phân bách khoa (UDC), hệ thống phân loại của Thư viện Quốc hội Mỹ (LC)... đều có cấu trúc toán học như trên.

Chẳng hạn cấu trúc toán học của hệ thống phân loại thập phân bách khoa (UDC) thể hiện như sau:

Trên tập T các tri thức của nhân loại, ta định nghĩa quan hệ R như sau:

$$R = \text{"có chủ đề } T_i\text{" với } i = 0, 1, 2, 3, \dots, 9.$$

Quan hệ đó chia T thành 10 lớp chính, biểu diễn bằng thuật ngữ của chủ đề:

$$T_0, T_1, T_2, \dots, T_9$$

Áp dụng nguyên tắc trên một lần nữa vào các lớp trên, ta được các lớp nhỏ:

$$T_{00}, T_{01}, T_{02}, \dots, T_{09}$$

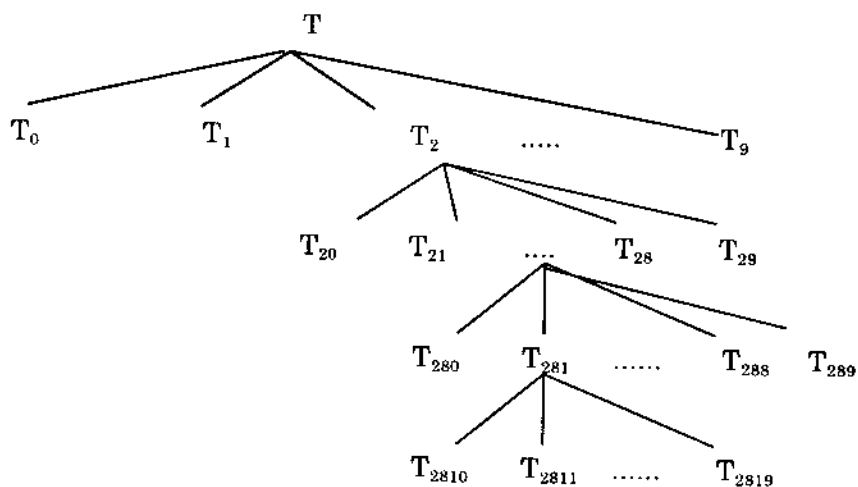
$$T_{10}, T_{11}, T_{12}, \dots, T_{19}$$

$$\dots$$

$$T_{90}, T_{91}, T_{92}, \dots, T_{99}$$

Quá trình ấy cứ tiếp tục, ta phân chia tri thức của nhân loại thành các lớp càng nhỏ, càng cụ thể, tính riêng của chúng càng rõ nét.

Nguyên tắc thập phân của UDC biểu hiện ở chỗ mỗi một lớp lại chia thành 10 lớp nhỏ hơn. Cấu trúc hình cây của UDC được biểu diễn bằng đồ thị sau:



Hình 13. Cấu trúc cây của UDC

Như trên hình vẽ, mỗi thuật ngữ chiếm một vị trí xác định và được biểu diễn bằng một chỉ số. Chỉ số này cho biết nó thuộc lớp nào, ở mức độ nào. Nguyên tắc từ chung đến riêng, từ lớn đến nhỏ của UDC thể hiện trên chỉ số là các chỉ số sau bao hàm trong chỉ số trước, chẳng hạn, với lớp T_{2819} ta có:

$$T_{2819} \subset T_{281} \subset T_{28} \subset T_2 \subset T$$

Như vậy tập hợp tất cả các lớp của UDC là một tập sắp thứ tự bộ phận.

Chính nhờ có cấu trúc như trên mà các hệ thống phân loại được dùng để phân loại tài liệu. Mục đích của phân loại là sắp xếp các tài liệu thành một số ít các lớp, giúp cho việc tổ chức các bộ phiếu và sắp xếp tài liệu trên giá nhằm phục vụ

cho việc tìm kiếm tài liệu theo nội dung chủ đề một cách nhanh chóng và thuận tiện.

2.4. Từ điển từ chuẩn

Các đặc trưng cơ bản

Từ điển từ chuẩn là một loại ngôn ngữ tư liệu có cấu trúc tổ hợp. Đó là một tập hợp các thuật ngữ có quan hệ ngữ nghĩa với nhau và được áp dụng cho một lĩnh vực riêng của tri thức. Đó là loại ngôn ngữ đánh chỉ số, nó có những nét đặc trưng cơ bản sau đây:

- Từ điển từ chuẩn xây dựng theo nguyên tắc tổ hợp, ở đó tập hợp các tri thức của một lĩnh vực chuyên môn được phân tích thành một số các khái niệm mà việc tổ hợp chúng theo những cách khác nhau cho phép diễn tả mọi khía cạnh của thông tin.

- Các khái niệm được ghi vào một danh mục các từ chuẩn có cấu trúc linh hoạt. Từ chuẩn là một từ hay một cụm từ diễn tả một khái niệm và được giữ lại sau khi loại bỏ các từ đồng nghĩa. Các từ chuẩn độc lập với nhau nhưng không loại trừ nhau và có thể kết hợp từ này với từ khác.

- Từ điển từ chuẩn nói chung mang tính chất chuyên ngành, nó thường áp dụng cho một lĩnh vực riêng của tri thức và được sử dụng khi có yêu cầu về những thông tin hẹp. Một từ điển từ chuẩn có thể bao gồm từ vài trăm từ đến 20 nghìn từ, trung bình thường có hai hoặc ba nghìn từ.

- Từ điển từ chuẩn có ưu điểm ở tính đặc thù, tính mềm dẻo linh hoạt và khả năng mô tả thông tin đầy đủ và linh hoạt của nó. Với một số lượng tương đối nhỏ các đơn vị từ

vựng, bằng các phép toán logic, ta có thể xây dựng được các “mệnh đề” mà trên thực tế nó có thể diễn tả được bất kỳ khái niệm nào. Tuy nhiên nó đòi hỏi sự biên tập khá công phu.

Mỗi từ điển từ chuẩn được xây dựng trên hai yếu tố cơ bản là quan hệ ngữ nghĩa (sémantic) và quy tắc cú pháp (syntax).

Các quan hệ ngữ nghĩa

Các yếu tố ngữ nghĩa của từ điển từ chuẩn bao gồm:

- Các từ chuẩn.
- Các từ kiểm tra (không phải từ chuẩn).
- Các từ rỗng.

Giữa các từ chuẩn có quan hệ ngữ nghĩa với nhau. Có ba loại quan hệ:

- Quan hệ tương đương.
- Quan hệ cấp bậc.
- Quan hệ lân cận.

Các quan hệ đó giúp ta sử dụng chính xác các từ chuẩn để biểu diễn thông tin. Các quan hệ này được mô tả bởi các mã, gọi là mã quan hệ. Dưới đây là các mã quan hệ dùng trong các nước nói tiếng Pháp:

TG: Thuật ngữ chung, còn gọi là thuật ngữ phái sinh (terme générique).

TS: Thuật ngữ riêng, còn gọi là thuật ngữ đặc thù (terme spécifique).

Để chỉ mức độ riêng (đặc thù) sâu hơn ta dùng TS1, TS2, TS3...

EP: có nghĩa là “dùng cho” (employer pour), để chỉ những từ của ngôn ngữ tự nhiên mà từ chuẩn đã được dùng thay thế cho nó.

VA: có nghĩa là cũng xem (voir aussi), để chỉ các từ gần gũi.

Đó là các mã cho ta thấy mối quan hệ giữa các từ chuẩn và làm rõ thêm ý nghĩa sử dụng của chúng.

Ví dụ:

DESSIN

EP Croquis

TG ARTS GRAPHIQUES

TS1 DESSIN TECHNIQUE

Croquis

VA DESSIN

Trong ví dụ trên DESSIN là từ chuẩn được dùng thay cho (EP) từ Croquis, DESSIN có thuật ngữ chung (TG) là ARTS GRAPHIQUES và thuật ngữ riêng (TS1) là DESSIN TECHNIQUE. Còn khi gặp từ Croquis thì hãy xem (VA) từ DESSIN để sử dụng.

Trình bày từ điển từ chuẩn

Việc trình bày một từ điển từ chuẩn có thể thực hiện dưới các dạng:

- *Các bảng tra từ vựng của từ chuẩn*, có thể trình bày dưới dạng danh mục hệ thống, danh mục theo cấp bậc, danh mục theo vần chữ cái, danh mục theo nguyên tắc hoán vị.

- *Các biểu diễn đồ thị*, có thể bằng sơ đồ mũi tên, sơ đồ hình cây hay sơ đồ đồng tâm.

Dưới đây là một ví dụ minh hoạ cho các cách trình bày trên. Để bảo đảm tính chuẩn xác của ví dụ, chúng tôi để nguyên các từ chuẩn dưới dạng gốc bằng tiếng Pháp.

Danh mục theo cấp bậc

1. BEAUX - ARTS

1.1. Architecture

VA Enseignement de l'architecture

1.2. Art graphique

1.2.1. Dessin

VA Bande dessinée

1.2.1.1. Dessin technique

1.2.2. Photographie

1.3. Art plastique

1.3.1. Sculpture

1.3.2. Gravure

Danh mục hệ thống

BEAUX - ARTS

Architecture

TG Beaux - Arts

VA Enseignement de l' architecture

Art graphique

TG Beaux - Arts

TS1 Dessin

VA Bande dessinée

TS2 Dessin technique

TS1 Photographie

Art plastique

TG. Beaux - Arts

TS1. Sculpture

TS1. Gravure

Danh mục chữ cái

Architecture

Art graphique

Art plastique

Bande dessinée

Beaux - Arts

Dessin

Dessin technique

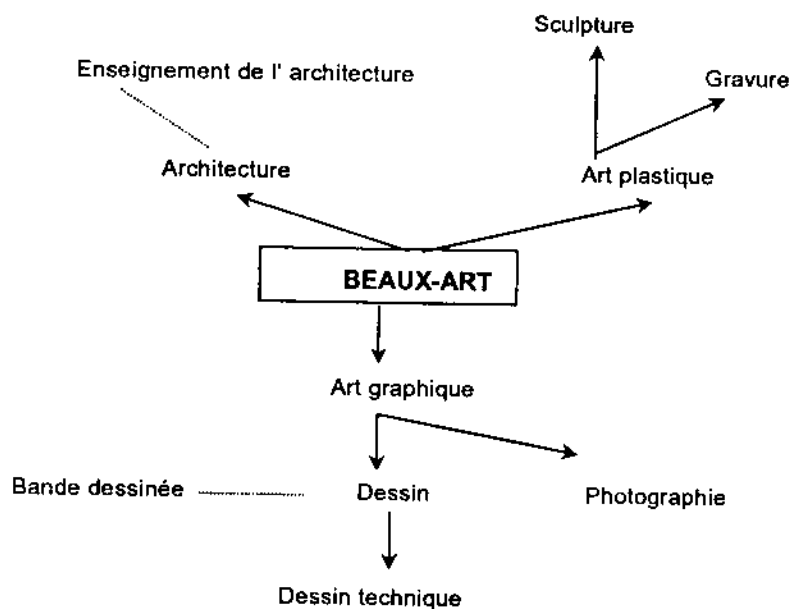
Enseignement de l' architecture

Gravure

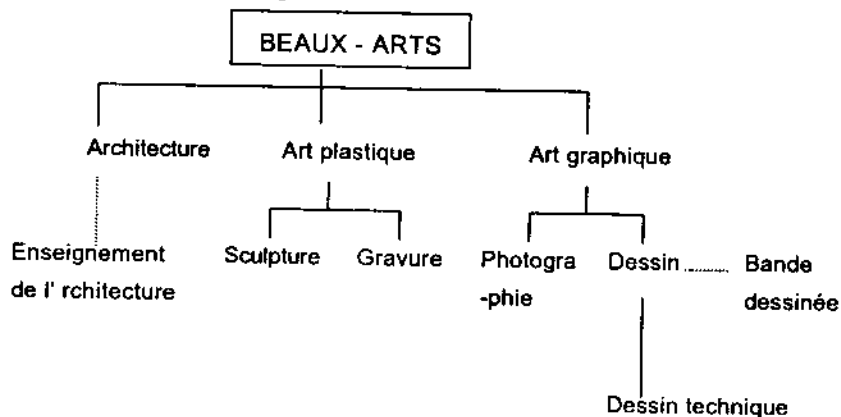
Photographie

Sculpture

Sơ đồ mũi tên



Sơ đồ hình cây



Hình 14. Biểu diễn đồ thị quan hệ giữa các từ chuẩn

Các quy tắc cú pháp

Thành phần cơ bản thứ hai của từ điển từ chuẩn là các quy tắc cú pháp. Các quy tắc cú pháp cho phép ta liên kết các từ chuẩn để diễn tả một thông tin và biết cách đặt một câu hỏi. Cú pháp được xây dựng trên cơ sở của các phép toán của logic mệnh đề, nó cho phép ta thực hiện ba loại quan hệ sau đây giữa các từ chuẩn:

- Quan hệ tương giao, dùng phép hội, còn gọi là phép toán VÀ. Nó cho phép nối hai từ chuẩn có trong bản chỉ mục của cùng một tài liệu.

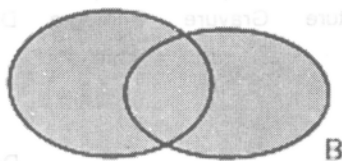
- Quan hệ kết hợp, dùng phép tuyển, còn gọi là phép toán HOẶC. Nó cho phép nối hai từ chuẩn mà ít nhất một trong hai từ đó có trong bản chỉ mục của tài liệu.

- Quan hệ loại trừ, dùng phép phủ định, còn gọi là phép toán KHÔNG. Nó cho phép nối hai từ chuẩn mà từ thứ nhất có trong bản chỉ mục của tài liệu còn từ thứ hai thì không.

Giống như các phép toán tập hợp, người ta có thể biểu diễn các phép toán trên đây bằng sơ đồ Venn.

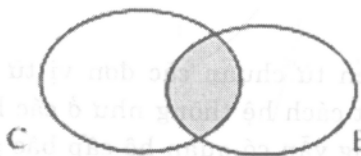
Ví dụ:

1. Phép HOẶC: “Thư viện HOẶC trung tâm thông tin” hay “A HOẶC B”



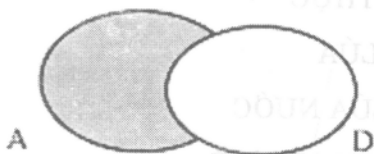
Tất cả các tài liệu đánh chỉ số A hoặc B là tương thích.

2. Phép VÃ: “Tự động hoá VÀ thư viện” hay “C VÀ A”



Tất cả các tài liệu đánh chỉ số C và A là tương thích.

3. Phép TRỪ: “Thư viện TRỪ thư viện trường học” hay “A TRỪ D”



Tất cả các tài liệu đánh chỉ số A trừ các tài liệu đánh chỉ số D là tương thích.

Sử dụng phối hợp các phép toán logic trên ta có thể diễn tả được một mệnh đề phức hợp bất kỳ. Chẳng hạn, mệnh đề: “Tài liệu nói về tự động hoá các thư viện hoặc trung tâm thông tin, trừ các thư viện trường học” có thể biểu diễn dưới dạng một mệnh đề phức hợp sau:

C VÀ (A HOẶC B) TRỪ D

Mệnh đề trên còn gọi là biểu thức tìm hay phương trình tìm.

Phương pháp xác định quan hệ ngữ nghĩa giữa các từ chuẩn

Trong các từ điển từ chuẩn các đơn vị từ ngữ không có quan hệ cấp bậc một cách hệ thống như ở các hệ thống phân loại. Tuy nhiên chúng vẫn có quan hệ cấp bậc giữa các thuật ngữ trong từng cụm từ liên quan đến một chủ đề hoặc một lĩnh vực của tri thức.

Ví dụ: Dưới đây là các từ chuẩn liên quan đến chủ đề "CÂY LƯƠNG THỰC":

CÂY LƯƠNG THỰC

CÂY LÚA

LÚA NƯỚC

LÚA NƯƠNG

HOA MẬU

CÂY NGÔ

CÂY KHOAI

CÂY SẮN

Để làm rõ bản chất quan hệ ngữ nghĩa trên đây và cho nó một biểu diễn toán học chặt chẽ, trước hết ta xây dựng khái niệm trường ngữ nghĩa.

Giả sử có một từ chuẩn nào đó diễn tả khái niệm của một lĩnh vực chuyên môn hay một chủ đề, ví dụ như các từ TIN HỌC, CÂY LƯƠNG THỰC. Khi đó các từ chuẩn đó có thể coi

là tên gọi của một tập hợp bao gồm các đơn vị từ vựng có quan hệ ngữ nghĩa với nhau, nó cho phép làm rõ nghĩa của các đơn vị từ vựng này. Ta gọi một tập hợp như vậy là một trường ngữ nghĩa.

Trong ví dụ nêu ra ở trên ta có trường ngữ nghĩa TIN HỌC, trường ngữ nghĩa CÂY LƯƠNG THỰC.

Ta dùng ký hiệu $\text{dom}(A)$ để chỉ trường ngữ nghĩa A. Với ký hiệu đó ta có $\text{dom}(\text{TIN HỌC})$ là trường ngữ nghĩa TIN HỌC, $\text{dom}(\text{CÂY LƯƠNG THỰC})$ là trường ngữ nghĩa CÂY LƯƠNG THỰC.

Như vậy một từ điển từ chuẩn có thể coi là một tập hợp bao gồm nhiều trường ngữ nghĩa. Các trường này không loại trừ nhau và có thể bao hàm trong nhau.

Ví dụ:

$$\text{dom}(\text{TIN HỌC TƯ LIỆU}) \subset \text{dom}(\text{TIN HỌC})$$

Quan hệ cơ bản trong trường ngữ nghĩa là quan hệ thuộc. Đó là quan hệ không đối xứng giữa từ chuẩn với trường ngữ nghĩa mà nó có liên quan.

Ví dụ:

$$\text{CÂY NGÔ} \in \text{dom}(\text{CÂY LƯƠNG THỰC})$$

Một từ chuẩn có thể liên quan đến nhiều lĩnh vực của tri thức, do đó nó có thể thuộc nhiều trường ngữ nghĩa.

Ví dụ:

$$\text{MÁY TÍNH ĐIỆN TỬ} \in \text{dom}(\text{TIN HỌC})$$

$$\text{MÁY TÍNH ĐIỆN TỬ} \in \text{dom}(\text{TỰ ĐỘNG HOÁ})$$

Quan hệ này giúp ta không sử dụng thừa các từ chuẩn khi đánh chỉ số tài liệu hay khi đặt ra một câu hỏi.

Chẳng hạn nếu trong một bản đánh chỉ số tài liệu đã có từ TIN HỌC thì không cần sử dụng từ MÁY TÍNH ĐIỆN TỬ.

Trong mỗi trường ngữ nghĩa còn luôn tồn tại một quan hệ cấp bậc giữa các từ chuẩn. Đó là quan hệ không đối xứng giữa hai từ chuẩn trong một trường ngữ nghĩa mà từ chuẩn này có mức độ ý nghĩa cao hơn từ chuẩn kia.

Ví dụ:

CÂY LƯƠNG THỰC có mức ý nghĩa cao hơn HOA MẪU

HOA MẪU có mức ý nghĩa cao hơn CÂY NGÔ

Nếu A có mức ý nghĩa cao hơn B, ta nói A là thuật ngữ phái sinh của B, ký hiệu là: A là TG của B (TG = terme générique). Khi đó B là thuật ngữ đặc thù của A, ký hiệu là: B là TS của A (TS = terme spécifique).

Để xác định B có là TS của A hay không ta cần trả lời một trong hai câu hỏi sau:

- B có là A không ?

Ví dụ: CÂY LÚA và CÂY LƯƠNG THỰC

- B có là bộ phận của A hay không?

Ví dụ: VIỆT NAM và ĐÔNG NAM Á

Với khái niệm trường ngữ nghĩa ta thấy ngay rằng: B là TS của A khi và chỉ khi $B \in \text{dom}(A)$ hoặc $\text{dom}(B) \subset \text{dom}(A)$.

Ví dụ: Vì CÂY LÚA $\in \text{dom}(\text{CÂY LƯƠNG THỰC})$ nên CÂY LÚA là TS của CÂY LƯƠNG THỰC.

Nhờ quan hệ cấp bậc giữa các từ chuẩn mà ta có thể sử dụng đúng các từ chuẩn khi đánh chỉ số mô tả nội dung tài liệu, tránh được hiện tượng mất mát thông tin cũng như hiện tượng dư thừa thông tin (nhiều).

Ví dụ: Do CÂY LÚA là TS của CÂY LƯƠNG THỰC nên nếu dùng từ CÂY LÚA để đánh chỉ số một tài liệu nói về CÂY LƯƠNG THỰC thì dẫn đến hiện tượng mất thông tin. Còn nếu dùng từ CÂY LƯƠNG THỰC để đánh chỉ số một tài liệu chỉ nói về CÂY LÚA thì thông tin bị nhiễu, do có nhiều thông tin không thích hợp.

Biên soạn một từ điển từ chuẩn

Việc biên soạn một ngôn ngữ tư liệu nói chung và một từ điển từ chuẩn nói riêng là một công việc nghiêm túc, đòi hỏi nhiều thời gian và công sức. Nó phải được thực hiện theo một kế hoạch được xây dựng trên cơ sở phân tích hệ thống tư liệu, điều kiện vật chất và khả năng chuyên môn sẵn có của đơn vị thông tin.

Để biên soạn một từ điển từ chuẩn ta phải thực hiện các bước sau:

1) Lập danh mục các từ để lựa chọn làm từ chuẩn. Có thể tiến hành theo hai phương pháp:

- Phương pháp phân tích: Rút ra các từ có trong các bản chỉ mục của các tài liệu. Ghi lại các từ có ý nghĩa của các câu hỏi hoặc các bài viết của các nhà nghiên cứu.

- Phương pháp tổng hợp: Tìm các từ có ý nghĩa của lĩnh vực đang xem xét có trong các nguồn tài liệu tra cứu như: các từ điển, các bản chỉ dẫn, các bảng mục lục, các bộ phiếu.

2) Lựa chọn các từ chuẩn theo các tiêu chuẩn lựa chọn xác định, dựa trên tần số sử dụng và giá trị sử dụng của các từ đó.

3) Xây dựng cấu trúc cho bản danh mục các từ chuẩn, bao gồm:

- Lập ra các quan hệ tương đương để chỉ ra các trường hợp đồng nghĩa và tránh sử dụng trùng lặp.
- Lập các quan hệ cấp bậc theo chủ đề hoặc theo diện.
- Lập các quan hệ kết hợp và gần gũi.

4) Tiến hành các bước kiểm tra bằng cách sử dụng để mô tả một số tài liệu tiêu biểu, đặt các câu hỏi và lấy ý kiến đánh giá của các nhà chuyên môn.

5) Tổ chức xuất bản từ điển từ chuẩn.

6) Theo dõi việc sử dụng và tiến hành bổ sung, sửa đổi, hoàn thiện từ điển từ chuẩn trong trường hợp cần thiết.

Việc biên soạn từ điển từ chuẩn là một công trình tập thể, đòi hỏi phải có sự phối hợp của các chuyên gia thông tin và các nhà chuyên môn trên nhiều lĩnh vực liên quan.

2.5. Phân loại, đánh chỉ số và tóm tắt

2.5.1. Phân loại

Phân loại là một công đoạn mô tả nội dung tài liệu, qua đó người ta xác định được nội dung chính của tài liệu và thể hiện nó bằng một thuật ngữ của khung phân loại. Công đoạn này được tiến hành ở giữa dây chuyền tư liệu, trước khi đưa tài liệu vào hệ thống lưu trữ và tìm tin.

Mục đích của phân loại là:

- Sắp xếp các thông tin thành một số lớp tương đối hẹp, mà mỗi lớp đó có thể tương ứng với một lĩnh vực quan tâm đặc biệt của một nhóm người dùng tin.

- Giúp cho việc tổ chức các bộ phiếu hệ thống và sắp xếp các tài liệu trên giá tạo điều kiện thuận lợi cho việc sử dụng vốn tài liệu.

Lợi ích của phân loại là nó cho phép:

- Phối hợp việc sắp xếp tài liệu trên giá với việc tổ chức các bộ phiếu, tạo điều kiện cho việc sử dụng vốn tài liệu.

- Sắp xếp các thông tin thành một số lớp tương đối hẹp, mà mỗi lớp đó có thể tương ứng với một lĩnh vực quan tâm đặc biệt của một nhóm người dùng tin.

Các bước tiến hành phân loại

Công tác phân loại được tiến hành theo các bước như mô hình chung của việc mô tả nội dung tài liệu. Nó bao gồm các bước chính sau:

- Xác định chủ đề chính của tài liệu.

- Xác định lớp của chủ đề chính.

- Tìm chỉ số tương ứng với lớp đã chọn.

- Xây dựng một mã số làm ký hiệu xếp giá cho tài liệu, ngoài chỉ số phân loại.

- Dán ký hiệu xếp giá lên tài liệu.

- Biểu thị chỉ số phân loại trên bản chỉ dẫn thư mục của tài liệu hay trên một bản kê.

Việc lựa chọn các chỉ số phân loại cho tài liệu là một công việc có tính chất phiên dịch, nó đòi hỏi phải có hiểu biết và kinh nghiệm.

Trường hợp đơn giản nhất là các thuật ngữ lựa chọn để mô tả nội dung chính của tài liệu có ngay trong khung phân loại đang dùng. Tuy nhiên nhiều khi thuật ngữ cần lựa chọn không có trong khung phân loại. Khi đó ta phải tìm một từ chuẩn tương ứng với mức độ chính xác nhất. Muốn vậy phải có hiểu biết về chủ đề, tra cứu thêm ở các từ điển, các bách khoa thư để hiểu biết thêm về chủ đề đó, sau đó tìm trong bảng danh mục hệ thống một từ chuẩn thích hợp nhất. Việc kiểm tra lại thông qua bản danh mục hệ thống xem từ chuẩn đang xét có thể hiện đúng ý nghĩa và nội dung của tài liệu hay không là rất cần thiết.

2.5.2. Đánh chỉ số

Đánh chỉ số là một dạng mô tả nội dung tài liệu, mà ở đó người ta chọn ra những thuật ngữ thích hợp nhất để trình bày nội dung và những khái niệm mà tài liệu đề cập tới.

Nội dung của tài liệu được thể hiện bằng một tập hợp các từ vựng của ngôn ngữ tư liệu mà hệ thống đó sử dụng như từ điển từ chuẩn, danh mục các từ khoá. Các thuật ngữ chọn ra được sắp xếp sao cho nó tạo thành như một dãy các chỉ dẫn giúp cho việc tìm tin.

Đánh chỉ số là một việc trung tâm của mọi hệ thống thông tin, nó cho phép ta có thể lưu trữ và tìm kiếm thông tin. Nó được tiến hành ở giữa dây chuyền tư liệu.

Sản phẩm của đánh chỉ số là một bản chỉ mục (index). Đó là một danh mục các thuật ngữ có ý nghĩa.

Các bản chỉ mục phục vụ cho việc tìm tin và lựa chọn ra các thông tin để trả lời yêu cầu của người dùng tin.

Việc đánh chỉ số có thể tiến hành ở nhiều mức độ khác nhau, có mức độ trung bình, có mức độ sâu sắc. Cũng như mọi hình thức mô tả nội dung, việc đánh chỉ số đôi khi cũng có tính chọn lọc, tức là nó chỉ giữ lại những nội dung của tài liệu mà người dùng tin quan tâm.

Các cán bộ được giao nhiệm vụ đánh chỉ số phải là những người có chuyên môn và kinh nghiệm, đồng thời phải am hiểu lĩnh vực chuyên môn mà tài liệu đề cập tới. Mặt khác cũng phải quan tâm nắm bắt được nhu cầu của người dùng tin.

Các bước tiến hành đánh chỉ số

Theo mô hình chung của mô tả nội dung tài liệu, đánh chỉ số được tiến hành theo các bước sau:

- Làm quen với tài liệu để có những hiểu biết tổng quát về tài liệu.

- Xác định nội dung chính của tài liệu, đồng thời phải đi đến phát hiện ra những khái niệm mà tài liệu đề cập tới. Tên tài liệu, lời nói đầu, bảng mục lục, các tóm tắt của tác giả, các bảng biểu và minh họa giúp ta rất nhiều để phát hiện ra các khái niệm đó.

- Rút ra những khái niệm thể hiện bằng ngôn ngữ tự nhiên. Các khái niệm rút ra phải có tính chất chọn lọc, tức là chỉ giữ những khái niệm mà người dùng tin quan tâm. Mặt khác lại phải có tính chất đặc trưng, tức là nó thể hiện những đặc trưng chủ yếu của nội dung tài liệu.

- Thể hiện các khái niệm đã rút ra bằng ngôn ngữ tư liệu. Nói chung ta phải cố gắng tìm những thuật ngữ có trong từ điển từ chuẩn hoặc danh mục các từ khoá. Người ta thường dùng danh từ số ít để diễn tả một khái niệm.

- Kiểm tra tính tương thích của các từ chuẩn đã chọn ra.

- Hình thành bản chỉ mục theo quy tắc trình bày đã quy định. Một bản chỉ mục thường bao gồm từ 10 đến 15 từ chuẩn.

Chất lượng của việc đánh chỉ số được đánh giá theo những tiêu chuẩn sau đây:

- Tính đầy đủ, tức là mọi chủ đề nội dung, mọi khái niệm mà tài liệu đề cập tới đều được trình bày đầy đủ trong bản chỉ mục.

- Tính chọn lọc, tức là chỉ giữ lại những thông tin mà người dùng tin quan tâm.

- Tính đặc trưng, tức là các từ chuẩn thể hiện sát nhất nội dung tài liệu và ít dùng các từ chuẩn quá chung hoặc quá riêng đối với những khái niệm thể hiện trong tài liệu.

- Tính đồng nhất, tức là nếu có một người nào khác cũng đánh chỉ số tài liệu ấy thì cũng theo một cách như vậy.

Ngày nay việc tự động hoá công việc đánh chỉ số bằng máy tính điện tử đang được thử nghiệm ở nhiều nước và có nhiều triển vọng.

2.5.3. Tóm tắt

Tóm tắt là một công đoạn mô tả nội dung tài liệu, ở đó người ta cô đọng nội dung tài liệu bằng một bài viết ngắn.

Sản phẩm của tóm tắt là một bản tóm tắt, thể hiện bằng ngôn ngữ tự nhiên.

Tóm tắt là một công việc cùng một lúc đem lại hai lợi ích. Một mặt nó cho phép giảm đáng kể khối lượng thông tin ban đầu. Mặt khác nó làm bật ra những khía cạnh mà người dùng tin quan tâm.

Người ta sử dụng các bản tóm tắt nhằm:

- Phân phối thông tin.
- Chọn lọc thông tin, do người dùng tin thực hiện.
- Tìm kiếm thông tin, nhất là trong các hệ thống thông tin tự động hóa.

Việc ứng dụng máy tính điện tử cho phép sử dụng các bản tóm tắt để rút ra các từ khoá giúp cho việc lưu trữ thông tin và so sánh các từ khoá này với các thuật ngữ của câu hỏi để tìm ra các câu trả lời trong quá trình tìm tin.

Các loại tóm tắt

Có nhiều loại tóm tắt, chúng khác nhau bởi các đặc trưng sau:

- Độ dài của tóm tắt, có thể từ vài chục từ đến vài trăm từ, nhưng đôi khi cũng vượt quá một nghìn từ.
- Mức độ chi tiết của nội dung bản tóm tắt, có loại mang tính chất chỉ dẫn, có loại chứa tính chất thông báo chứa đựng những điểm mà người dùng tin quan tâm.
- Tóm tắt có hay không các yếu tố mang tính chất đánh giá, phân tích và phê phán.

- Tóm tắt toàn bộ tài liệu hay tóm tắt từng phần tài liệu mà người dùng tin quan tâm.

- Tóm tắt do tác giả tài liệu biên soạn hay tóm tắt do các bộ thông tin biên soạn.

Ta có thể kể ra các loại tóm tắt sau đây:

1- Tóm tắt sơ lược: Đó là một tóm tắt chỉ dẫn, trình bày ngắn gọn nội dung tài liệu. Nó đề cập tới tất cả các chủ đề có trong tài liệu và bao gồm khoảng từ 10 đến 50 từ.

2- Tóm tắt thông báo: Bao gồm từ 50 đến 200 từ. Trong trường hợp này ngoài việc đề cập tới tất cả các chủ đề có trong tài liệu, bản tóm tắt còn đề cập tới các quan điểm khác nhau và sự phát triển của tài liệu.

3- Tóm tắt phân tích: Bao gồm từ 100 đến 500 từ. Bản tóm tắt phân tích sâu sắc hơn, ngoài tóm tắt nội dung nó còn có những nhận xét, phân tích và đánh giá. Người ta gọi đó là tóm tắt mang tính chất thông tin.

4- Tóm tắt chọn lọc: Đó là tóm tắt của các cơ quan thông tin chuyên ngành, người ta chỉ trích ra các chủ đề liên quan tới diện đề tài của lĩnh vực chuyên môn mà người dùng tin quan tâm.

5- Tóm tắt phê phán: Trong đó có phần bình luận, phê phán của người phân tích.

6- Bản thu gọn: Đó là một tóm tắt dài, nó trình bày nội dung tài liệu một cách cô đọng và thường dài bằng 10% đến 20% tài liệu.

7- Trích dẫn: Trích ra dưới dạng tài liệu gốc những phần của tài liệu mà người dùng tin quan tâm.

8- Tóm tắt của tác giả: Do tác giả thực hiện để giới thiệu tài liệu khi xuất bản.

Nội dung của tóm tắt

Tóm tắt là bản đúc kết một cách cô đọng nội dung của tài liệu gốc. Tùy theo từng trường hợp nó có thể bao gồm những điểm sau đây:

- Nội dung chủ đề của tài liệu.
- Bản chất của tài liệu (lý thuyết hay thực nghiệm).
- Các phương pháp sử dụng.
- Các kết quả thu được.
- Kết luận và những triển vọng mà tác giả đưa ra.
- Thời gian, địa điểm, bối cảnh của sự việc.
- Đánh giá mức độ giá trị của tài liệu trong trường hợp tóm tắt phê phán.

Nội dung của bản tóm tắt phải được thể hiện sao cho khi đọc tóm tắt người dùng tin có thể biết được nội dung chủ yếu của tài liệu và qua đó xác định xem có cần tìm đến tài liệu gốc hay không. Trong nhiều trường hợp bản tóm tắt phải thay thế được tài liệu gốc trong việc thoả mãn thông tin ban đầu.

Phương pháp làm tóm tắt

Các bước chuẩn bị cho việc làm tóm tắt cũng được tiến hành như các công việc mô tả nội dung khác. Nếu việc làm tóm tắt được tiến hành sau khi đã phân loại và đánh chỉ số tài liệu thì ta có thể sử dụng các kết quả của các bước trên.

- Việc rút ra các thuật ngữ đặc trưng của tài liệu phải được tiến hành một cách có thứ tự. Ở đây ta nên theo dần ý trình bày của tác giả được thể hiện trong các chương mục của tài liệu.

- Các thuật ngữ rút ra không được trích ra dưới dạng riêng lẻ cô lập mà phải được trích ra theo các nhóm hoặc dưới dạng các câu của bản gốc hoặc tạo thành các câu mới thể hiện mối quan hệ giữa chúng.

- Ngôn ngữ của tóm tắt nói chung nên theo ngôn ngữ của tác giả tài liệu gốc. Cách diễn đạt phải sáng sủa, gọn gàng và chặt chẽ. Các thuật ngữ không những phải chính xác mà phải mang một thông tin thực và dễ hiểu đối với người dùng tin. Tránh dùng các từ viết tắt hoặc các ký hiệu không có trong ngôn ngữ hàng ngày.

Chất lượng của tóm tắt thể hiện ở các mặt sau đây:

- Trước hết phải gọn gàng, dù bản tóm tắt dài hay ngắn. Người ta tránh dùng các mệnh đề và các câu mà có thể thay thế bằng các từ. Tuy nhiên yêu cầu gọn gàng không được ảnh hưởng đến tính chính xác. Người ta phải dùng những cách thể hiện chính xác, đặc trưng và ngắn gọn.

- Tính đầy đủ, tức là bản tóm tắt phải thể hiện đầy đủ nội dung tài liệu và dễ hiểu.

- Tính khách quan, tức là nội dung tài liệu phải được mô tả một cách khách quan. Phải loại trừ mọi đánh giá, bình luận có tính chủ quan của người làm tóm tắt. Trong trường hợp tóm tắt có phê phán, tính khách quan cũng phải thể hiện rõ ràng.

3. CÁC SẢN PHẨM THÔNG TIN

Sản phẩm thông tin là kết quả của quá trình xử lý thông tin (bao gồm biên mục, phân loại, đánh chỉ số, làm tóm tắt, phân tích tổng hợp tin) nhằm thoả mãn nhu cầu thông tin của xã hội.

Dưới đây là một số sản phẩm thông tin chủ yếu.

3.1. Ấn phẩm thông tin thư mục

Ấn phẩm thông tin thư mục là kết quả của quá trình biên mục. Đó là các ấn phẩm thư mục do cơ quan thông tin xuất bản.

Có ấn phẩm thư mục quốc gia, ấn phẩm thư mục chuyên ngành, ấn phẩm thư mục chuyên đề.

Việc sử dụng máy tính trong công tác thông tin thư viện cho phép các đơn vị thông tin - thư viện xây dựng các cơ sở dữ liệu thư mục. Các biểu ghi tra cứu thư mục ở đây thường bao gồm cả tóm tắt, chú giải và tập hợp các từ khoá. Cơ sở dữ liệu thư mục trở thành một công cụ lưu trữ và tìm kiếm thông tin tự động hoá không thể thiếu được trong các cơ quan thông tin thư viện hiện nay.

Từ các cơ sở dữ liệu thư mục người ta có thể in ra các phiếu mục lục, xuất bản các ấn phẩm thư mục như là một sản phẩm đầu ra.

Với công nghệ thông tin hiện đại, ngày nay nhiều cơ sở dữ liệu thư mục được sản xuất dưới dạng đĩa quang CD-ROM. Với những ưu điểm sẵn có như kích thước nhỏ, dung lượng lớn, độ bền cao, có phần mềm khai thác trực tiếp nằm ngay trên đĩa, các CD-ROM cơ sở dữ liệu thư mục trở thành một

sản phẩm thông tin quan trọng trong các dịch vụ phổ biến thông tin hiện nay.

Ví dụ: Ấn phẩm thư mục British National Bibliography (BNB) xuất bản từ năm 1950, được biên tập từ tất cả các xuất bản phẩm nộp lưu chiểu cho Thư viện Quốc gia Anh từ trên 1000 nhà xuất bản ở Anh quốc. BNB được xuất bản hàng tuần và bản cuối cùng của mỗi tháng lại chứa bản chỉ mục theo tác giả, nhan đề và chủ đề của cả tháng. Từ năm 1989 BNB bắt đầu được phát hành dưới dạng CD-ROM.

3.2. Tạp chí tóm tắt

Tạp chí tóm tắt là sản phẩm thông tin, được thể hiện dưới dạng một ấn phẩm định kỳ, trong đó trình bày các bản tóm tắt của các công trình nghiên cứu khoa học đã được công bố.

Tạp chí tóm tắt là công cụ đắc lực của các nhà nghiên cứu, các cơ quan thông tin, xuất bản trong việc khắc phục những trở ngại do quá trình phân tán thông tin gây ra và là công cụ giúp họ trong việc tìm tin theo nội dung được nhanh chóng và thuận lợi.

Chất lượng của tạp chí tóm tắt thể hiện ở các yếu tố sau đây: khả năng bao quát nguồn tin, tính đầy đủ, tính cập nhật của thông tin được cung cấp và chất lượng các bài tóm tắt.

Khi xây dựng tạp chí tóm tắt người ta phải quan tâm đến cách trình bày, thể hiện qua cách sắp xếp trong bảng chính và các bảng tra cứu bổ sung sao cho người dùng tin dễ truy nhập tới các phần tử thông tin trong tạp chí tóm tắt.

Ngoài dạng ấn phẩm truyền thống in trên giấy, ngày nay với công nghệ thông tin hiện đại, tạp chí tóm tắt còn được

xuất bản dưới dạng CSDL on-line và CSDL lưu trữ trên CD-ROM.

Ví dụ: "Tập chí tóm tắt tài liệu khoa học và công nghệ Việt Nam" là ấn phẩm tóm tắt do Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia tổ chức biên soạn và phát hành 12 số/năm, bao quát thông tin về tất cả các dạng tài liệu của Việt Nam như bài tạp chí, tài liệu hội nghị, báo cáo kết quả nghiên cứu, luận án tiến sĩ, phó tiến sĩ bảo vệ trong nước, đề tài nghiên cứu trong tất cả các lĩnh vực khoa học, công nghệ và môi trường. Tập chí tóm tắt gồm hai phần chính: phần tóm tắt và các bảng mục lục (mục lục tác giả, mục lục chủ đề và mục lục địa lý).

Các yếu tố mô tả trong tạp chí tóm tắt bao gồm:

- 1- Chỉ số đề mục quốc gia
- 2- Mã số tài liệu
- 2- Số thứ tự trong tạp chí tóm tắt
- 4- Tác giả
- 5- Nhan đề tài liệu
- 6- Nguồn
- 7- Số liệu xuất bản
- 8- Tóm tắt (chú thích)
- 9- Từ khoá chủ đề

Từ năm 1987 trung tâm đã sử dụng chương trình CDS/ISIS để xây dựng cơ sở dữ liệu Tài liệu KHCN Việt Nam, viết tắt là STD cho Tập chí tóm tắt. Đến năm 1999 STD đã có khoảng 40.000 biểu ghi và ngày nay thông tin trong Tập chí

tóm tắt có thể tra cứu trực tuyến trên mạng thông tin VISTA theo nhiều điểm tiếp cận thông tin khác nhau như: chỉ số của khung đề mục, tác giả, nhan đề hoặc từ khóa...

3.3. Tổng luận khoa học

Tổng luận là bài trình bày tổng hợp một cách khoa học và có hệ thống về các vấn đề được xem xét, cùng với sự phân tích hiện trạng, mức độ và xu hướng phát triển của chúng.

Một hoặc một số bài tổng luận có thể được tập hợp lại và xuất bản dưới hình thức một tài liệu, gọi là ấn phẩm tổng luận.

Tổng luận thường được các nhà khoa học có trình độ cao biên soạn từ nhiều nguồn tài liệu khác nhau nên là loại sản phẩm thông tin đặc biệt có giá trị, nhất là đối với các nhà nghiên cứu và các cán bộ lãnh đạo, quản lý. Có thể coi tổng luận là sản phẩm giáp ranh của hai lĩnh vực hoạt động: nghiên cứu và thông tin.

Căn cứ vào phạm vi bao quát của nội dung, người ta phân biệt: tổng luận khoa học chuyên ngành và tổng luận khoa học đa ngành.

Ngoài ra, còn có một loại tổng luận đặc biệt, thường do các trung tâm phân tích tổng hợp tin, các trường đại học, các cơ quan nghiên cứu khoa học... tổ chức biên soạn và xuất bản đều đặn hàng năm, theo một chu kỳ nhất định, gọi là ấn phẩm tổng luận hàng năm.

Ví dụ: Dưới đây là một số tổng luận khoa học và công nghệ do Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia tổ chức biên soạn và xuất bản:

- Môi trường châu Á - Thái Bình Dương.
- Nhà nước trong một thế giới đang chuyển biến.
- Chính sách và chiến lược quốc gia về Công nghệ thông tin.
- Môi trường biển và quản lý môi trường vùng ven bờ biển Việt Nam.
- Dịch vụ tư vấn kỹ thuật ở các nước đang phát triển.
- Năng lực công nghệ và chiến lược đổi mới doanh nghiệp.
- Tình hình và xu thế phát triển giáo dục đại học ở một số nước khu vực châu Á - Thái Bình Dương.

v.v...

Loại ấn phẩm này trước đây chỉ phát hành trên giấy, nay được phát hành dưới dạng tài liệu điện tử và người dùng tin có thể đọc trên mạng thông tin VISTA.

3.4. Cơ sở dữ liệu

Cơ sở dữ liệu là tập hợp các dữ liệu về các đối tượng cần được quản lý, được lưu trữ đồng thời trên các vật mang tin của máy tính điện tử và được quản lý theo một cơ chế thống nhất, nhằm giúp cho việc truy nhập và xử lý dữ liệu được dễ dàng và nhanh chóng.

CSDL được quản lý bởi một hệ quản trị CSDL. Đó là một hệ thống các phần mềm, bao gồm các chương trình giúp người sử dụng quản lý và khai thác CSDL theo ba chức năng sau đây:

- Mô tả dữ liệu.

- Cập nhật dữ liệu.
- Tìm kiếm dữ liệu.

Có nhiều cách phân loại CSDL. Trong lĩnh vực thông tin - thư viện, căn cứ theo bản chất của thông tin được lưu giữ, người ta phân biệt ba loại CSDL chính sau:

CSDL thư mục

Dữ liệu trong CSDL thư mục là tin tức về bản thân tài liệu. Chúng chứa các thông tin cấp hai, tức là các dữ liệu thư mục và các dữ liệu bổ sung, chứ không phải là văn liệu gốc. CSDL thư mục được trình bày tương tự như trong ấn phẩm thư mục hay trong tạp chí tóm tắt. Nó bao gồm các dữ liệu thư mục (tác giả, nhan đề, các yếu tố xuất bản, các đặc trưng số lượng,...), các chỉ số phân loại, tóm tắt, chú giải, từ khoá... Đối tượng xử lý là các tài liệu chuyên khảo, bài trích báo, tạp chí, kỷ yếu hội nghị, hội thảo khoa học, báo cáo khoa học, luận án, sáng chế... Nói tóm lại là mọi đối tượng đã được tư liệu hoá.

CSDL thư mục bao gồm các tra cứu về tài liệu gốc, cùng với các chỉ dẫn giúp người dùng tin có thể nhận được tài liệu gốc ở một thư viện.

Các CSDL thư mục đưa lại nhiều lợi ích. Trước hết nó cho phép người sử dụng truy nhập trực tiếp và tức thì các thông tin thư mục trong kho dữ liệu và thông tin cần tìm hiển thị ngay trên màn hình. Mặt khác nó cho phép tạo ra nhiều sản phẩm trung gian như là những sản phẩm đầu ra của CSDL, đó là các ấn phẩm thư mục, các tạp chí tóm tắt, các bộ phiếu mục lục hay các tài liệu vi dạng (microfilm, microfiche).

CSDL dữ kiện

Khác với CSDL thư mục, CSDL dữ kiện chứa các thông tin cấp một. Nội dung thông tin chứa trong CSDL là các thông tin dữ kiện. Đó là các số liệu hay các dữ kiện cụ thể về các đối tượng, các quy trình hoặc phương pháp. Thông tin dữ kiện được lưu trữ chủ yếu dưới dạng số, ngoài ra cũng được trình bày dưới dạng văn bản, hình ảnh, biểu đồ, đồ thị,...

Trong CSDL dữ kiện, thông tin đã được xử lý, biên tập và có thể dùng trực tiếp, không cần tới một nguồn dữ liệu bổ sung nào. Lợi ích của nó là người dùng tin không cần mất thời gian tìm đến tài liệu gốc. Điều đó làm cho CSDL dữ kiện trở thành một công cụ hữu ích cho các nhà quản lý, các nhà khoa học và các kỹ sư.

CSDL toàn văn

CSDL toàn văn là sự mở rộng logic tới các hệ thống thông tin hiện đại. Chúng chứa các thông tin cấp một, đó là toàn văn bản của tài liệu cùng với các dữ liệu thư mục và các dữ liệu chủ đề khác. Ngày nay một số tạp chí và dịch vụ báo chí đã được cung cấp dưới dạng CSDL toàn văn. Đó là trường hợp của các báo điện tử.

CSDL toàn văn có nhược điểm là nó không chứa các bảng biểu và các hình ảnh minh họa. Ngoài ra CSDL toàn văn đòi hỏi phải có bộ nhớ với dung lượng lớn, đầu tư ban đầu tốn kém. Tuy nhiên ngày nay, với sự phát triển của CNTT, vấn đề này đã và đang được giải quyết.

Các CSDL toàn văn đã cải thiện đáng kể việc cung cấp các nguồn thông tin cấp một, mở ra triển vọng mới trong hoạt động của các hệ thống giao lưu thông tin.

Theo phạm vi bao quát của đề tài, các CSDL phân thành: CSDL đa ngành, CSDL đơn ngành, CSDL theo chủ đề. Ví dụ: CSDL mỏ và địa chất, CSDL về môi trường, CSDL về dân số, v.v...

Các CSDL cũng có thể phân theo loại hình tài liệu. Những CSDL này thường hướng vào một loại hình tài liệu cụ thể và thường có nội dung đa ngành. Ví dụ: CSDL sách, CSDL tạp chí, CSDL về các phát minh sáng chế, CSDL về các công trình khoa học, v.v...

Các CSDL được lưu trữ trong bộ nhớ của máy tính điện tử, tức là ghi trên các băng từ, đĩa từ. Từ đó CSDL cũng có thể được xuất bản dưới dạng một ấn phẩm, tức là được ghi ra trên giấy. Ngày nay, với sự phát triển của công nghệ thông tin hiện đại, CSDL còn được ghi và lưu trữ trên đĩa quang CD-ROM. Các CSDL toàn văn trên CD-ROM có thể lưu trữ các thông tin dưới dạng văn bản, âm thanh, hình ảnh tĩnh và động và truy nhập rất thuận lợi.

Với những ưu điểm nổi bật như kích thước nhỏ, dung lượng lớn, độ bền cao, có phần mềm khai thác thông tin ghi ngay trên đĩa... các CD-ROM cơ sở dữ liệu trở thành một sản phẩm thông tin quan trọng, không thể thiếu được trong các dịch vụ phổ biến thông tin hiện nay.

Những CSDL đầu tiên xuất hiện vào cuối những năm 60 và nhanh chóng trở thành một nguồn thông tin mới bên cạnh các kho lưu trữ thông tin truyền thống của các thư viện và cơ quan lưu trữ. Sự ra đời và phát triển của các CSDL tạo nên một biến chuyển về chất trong hoạt động của các cơ quan thông tin - thư viện. Đánh giá về sự kiện này, tại hội thảo lần thứ 41 của Liên đoàn Tư liệu Quốc tế (FID) về chủ đề "Thách

thức của CNTT đối với hoạt động thông tin - thư viện", tổ chức tại Hồng Kông năm 1982, J.Martin - một chuyên gia thông tin học người Mỹ đã nói: "Những nhà sử học sẽ nhìn nhận sự xuất hiện của các CSDL, cũng như khả năng phát triển của chúng là một bước biến đổi về chất của sự tiến hoá xã hội và có lẽ nó có tầm quan trọng lớn hơn so với sự tiến bộ của báo in".

Để thấy rõ tầm vóc đó của các CSDL trong hoạt động thông tin tư liệu, ta nêu ở đây một ví dụ: Viện Thông tin Khoa học Kỹ thuật Quốc gia Pháp (INIST) sản xuất hai CSDL rất nổi tiếng là PASCAL và FRANCIS.

PASCAL là CSDL thư mục, đa ngành, đa ngôn ngữ, bao quát các lĩnh vực khoa học công nghệ và y học:

- Bắt đầu xây dựng từ năm 1973.
- Số biểu ghi tra cứu hơn 10 triệu.
- Mức tăng trưởng hàng năm khoảng 600000 biểu ghi.
- Ngôn ngữ sử dụng: Pháp, Anh, Tây Ban Nha.

Dữ liệu trong CSDL PASCAL là các bài báo của 8200 tạp chí khoa học của Pháp và thế giới, các báo cáo khoa học, các luận án, các tổng kết hội nghị, các công trình nghiên cứu.

Từ năm 1987, CSDL PASCAL được ghi trên CD-ROM. Mỗi đĩa chứa nội dung một năm của CSDL này, khoảng 600000 biểu ghi. Ngoài ra hàng tháng CSDL PASCAL còn xuất bản 63 ấn phẩm thư mục theo từng lĩnh vực, dưới dạng giấy hay microfiche.

FRANCIS là CSDL thư mục, đa ngành, đa ngôn ngữ, bao quát các lĩnh vực khoa học xã hội, nhân văn và kinh tế:

- Bắt đầu xây dựng từ năm 1972.
- Số biểu ghi tra cứu: 1,5 triệu.
- Mức tăng trưởng hàng năm khoảng 80 000 biểu ghi.
- Ngôn ngữ sử dụng: Pháp, Anh.

85% dữ liệu trong CSDL FRANCIS là các bài báo của 7200 tạp chí khoa học của Pháp và thế giới.

Đĩa quang đầu tiên của FRANCIS ghi lại nội dung của CSDL này trong 7 năm, từ 1984 đến 1990, chứa khoảng 500000 biểu ghi. Hàng năm 1 đĩa cập nhật toàn bộ nội dung của FRANCIS được xây dựng trong năm đó. Ba tháng một lần CSDL FRANCIS xuất bản 16 ấn phẩm thư mục theo từng lĩnh vực, dưới dạng giấy hay microfiche.

Các CSDL trên đều có thể truy nhập thông qua các cơ quan dịch vụ thông tin, minitel hoặc CD-ROM.

Ví dụ: CSDL PASCAL do các cơ quan dịch vụ thông tin Questel, Dialog, ESA-IRS và Datastar đảm nhiệm, còn CSDL FRANCIS thì do Questel đảm nhiệm.

Chúng ta có thể kể thêm cơ sở dữ liệu AGRIS do Tổ chức nông nghiệp và lương thực của Liên hợp quốc (FAO) xây dựng. AGRIS là CSDL của hệ thống thông tin nông nghiệp quốc tế, bao quát các lĩnh vực nông nghiệp, lâm nghiệp, thủy sản và công nghiệp thực phẩm.

- Bắt đầu xây dựng từ năm 1975.
- Số biểu ghi tra cứu: 2 triệu.
- Mức tăng trưởng hàng năm: 100000 biểu ghi.

Các dữ liệu nhập vào AGRIS không chỉ bao gồm các bài báo trong các ấn phẩm định kỳ, các sách khoa học kỹ thuật mà cả các tài liệu không công bố như: các báo cáo tổng kết, các công trình nghiên cứu khoa học, các phát minh sáng chế, các tiêu chuẩn, bản đồ và cả các tài liệu khó với tới khác trên các lĩnh vực phát triển của nông, lâm nghiệp.

Chương 5

LƯU TRỮ VÀ TÌM KIẾM THÔNG TIN

1. LƯU TRỮ THÔNG TIN

1.1. Nguyên tắc lưu trữ thông tin

Sau khi xử lý hình thức, nội dung tài liệu (mô tả thư mục, phân loại, đánh chỉ số, làm tóm tắt) ta phải lưu trữ các thông tin đã có. Ta gọi đó là đưa thông tin vào bộ nhớ của hệ thống.

Việc lưu trữ thông tin được thực hiện trên các vật mang tin khác nhau:

- Các bộ phiếu truyền thống (phương tiện thủ công).
- Các phiếu lỗ mép, phiếu lỗ soi (phương tiện bán tự động).
- Các biểu ghi trong các tệp dữ liệu trên các đĩa từ, đĩa quang (phương tiện tự động hoá).

Để xây dựng mô hình chung cho các hình thức lưu trữ thông tin kể trên, trước hết ta phải tìm hiểu và làm rõ: những đặc trưng cơ bản của một tài liệu là gì ?

Mỗi tài liệu tự trị trung đều có hai đặc trưng cơ bản:

1- Đặc trưng hình thức, thể hiện bằng các dữ liệu thư mục như: tác giả, nhan đề, các yếu tố xuất bản, dạng của tài liệu, v.v... Những dữ liệu này là những đặc điểm vốn có ở mỗi tài liệu, nó cho phép mô tả tài liệu đó và nhận biết tài liệu này một cách chính xác. Và để nhận biết được tài liệu ta chỉ cần dùng một mã số gán cho mỗi tài liệu, đó thường là số ký hiệu nhập của tài liệu.

2- Đặc trưng nội dung, có thể thể hiện bằng một chỉ số phân loại, một bản tóm tắt, hay tiện lợi nhất là một bản chỉ mục, nó bao gồm các từ khóa hoặc từ chuẩn thể hiện nội dung chủ đề của tài liệu đó. Những yếu tố dữ liệu này cho phép lưu trữ và tìm những thông tin có trong tài liệu.

Gọi D là tập hợp các tài liệu thể hiện qua số ký hiệu của các tài liệu đó, T là tập hợp các chủ đề của tài liệu thể hiện qua các từ khóa. Khi đó giả sử ta có:

$$D = \{ 1, 2, 3, 4, 5, 6, 7, 8 \}$$

$$T = \{ A, B, C, D, E, F, G, H \}$$

Ở đây để định ý ta giả sử các từ khóa:

$$A = \text{"Cây lúa"} \quad E = \text{"Cây ngô"}$$

$$B = \text{"Hoa mầu"} \quad F = \text{"Thuỷ lợi"}$$

$$C = \text{"Sâu bệnh"} \quad G = \text{"Kỹ thuật trồng trọt"}$$

$$D = \text{"Đất phèn"} \quad H = \text{"Chăn nuôi"}$$

Giữa tập D các tài liệu và tập T các từ khóa ta xây dựng một quan hệ hai ngôi, ký hiệu là R , xác định như sau:

$i \in R \times x$ nếu "Tài liệu i nói về chủ đề x "

Quan hệ này xác định một tập con R của tích Đề các $D \times T$ và có thể biểu diễn bằng một ma trận gồm hai đầu vào D và T , ở đó nếu cặp $(i, x) \in R$ thì trên ô tương ứng của ma trận ta đặt một dấu +. Bằng cách đó giả sử ta lập được bảng sau:

Từ khoá Tài liệu	A	B	C	D	E	F	G	H
1	+		+				+	
2		+		+	+			+
3	+					+	+	+
4		+		+		+		
5	+	+					+	
6			+			+		
7		+	+	+	+		+	
8	+	+	+			+		

Quan hệ này thể hiện mối liên quan giữa hai nguyên tắc lưu trữ thông tin, đó là:

- Lưu trữ theo tài liệu.
- Lưu trữ theo nội dung chủ đề của tài liệu.

Ma trận cho bởi bảng trên biểu diễn quan hệ giữa hai hệ thống sắp xếp vốn tư liệu tương ứng của đơn vị thông tin, gọi là "ma trận tư liệu".

Nếu cắt ma trận theo chiều ngang ta thấy ở đây thể hiện nguyên tắc: mỗi tài liệu ứng với một phiếu nêu lên những khái niệm và nội dung mà tài liệu đề cập tới.

Ví dụ: Tài liệu 1 ứng với $t_1 = \{ A, C, G \}$

Tài liệu 2 ứng với $t_2 = \{ B, D, E, H \}$

Tài liệu 3 ứng với $t_3 = \{ A, F, G, H \}$

Tài liệu 4 ứng với $t_4 = \{ B, D, F \}$

...

Ở đây $t_1, t_2, t_3, \dots$ là các tập con của T , tức là các phần tử của tập $P(T)$ bao gồm tất cả các bộ phận của T . Trong công tác tư liệu thì đó chính là các bản chỉ mục của các tài liệu tương ứng. Vì vậy việc đánh chỉ số tài liệu có thể coi là một ánh xạ I từ tập D các tài liệu đến tập $P(T)$ các bộ phận của T , trong đó T là tập các từ khoá thể hiện nội dung và chủ đề của các tài liệu:

$$I: D \rightarrow P(T)$$

Ánh xạ này ứng tài liệu thứ i với bản chỉ mục t_i của tài liệu đó. Bản chỉ mục này thường được ghi trên các phiếu mô tả thư mục hoặc trên các biểu ghi của các cơ sở dữ liệu thư mục, ở mục "Từ khoá". Các phiếu này được lập theo thứ tự nhập của tài liệu và chúng phục vụ cho việc lưu trữ và tìm kiếm thông tin nên được gọi là các mẫu tìm.

Nếu cắt ma trận theo chiều dọc thì ta thấy ở đây thể hiện nguyên tắc: mỗi từ khoá đặc trưng cho nội dung chủ đề của tài liệu ứng với một phiếu, ghi số tất cả các tài liệu có nội dung đề cập đến chủ đề đó.

Ví dụ:

Chủ đề A ứng với $d_A = \{ 1, 3, 5, 8 \}$

Chủ đề B ứng với $d_B = \{ 2, 4, 5, 7, 8 \}$

Chủ đề C ứng với $d_C = \{ 1, 6, 7, 8 \}$

Chủ đề D ứng với $d_D = \{ 2, 4, 7 \}$

Ta nhận thấy $d_A, d_B, d_C, \dots$ chính là các tập con của tập D các tài liệu. Do đó $d_A, d_B, d_C, \dots$ là các phần tử của tập $P(D)$ các bộ phận của D. Vì vậy phép tương ứng trên có thể coi là một ánh xạ J từ tập T các từ khoá vào tập $P(D)$:

$$J : T \rightarrow P(D)$$

Trong công tác thông tin tư liệu các từ khoá A,B,C... gọi là các điểm tiếp cận thông tin, còn ảnh của chúng qua ánh xạ J là $d_A, d_B, d_C, \dots$ được gọi là các bảng đảo. Bảng đảo cũng được gọi là các mẫu tìm. Tập hợp các bảng đảo lập thành bộ phiếu đảo. Chúng có thể là:

- Các phiếu uniterm.
- Các phiếu lỗi soi.
- Các điểm tiếp cận thông tin ghi trong bộ phiếu đảo của các cơ sở dữ liệu thư mục.

Nguyên tắc chung để xây dựng các phiếu uniterm và phiếu lỗi soi là: mỗi phiếu ứng với một từ khoá hoặc từ chuẩn đặc trưng cho nội dung chủ đề của tài liệu. Trên mỗi phiếu người ta ghi số ký hiệu (địa chỉ) của các tài liệu đề cập đến nội dung chủ đề đó. Còn bộ phiếu đảo của các cơ sở dữ liệu thì bao gồm một danh mục các điểm tiếp cận thông tin (các từ khoá, từ chuẩn) có trong cơ sở dữ liệu và nói chung được sắp

xếp theo vần chữ cái. Các điểm tiếp cận thông tin này đều có ghi kèm theo số ký hiệu nhập của các tài liệu mà nội dung của nó có chứa điểm tiếp cận thông tin nêu ra.

1.2. Các phương tiện lưu trữ thông tin truyền thống

Trong các thư viện hay đơn vị thông tin, phương tiện lưu trữ thông tin truyền thống là các bộ phiếu tra cứu hay mục lục truyền thống. Đó là tập hợp có thứ tự các phiếu tra cứu, nó liệt kê và mô tả tài liệu cùng với nội dung chủ đề của chúng.

Mọi hoạt động dịch vụ phục vụ người dùng tin, nhiều khâu trong quá trình xử lý thông tin đều phải tiến hành trên cơ sở khai thác các bộ phiếu. Việc sử dụng các bộ phiếu có thể cho phép ta:

- Xác định tài liệu gốc.
- Tìm tài liệu gốc theo tên tác giả, theo chủ đề hoặc theo địa danh.
- Quản lý vốn tài liệu, vì các bộ phiếu cho ta biết chi tiết vào bất cứ lúc nào cấu tạo và thành phần của vốn tài liệu.

Việc xây dựng các bộ phiếu hay mục lục được thực hiện nhờ các chỉ dẫn thư mục. Các mô tả thư mục và mô tả nội dung cho ta một phiếu chính hay phiếu cơ bản. Nội dung của các mục lục có thể dựa vào một số yếu tố nào đó của chỉ dẫn thư mục, bao gồm cả việc phân loại và đánh chỉ số. Tuy nhiên người ta chỉ xây dựng những mục lục sao cho nó có thể phục vụ cho việc tìm tài liệu và quản lý vốn tài liệu và nó giúp cho người dùng tin có thể tiếp cận các nguồn tin. Mỗi đơn vị

thông tin, tùy theo yêu cầu và đặc điểm của mình, phải xác định xem cần xây dựng loại mục lục nào.

Nói chung, mỗi đơn vị thông tin cần phải xây dựng ít nhất một mục lục tác giả và một mục lục chủ đề. Các mục lục khác thường thấy là mục lục theo địa lý, mục lục theo thời gian và mục lục xếp kho. Hệ thống mục lục chính là bộ nhớ lưu trữ thông tin và chúng trở thành một bộ phận của bộ máy tra cứu của các đơn vị thông tin.

Các loại mục lục

1- Mục lục tác giả: Trình bày theo thứ tự chữ cái các chỉ dẫn về tên tác giả hay tên tài liệu trong trường hợp tác giả khuyết danh hay có từ 3 tác giả trở lên. Người ta cũng đưa vào mục lục tác giả các tác giả phụ như các nhà dịch thuật, các nhà hiệu đính...

Mục lục tác giả cho phép trả lời câu hỏi:

- Đơn vị thông tin có tài liệu này của tác giả này hay không?

- Đơn vị thông tin có các tài liệu do tác giả này viết hay không?

2- Mục lục chủ đề: Trình bày theo thứ tự chữ cái các chỉ dẫn về các tiêu đề hay các từ chuẩn mô tả nội dung tài liệu.

Trong phạm vi mỗi khung đề mục, các chỉ dẫn này được sắp xếp theo thứ tự chữ cái của tên tác giả, tên sách hay theo thứ tự nhập của tài liệu.

Có hai loại mục lục chủ đề:

- Mục lục chủ đề chữ cái, ở đó các điểm tiếp cận được sắp xếp theo thứ tự chữ cái như trong từ điển. Loại mục lục này

cho phép tìm những tài liệu mà người ta chỉ biết chủ đề của nó, hoặc biết những tài liệu về một chủ đề mà đơn vị thông tin có.

- Mục lục chủ đề hệ thống, ở đó các chỉ dẫn được sắp xếp theo một khung phân loại có trước hoặc theo tên hoặc theo chỉ dẫn tương ứng với chủ đề. Loại mục lục này cho phép ta biết diện đề tài của đơn vị thông tin, hoặc tìm tài liệu theo loại chủ đề hoặc lập các bản kê tài liệu mà đơn vị có.

3- Mục lục địa lý: Kê ra các chỉ dẫn liên quan đến tên của một đất nước, một khu vực hành chính hay một vùng sinh thái tự nhiên mà tài liệu đề cập tới. Các điểm tiếp cận có thể sắp xếp theo vần chữ cái hay theo hệ thống.

Loại mục lục này cho phép ta tìm tài liệu liên quan đến một địa danh. Nếu có yêu cầu người ta có thể xây dựng một mục lục theo tên của nơi xuất bản.

4- Mục lục thời gian: Trình bày các chỉ dẫn theo thời gian xuất bản của tài liệu, theo thời gian nhập của tài liệu, hay theo số thứ tự bổ sung tài liệu.

Loại mục lục này thường được sử dụng như một bộ phiếu chính vì nó có thể cập nhật dễ dàng. Nó cho phép có thể tìm tài liệu theo thời gian.

5- Mục lục xếp kho: Sắp xếp các chỉ dẫn theo thứ tự sắp xếp các tài liệu trên giá, nghĩa là theo ký hiệu xếp kho. Mục lục này cần thiết cho việc kiểm kê và giúp cho nhân viên thư viện có thể tìm tài liệu gốc một cách dễ dàng.

6- Mục lục theo loại hình tài liệu: Cho phép tìm dễ dàng một tài liệu theo bản chất của nó, chẳng hạn như: tài liệu phát minh sáng chế, ấn phẩm định kỳ, bản đồ, băng đĩa

nhạc,... Nó được tổ chức theo tên hay theo số đăng ký của tài liệu.

7- Mục lục liên hợp: Tập hợp theo loại hình tài liệu hay theo chủ đề các mục lục của nhiều đơn vị thông tin. Các mục lục liên hợp thường được tổ chức theo tác giả hoặc theo tên tài liệu và được xuất bản dưới dạng một tuyển tập với sự hợp tác của các đơn vị liên quan. Mục lục liên hợp là không thể thiếu được trong sự hợp tác và trao đổi vốn tài liệu giữa các đơn vị thông tin.

1.3. Các phương tiện lưu trữ thông tin bán tự động

1.3.1. Phiếu lỗ mép

Nguyên tắc

Phiếu lỗ mép là một hệ thống lưu trữ thông tin theo nguyên tắc “lưu trữ theo tài liệu”, tức là mỗi tài liệu được trình bày bằng một phiếu. Phiếu này là cơ cấu cơ bản của bộ nhớ và được chia làm hai vùng:

- Vùng giữa của bộ phiếu dùng để mô tả thư mục, ghi những chỉ dẫn thư mục của tài liệu.

- Vùng xung quanh mép phiếu là vùng đục lỗ, từ đó theo một mã số đã quy định, người ta có thể ghi dưới dạng các lỗ khuyết những đặc trưng nội dung và khái niệm của tài liệu, giúp cho việc tìm kiếm tài liệu sau này.

Việc lựa chọn được thực hiện trên toàn bộ phiếu mà người ta lấy ra từng tập một. Người ta cầm lấy một tập phiếu và xiên một que vào lỗ tương ứng với đặc trưng định tìm, rồi bỏ tay ra cho tập phiếu rơi tự do. Những phiếu có bầm lỗ khuyết sẽ rơi xuống và được coi là kết quả lựa chọn.

Nếu một đặc trưng được trình bày bằng nhiều lỗ và nếu người ta muốn tìm nhiều đặc trưng thì phải lập lại thủ thuật trên cho mỗi lỗ, mỗi lần chỉ lấy những phiếu đã rơi xuống và được lựa chọn cho lần trước.

Phiếu được làm bằng bìa cứng, bìa mỏng hoặc giấy Bristol. Loại này do các nhà xuất bản khác nhau cung cấp, có kích thước từ 75×125 mm đến 210×297 mm. Các lỗ có đường kính thay đổi từ 2,5 đến 3 mm và khoảng cách giữa chúng là 5,6 mm hoặc 6,35 mm.

Việc lựa chọn có thể thực hiện bằng tay hoặc bằng một máy chọn có bàn que xiên và có bộ rung điện. Một bộ phiếu như vậy có thể tập hợp được khoảng một nghìn phiếu một lần chọn.

Mã số đăng ký

Mỗi lỗ khuyết hoặc một tổ hợp lỗ khuyết dành cho một đặc trưng chữ số hoặc chữ cái tùy theo mã số được tạo nên trong quá trình xây dựng phiếu.

Người ta đưa ra những cách mã hoá cho phép đăng ký một số lượng tin lớn và có khả năng sử dụng một cách rất mềm dẻo và linh hoạt. Các loại mã số có khá nhiều. Ở đây ta chỉ nêu một vài ví dụ.

Để trình bày các con số ta có thể dùng mã 1.2.4.7. Ở đây người ta bấm lỗ những con số, mà tổng của nó đặc trưng cho số định dùng.

Ví dụ: Muốn trình bày số 193 người ta bấm lỗ vào vị trí các lỗ có đánh dấu như sau:

Người ta xây dựng mỗi phiếu ứng với một từ khoá hoặc từ chuẩn đặc trưng cho nội dung tài liệu (lưu trữ theo nội dung). Mỗi lỗ tượng trưng cho một số đăng ký của tài liệu trong hệ thống. Để ghi một tài liệu vào bộ nhớ, ta phải đục lỗ ở vị trí dành cho nó trong những phiếu ứng với khái niệm mà tài liệu có đề cập tới. Như vậy người ta có trong hệ thống một bộ phiếu giống như phiếu uniterme.

Việc lựa chọn được tiến hành bằng cách đặt chồng các phiếu có những đặc trưng cần tìm trước một nguồn sáng. Những tài liệu giải đáp được câu hỏi sẽ là những phiếu có cùng vị trí đục lỗ trên toàn bộ các phiếu chồng lên nhau đó và vì vậy mà ánh sáng xuyên qua được.

Phiếu lỗ soi làm bằng giấy Bristol nhẹ, có nhiều loại kích thước khác nhau. Trên mỗi phiếu có một mạng các vị trí được xây dựng bằng phương pháp toạ độ. Mỗi vị trí ứng với số một số đăng ký của tài liệu đã được gán cho từ trước. Số lượng các vị trí trên mỗi phiếu ứng với số lượng của tài liệu có thể đưa vào bộ nhớ. Dung lượng của mỗi phiếu có thể từ 5000 tài liệu, như mẫu phiếu SPHINXO, hoặc lên tới 14.000 tài liệu như mẫu phiếu SELECTO (do kỹ sư Gérard Cordonnier đưa ra).

Trong phiếu SELECTO, mỗi phiếu có kích thước 150×200mm. Phiếu bao gồm các ô vuông đánh số theo hoành độ từ 0 đến 99 và theo tung độ từ 0 đến 139. Như vậy sẽ có 14000 vị trí ứng với mỗi ô vuông.

Đưa thông tin vào bộ nhớ

Mỗi tài liệu trước khi đưa vào hệ thống đều được ghi một số đăng ký theo thời gian. Ví dụ muốn đục lỗ một tài liệu nói về "Sâu bệnh của cây lúa ở Việt Nam năm 1975", ta lấy trong

hộp phiếu ra các phiếu ứng với các từ khoá: SÂU BỆNH, CÂY LÚA, VIỆT NAM, 1975.

Nếu tài liệu đưa vào hệ thống có số thứ tự 1509 thì mỗi phiếu trích ra sẽ được đục lỗ ở ô vuông có toạ độ:

$$x = 15, y = 09$$

Việc đục lỗ có thể thực hiện dễ dàng bằng một máy đục lỗ điện tử.

Việc đọc được thực hiện bằng cách sử dụng một cái khung chiếu sáng, cho phép nhìn thấy dễ dàng các ô có lỗ chung với nhau trên các phiếu đặt chồng lên nhau. Người ta có thể chồng 30 phiếu lên nhau, nhưng ít khi có một câu hỏi phức tạp phải dùng đến nhiều từ khoá như vậy.

Sử dụng phiếu lỗ soi có những thuận lợi sau:

- Thời gian đưa vào bộ nhớ không phụ thuộc vào số tài liệu cần nhớ, mà chỉ phụ thuộc vào số lượng các phiếu chồng lên nhau.

- Việc lựa chọn có thể thực hiện với nhiều khái niệm cùng một lúc.

- Tất cả các khái niệm chứa trong một tài liệu đều lần lượt được tham dự việc tra tìm.

Tuy nhiên hệ thống này cũng có bất tiện là sự hạn chế về số lượng đăng ký của nó. Mặt khác kết quả lựa chọn được không phải là bản mô tả tài liệu trả lời câu hỏi mà chỉ là chỉ dẫn về số thứ tự của tài liệu đó mà thôi.

Ngày nay với công nghệ thông tin hiện đại, hầu như người ta không còn dùng phiếu lỗ mép và phiếu lỗ soi nữa. Nhưng nguyên tắc xây dựng của chúng có ý nghĩa rất quan

trọng. Đó cũng chính là những nguyên tắc lưu trữ thông tin trên máy tính điện tử mà ta sẽ trình bày dưới đây. Ta sẽ thấy phiếu lỗ mép tương ứng với bộ phiếu chủ (tệp chủ), còn phiếu lỗ soi tương ứng với bộ phiếu đảo (tệp đảo) trong các CSDL thư mục của các hệ thống tìm tin tự động hoá.

1.4. Các phương tiện lưu trữ thông tin tự động hoá

1.4.1. Lưu trữ thông tin trên máy tính điện tử

Trong các hệ thống thông tin tự động hoá, phương tiện lưu trữ thông tin chính là các thiết bị nhớ của máy tính điện tử. Đó có thể là các băng từ, đĩa từ hoặc đĩa quang. Đĩa từ lại có hai loại đĩa cứng và đĩa mềm. Ở đây thông tin được biểu diễn dưới dạng các số nhị phân, tức là chỉ gồm hai chữ số 0 và 1, ứng với hai trạng thái của các vật mang tin từ tính. Mỗi yếu tố thông tin 0 và 1 ứng với một đơn vị thông tin nhị phân, gọi là bit. Khi đó các chữ cái, chữ số, các ký hiệu của các hệ thống ngôn ngữ viết đều có thể biểu diễn dưới dạng một dãy các số 0 và 1. Một dãy 8 bit gọi là một byte. Mỗi byte biểu diễn một ký tự. Bảng tương ứng giữa các ký tự với các dãy số nhị phân lập thành một hệ thống mã. Với một hệ thống mã 8 - bit ta có thể biểu diễn được 256 ký tự khác nhau. Có hai hệ thống mã quan trọng là:

- EBCDIC (Extended Binary Coded Decimal Interchange Code).
- ASCII (American Standart Code for Information Interchange).

Ví dụ trong hệ mã ASCII :

số 1 có mã 00110001

chữ A có mã 01000001

chữ a có mã 01100001

Tập dữ liệu

Trong máy tính điện tử thông tin được tổ chức và lưu trữ dưới dạng các tập dữ liệu. Các tập dữ liệu lại có thể cấu trúc thành các biểu ghi, mỗi biểu ghi lại gồm nhiều trường. Trường có thể coi là đơn vị dữ liệu được lưu trữ, nó cung cấp thông tin liên quan tới một khía cạnh hay thuộc tính của thực thể được mô tả bởi tập dữ liệu còn biểu ghi cho ta thông tin về một đối tượng của thực thể.

Trong công tác thông tin tư liệu, ta thường gặp các tập dữ liệu lưu trữ thông tin thư mục về các tài liệu cần quản lý. Các tập dữ liệu này bao gồm các biểu ghi thư mục. Biểu ghi thư mục là một dữ liệu có cấu trúc, mà mỗi chỉ dẫn thư mục là một trường.

Ví dụ: Biểu ghi tra cứu thư mục của cơ sở dữ liệu sách bao gồm các trường sau:

- Tác giả
- Tên sách
- Nơi xuất bản
- Nhà xuất bản
- Năm xuất bản
- Số trang
- Ký hiệu phân loại
- Ký hiệu kho
- Tóm tắt
- Từ khoá.

Tập hợp các biểu ghi thư mục được tổ chức thành tập dữ liệu thư mục. Đó là thành phần cơ bản của một cơ sở dữ liệu thư mục.

Có nhiều cách tổ chức các tập dữ liệu, tuy nhiên chúng phải được tổ chức sao cho dễ truy nhập tới các phần tử dữ liệu.

Các tập có thể tổ chức theo nhiều hình thức sau:

Tập kế tiếp

Tập kế tiếp (sequential file) là tập được trình bày thành một dãy các biểu ghi liên tiếp. Để đọc một biểu ghi, ta phải đọc lần lượt các biểu ghi trước nó.

Tập truy nhập trực tiếp

Tập truy nhập trực tiếp (direct access file) có thể coi là một bộ sưu tập các biểu ghi có đánh số thứ tự. Nó có các tính chất sau:

- Việc tra cứu một biểu ghi được tiến hành bằng cách chỉ rõ số thứ tự của nó.

- Người ta có thể thay thế, loại bỏ hoặc bổ sung dễ dàng một biểu ghi.

Tập đảo

Cấu trúc tập đảo bao gồm hai phần:

- Phần thứ nhất là tập chứa tất cả các biểu ghi thư mục của CSDL, gọi là tập sơ cấp hay còn gọi là tập chủ.

- Phần thứ hai là tập đảo kết hợp với tập sơ cấp. Nó bao gồm các giá trị của một số trường của tập sơ cấp, được chọn là các điểm truy nhập thông tin. Đó có thể coi như là một tập

hợp các bảng đảo của các trường trong các biểu ghi mà ta cần tìm.

Như vậy tổ chức của tệp đảo xuất phát từ giá trị của các trường. Nó cho phép ta tiếp cận biểu ghi của tệp sơ cấp mà trường đang xét có giá trị riêng đó.

Để tổ chức tệp đảo thì ở tệp sơ cấp mỗi biểu ghi phải được định vị bằng một “địa chỉ”. Nhờ các “địa chỉ” này mà xuất phát từ một giá trị của bảng đảo ta có thể xác định được biểu ghi chứa các thông tin mà ta cần tìm. “Địa chỉ” này cũng giống như số trang trong mục lục hay trong các bảng chỉ mục (index) của một cuốn sách.

Từ quản lý tệp đến hệ thống quản trị dữ liệu

Nhu cầu tích lũy và xử lý dữ liệu là nhu cầu cơ bản nảy sinh từ hoạt động thực tiễn của con người. Tự động hoá xử lý dữ liệu là một đòi hỏi tất yếu khi cần xử lý nhanh một khối lượng lớn dữ liệu.

Khởi đầu công tác tự động hoá xử lý dữ liệu, người ta lưu trữ thông tin trên các tệp dữ liệu và dùng các chương trình để tìm kiếm, thao tác trên các tệp dữ liệu đó. Đó là tiền thân của các hệ thống cơ sở dữ liệu.

Việc quản trị dữ liệu trên máy tính được thực hiện nhờ hai dạng chương trình sau:

- Các hệ thống quản lý tệp.
- Các hệ thống quản trị cơ sở dữ liệu.

Hệ thống quản lý tệp là chương trình lưu trữ, xử lý và in dữ liệu chứa trong các tệp tách biệt. Hệ quản trị tệp lưu trữ dữ liệu trên bộ nhớ ngoài (băng từ, đĩa từ, đĩa quang) và khai

thác dữ liệu bằng các chương trình. Chương trình là tập hợp các lệnh được viết bằng ngôn ngữ bậc cao, như ngôn ngữ PASCAL, C... cho phép mô tả và xử lý dữ liệu. Các chương trình cần thoả mãn yêu cầu phục vụ nhiều tệp dữ liệu khác nhau. Chương trình đó bảo đảm tính độc lập so với dữ liệu, tức là có thể thay đổi dữ liệu của bộ nhớ ngoài mà không cần thay đổi chương trình. Một khi tệp được phân chia thành các biểu ghi thì tệp được xem là tập hợp các biểu ghi. Tuy nhiên không phải tệp nào cũng được sắp đặt như vậy. Tại một thời điểm chỉ có một tệp được truy nhập và xử lý.

Chức năng cơ bản của hệ quản lý tệp là xử lý các tệp dữ liệu, bao gồm:

- Tạo tệp mới.
- Mở, đóng các tệp.
- Xoá tệp.

Hệ thống quản trị tệp không tự trang bị mà nhờ hệ thống điều hành máy tính thực hiện các chức năng trên. Nó quan tâm nhiều đến các phương pháp tổ chức dữ liệu trên bộ nhớ ngoài. Có hai phương pháp chính tổ chức dữ liệu trên tệp là tổ chức tuần tự và tổ chức khai thác trực tiếp.

Sử dụng hệ thống quản lý tệp có nhiều hạn chế. Thứ nhất nó tốn thời gian khi phải khai thác các tệp chứa các thông tin khác. Thứ hai là nhiều khi phải lưu trữ lại một số dữ liệu, ví dụ như tên và mã số của các nhân viên. Thứ ba là độ tin cậy của dữ liệu không cao, ví dụ một thông tin phải lưu trữ ở nhiều nơi, do sơ suất không cập nhật được ở tất cả các nơi khác nhau, tạo nên sự không nhất quán về thông tin.

Khái niệm cơ sở dữ liệu được hình thành để giải quyết vấn đề liên quan tới hệ thống quản lý tệp. Một CSDL là một tập hợp các tệp dữ liệu có liên quan với nhau, được lưu trữ đồng thời trên các vật mang tin của MTĐT và được quản lý theo một cơ chế thống nhất, giúp cho việc truy nhập và xử lý dữ liệu được dễ dàng và nhanh chóng.

CSDL bao gồm một tệp hoặc một tập hợp các tệp dữ liệu. Thông tin trong các tệp này có thể chia nhỏ thành các biểu ghi, mỗi biểu ghi lại bao gồm một hoặc nhiều trường. Trường là đơn vị cơ sở của dữ liệu và mỗi trường thường chứa các thông tin liên quan một khía cạnh hay một thuộc tính của thực thể được mô tả bởi CSDL. Sử dụng các từ khoá hoặc các lệnh tìm, người dùng tin có thể nhanh chóng lựa chọn ra các biểu ghi thoả mãn yêu cầu tìm tin đặt ra.

Các CSDL được xây dựng và quản lý nhờ các hệ quản trị CSDL. Đó là các phần mềm cho phép ta xây dựng cấu trúc của CSDL, cập nhật dữ liệu (bổ sung, sửa đổi và loại bỏ dữ liệu), tìm kiếm, khai thác và trao đổi dữ liệu. Trong CSDL các tệp được xử lý và khai thác đồng thời.

Để xây dựng các CSDL thoả mãn những nét đặc trưng trên, người ta đã thiết kế những mô hình để bố trí dữ liệu về mặt logic theo một quan điểm hệ thống nhất định. Có ba loại mô hình dữ liệu cơ bản là:

- Mô hình mạng.
- Mô hình phân cấp.
- Mô hình quan hệ.

Trong mô hình phân cấp các dữ liệu được tổ chức dưới dạng cấu trúc cây. Cấu trúc mạng cho phép thể hiện mối

quan hệ liên kết giữa dữ liệu này với một nhóm dữ liệu khác. Trong mô hình quan hệ người ta sử dụng nhiều quan hệ, trong đó mỗi quan hệ là một bảng hai chiều, mà mỗi dòng là một bản tin, mỗi cột là một mục trong bản tin.

Các hệ quản trị CSDL khác nhau được phát triển dựa trên các mô hình dữ liệu khác nhau. Mỗi loại hệ quản trị CSDL đều được xây dựng trên một mô hình dữ liệu xác định. Các hệ quản trị CSDL đều cung cấp một ngôn ngữ hình thức để trợ giúp người sử dụng và làm công cụ giao tiếp giữa người và máy. Đó là ngôn ngữ dữ liệu, nó cho phép người sử dụng xây dựng cấu trúc của CSDL, mô tả, cập nhật và tìm kiếm dữ liệu.

Trong các phần mềm quản lý CSDL văn bản ta thấy các hệ quản trị CSDL có cấu trúc tệp đảo có nhiều ưu điểm nổi bật và được sử dụng nhiều hơn cả, vì nó cho khả năng truy nhập dữ liệu một cách mềm dẻo và linh hoạt. CDS.ISIS chính là hệ quản trị CSDL có cấu trúc tệp đảo, đang được sử dụng rộng rãi trong các cơ quan thông tin và thư viện ở nước ta.

Từ quản lý tệp đến quản lý dữ liệu bằng các hệ quản trị cơ sở dữ liệu là một bước tiến quan trọng của tin học, góp phần mở rộng các ứng dụng của tin học trong công tác quản lý và điều hành.

1.4.2. Lưu trữ thông tin trên CD-ROM

Việc lưu trữ thông tin trên đĩa từ dựa trên cơ sở công nghệ điện từ. Từ đầu những năm 80 xuất hiện một loại phương tiện lưu trữ thông tin hoàn toàn khác, đó là đĩa quang. Đĩa quang được chế tạo dựa trên cơ sở công nghệ laser. Ở đây các dữ liệu số được ghi lại bằng cách đốt cháy một dãy các lỗ cực nhỏ, nhờ một chùm tia laser chiếu vào một

phim kim loại mỏng láng trên một đĩa nhựa có đường kính 12 cm. Bằng cách đó các thông tin trên đĩa từ được mã hoá và ghi lại trên một đĩa chủ, rồi từ đó được nhân bản lên. CD-ROM là thiết bị nhớ có khả năng ghi một lần nhưng có thể đọc nhiều lần (ROM - Read Only Memory). Nó thường được dùng để lưu trữ các thông tin với khối lượng lớn nhưng không có nhu cầu phải thay đổi hoặc cập nhật bởi người sử dụng.

Với các lợi thế nổi bật như kích thước nhỏ, dung lượng lớn, độ bền cao, đã được tiêu chuẩn hoá, có phần mềm khai thác nằm ngay trên đĩa, các đĩa quang CD-ROM ngày nay được sử dụng rộng rãi để lưu trữ, phân phối thông tin và triển khai các ứng dụng multimedia.

CD-ROM có dung lượng lớn, khoảng 600 MB thông tin, tương đương với khoảng 300.000 trang in (600 quyển sách, mỗi quyển 500 trang), do đó ứng dụng chính của CD-ROM là dùng để lưu trữ thông tin. Bất cứ dạng dữ liệu nào, chữ viết, âm thanh, hình ảnh tĩnh và động, sau khi số hoá đều có thể được lưu trữ trên CD-ROM dưới những dạng thức nhất định. Vì vậy sách báo, tạp chí, âm nhạc, phim ảnh, chương trình máy tính v.v... đều có thể được ghi vào đĩa CD-ROM. Mỗi đĩa CD-ROM chứa được khoảng 700 triệu ký tự hoặc hàng ngàn hình ảnh đồ hoạ hoặc 18 giờ âm thanh hoặc 74 phút phim với hình ảnh động.

Điều đặc biệt có ý nghĩa là mỗi CD-ROM đều có một phần mềm khai thác lưu trữ ngay trên đĩa (thường chiếm khoảng 1/2 MB), với giao diện sử dụng thích hợp, giúp ta dễ dàng tìm kiếm và truy nhập tới các thông tin ghi trên đĩa.

Trong những năm qua, lượng thông tin khổng lồ chứa trong các CSDL, các sách tham khảo, cẩm nang kỹ thuật, các

phần mềm máy tính, v.v... đã được phát hành dưới dạng CD-ROM.

Công nghệ CD-ROM cung cấp một phương tiện tốt để lưu trữ các bộ bách khoa toàn thư, các bộ từ điển lớn. Một bách khoa toàn thư hàng chục tập chỉ cần lưu trữ trên một CD-ROM. Các thông tin ở đây có thể phân phối dưới dạng có thể đọc lướt qua, được dò tìm theo các đề mục, các thông tin tham khảo chéo với hình ảnh - âm thanh hay kết nối với các thông tin khác chi tiết hơn.

2. TÌM TIN

2.1. Khái niệm chung về tìm tin

Tìm tin là thuật ngữ chung để chỉ việc tìm tài liệu hay nguồn của tài liệu, cũng như những thông tin về dữ liệu và sự kiện mà tài liệu đó cung cấp.

Các công việc ấy được tiến hành ở gần cuối dây chuyền tư liệu. Nó chuẩn bị cho việc phổ biến thông tin. Đó là mục đích cơ bản của hoạt động của các cơ quan thông tin. Có thể nói hầu hết các công việc của một đơn vị thông tin đều nhằm phục vụ cho việc tìm tin, đặt cơ sở cho các dịch vụ phổ biến thông tin.

Nếu nị tới dùng tin hỏi: "Chỗ anh (chị) có tài liệu này không?" thì chỉ cần chỉ cho họ chỗ tìm tài liệu là hệ thống mục lục. Nếu hỏi: "Ở đây có tài liệu về đề tài này không?" thì phải tìm trong thư mục các tài liệu nói về đề tài đó. Việc tìm tin như thế gọi là tìm tin hồi cố. Đó là công việc truyền thống của các cán bộ thư viện.

Nhưng trong sự phát triển hiện nay, xuất hiện hàng loạt các yêu cầu khác nhau: “Có gì mới về đề tài này không?” hoặc “Nội dung của tài liệu này là như thế nào?”, tức là người ta hỏi về những thông tin có trong tài liệu hay nói về một đề tài nào đó. Khi đó ta phải xác định tất cả các tài liệu có chứa các thông tin mới được bổ sung trong thời gian gần nhất.

Như vậy từ hoạt động thư viện đến hoạt động thông tin, nội dung phục vụ đã thay đổi khác trước và việc phục vụ thông tin bao gồm ba mức độ khác nhau:

- Cung cấp tài liệu cấp một.
- Cung cấp các thông tin chỉ dẫn.
- Cung cấp các thông tin số liệu, dữ kiện, các thông tin trích dẫn, các bản tóm tắt, phân tích, tổng hợp.

2.2. Phương thức tìm tin

Phương thức cơ bản của quá trình tìm tin là phương trình tìm hay còn gọi là biểu thức tìm. Cú pháp của phương trình tìm được xây dựng trên cơ sở các phép toán của logic mệnh đề, nó cho phép ta thực hiện ba loại quan hệ cơ bản sau giữa các từ chuẩn:

- Quan hệ tương giao, dùng phép hội, còn gọi là phép toán và. Nó cho phép nối hai từ chuẩn có trong bản chỉ mục của cùng một tài liệu.

- Quan hệ kết hợp, dùng phép tuyển, còn gọi là phép toán hoặc. Nó cho phép nối hai từ chuẩn mà ít nhất một trong hai từ đó có trong bản chỉ mục của tài liệu.

- Quan hệ loại trừ, dùng phép phủ định, còn gọi là phép toán không. Nó cho phép nối hai từ chuẩn mà từ thứ nhất có trong bản chỉ mục của tài liệu còn từ thứ hai thì không.

Các ví dụ tìm tin dưới đây đều thực hiện trên các mẫu tìm là các bảng đảo đã được thiết lập từ ma trận tư liệu, trình bày ở phần nguyên tắc lưu trữ thông tin.

Ví dụ:

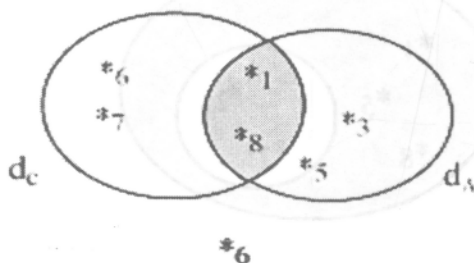
1- Phép VÀ:

"Sâu bệnh và Cây lúa" hay "C VÀ A".

Tất cả các tài liệu đánh chỉ số C và A là tương thích.

Như vậy các tài liệu thoả mãn mệnh đề trên chính là giao của hai mẫu tìm tương ứng:

$$d_C \cap d_A = \{ 1, 8 \}$$

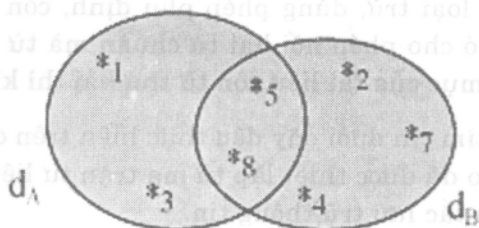


2- Phép HOẶC:

"Cây lúa hoặc Hoa màu" hay "A HOẶC B".

Tất cả các tài liệu đánh chỉ số A hoặc B là tương thích. Đó là hợp của hai mẫu tìm tương ứng:

$$d_A \cup d_B = \{ 1, 2, 3, 4, 5, 7, 8 \}$$

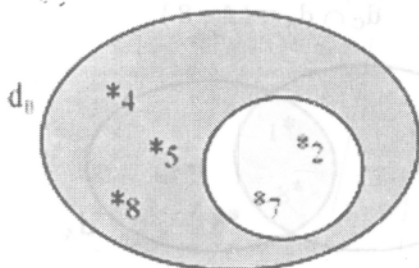


3- Phép KHÔNG:

"Hoa màu trừ Cây ngô" hay " $B \text{ KHÔNG } E$ ".

Tất cả các tài liệu đánh chỉ số B trừ các tài liệu đánh chỉ số E là tương thích. Đó là hiệu của hai mẫu tìm tương ứng:

$$d_B \setminus d_E = \{ 4, 5, 8 \}$$



Sử dụng phối hợp các phép toán trên đây ta có thể diễn đạt một yêu cầu thông tin bất kỳ, thể hiện bằng một mệnh đề phức hợp, gọi là biểu thức tìm hay phương trình tìm.

Ví dụ: Giả sử có một yêu cầu Q như sau: "Cần tìm tài liệu nói về sâu bệnh của cây lúa hoặc hoa màu, trừ cây ngô".

Xuất phát từ nội dung chủ yếu của câu hỏi, người cán bộ tra cứu trước hết phải thể hiện các thuật ngữ của câu hỏi thành các thuật ngữ của ngôn ngữ tư liệu: các từ chuẩn, các từ khoá. Đó là các dấu hiệu tìm kiếm của câu hỏi hay còn gọi là các điểm tiếp cận thông tin. Chẳng hạn với câu hỏi trên ta có thể rút ra các từ khoá tương thích với các dấu hiệu tìm kiếm như sau:

A = Cây lúa C = Sâu bệnh

B = Hoa màu E = Cây ngô

Sau đó dùng các phép toán logic liên kết các từ khoá trên thành phương trình tìm:

Q = Sâu bệnh VÀ (Cây lúa HOẶC Hoa màu) KHÔNG Cây ngô
hay gọn hơn:

Q = C VÀ (A HOẶC B) KHÔNG E

Nếu dùng ký hiệu của phép toán logic thì có thể viết:

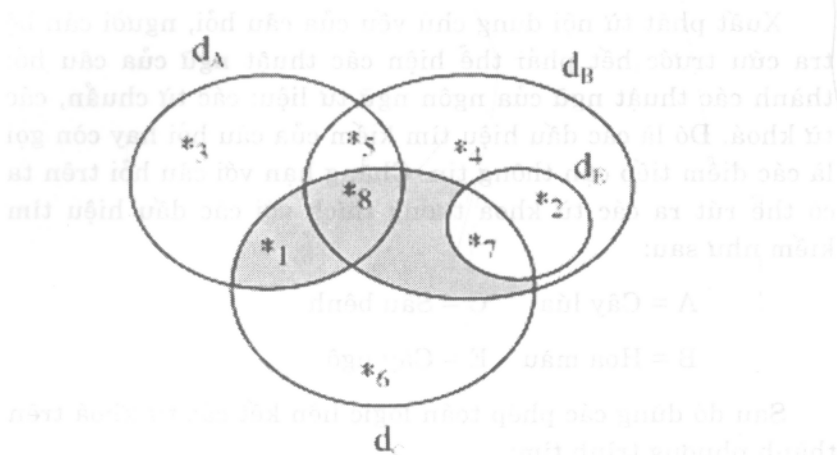
$Q = C \wedge (A \vee B) \setminus E$

Với ví dụ đang xét thì câu trả lời của câu hỏi trên là tập hợp:

$d_Q = d_C \cap (d_A \cup d_B) \setminus d_E = \{ 1, 8 \}$

Vậy các tài liệu số 1 và tài liệu số 8 thoả mãn yêu cầu tìm tin đặt ra.

Ta có thể minh hoạ kết quả tìm tin trên bằng sơ đồ Venn sau đây:



Ngoài các phép toán cơ bản trên, trong ngôn ngữ hỏi còn có một số toán tử khác, cho phép ta tìm tin thuận lợi hơn hoặc để hoàn chỉnh thêm biểu thức tìm. Trong số đó ta có thể kể:

- Toán tử chặt cứng, cho phép ta tìm một từ bằng một nhóm chữ, không kể đến phân đầu và phân đuôi của từ đó.

Ví dụ: Trong ngôn ngữ hỏi của chương trình CDS.ISIS toán tử chặt cứng ký hiệu là \$. Khi đó trong biểu thức tìm nếu viết:

Nguyễn \$

thì câu trả lời sẽ cho tài liệu của tất cả các tác giả họ Nguyễn như Nguyễn Đình Thi, Nguyễn Văn Bổng, Nguyễn Thi, v.v...

- Toán tử so sánh, cho phép tìm theo các dấu hiệu về lượng.

Ví dụ: Trong một CSDL về sách, sử dụng chương trình CDS.ISIS, có trường Năm xuất bản với format hiện hình là V3^c. Để tìm các sách xuất bản sau năm 1976 ta viết biểu thức tìm:

`?val(V03^c)>1976`

2.3. Quá trình tìm tin

Tìm tin là một quá trình bao gồm nhiều bước kế tiếp nhau, theo một trình tự hợp lý. Quá trình tìm tin bao gồm các bước sau:

1) Xác định câu hỏi:

Xuất phát từ nhu cầu thông tin, người dùng tin đề ra các câu hỏi. Người cán bộ tra cứu cần xác định nội dung của câu hỏi. Khi cần phải tiếp xúc với người dùng tin để làm rõ tất cả các khía cạnh của câu hỏi.

2) Thể hiện câu hỏi bằng ngôn ngữ tư liệu:

Ở đây các yếu tố của câu hỏi được thể hiện thành các yếu tố của ngôn ngữ tư liệu và được viết thành biểu thức tìm, tương thích với các phương tiện tra cứu và các nguồn thông tin cấp hai sẵn có của đơn vị thông tin.

Ở giai đoạn này cần sử dụng các loại từ điển khác nhau, các từ điển từ chuẩn, các khung đề mục.

3) Vạch ra chiến lược tìm:

Người cán bộ tra cứu vạch ra con đường để tiếp cận tài liệu. Muốn vậy điều chủ yếu là phải xác định được các nguồn tài liệu cấp hai thích hợp nhất: các thư mục, các bộ phiếu mục lục, các bảng tra cứu, các cơ sở dữ liệu thư mục...

4) Thực hiện tìm:

Thực hiện tìm tin thông qua các công cụ tìm kiếm khác nhau, dựa trên cơ sở các biểu thức tìm. Các công cụ tìm kiếm có thể là các mục lục, các thư mục, các ấn phẩm tra cứu khác nhau hay máy tính điện tử trong các hệ thống thông tin tự động hoá.

5) Phân tích các kết quả của chiến lược tìm:

Kiểm tra đánh giá kết quả tìm được, từ đó đánh giá tính đúng đắn của chiến lược tìm. Nếu thấy cần thiết thì quay lại giai đoạn ba để sửa đổi và xác định lại chiến lược tìm.

6) Chuyển kết quả tìm được cho người dùng tin:

Cung cấp cho người dùng tin tài liệu và câu trả lời dưới dạng mà họ yêu cầu.

7) Đánh giá tính phù hợp thông tin nhận được:

Sau khi nhận được kết quả, người dùng tin đánh giá chất lượng và giá trị của thông tin nhận được. Nếu kết quả không phù hợp thì quay lại giai đoạn một, xác định lại câu hỏi và đề ra câu hỏi mới. Ở đây đơn vị thông tin ghi nhận những yêu cầu và đánh giá của người dùng tin để điều chỉnh cho thích hợp.

Việc tìm tin có thể do người dùng tin trực tiếp thực hiện, cũng có thể do cán bộ thông tin thực hiện với vai trò là người trợ giúp trung gian.

Trong trường hợp tìm trực tiếp, người dùng tin phải thực hiện hầu hết các công việc của quá trình tìm tin. Đơn vị thông tin chỉ hỗ trợ chỉ dẫn các nguồn và cung cấp tài liệu.

Trong trường hợp tìm tin do cán bộ tra cứu đảm nhận thì giai đoạn một - xác định câu hỏi - là quan trọng nhất vì bản chất của việc tìm tin là tìm câu trả lời cho một câu hỏi. Do đó người cán bộ tra cứu phải hợp tác với người dùng tin để xác định chính xác yêu cầu nội dung của câu hỏi đặt ra.

Quá trình tìm tin có thể biểu diễn bằng sơ đồ ở hình 17.



Hình 17. Lược đồ của quá trình tìm tin

2.4. Một số dạng tìm tin cơ bản

Các công cụ tìm tin có thể rất khác nhau từ các bộ phiếu thủ công đến các tệp dữ liệu trong các hệ thống thông tin tự động hoá cho phép tìm tin bằng hội thoại. Mặc dù cách tổ chức, nội dung hình thức của các bộ phiếu, bao hàm kỹ thuật tìm tin, ít nhiều phức tạp khác nhau, nhưng phương thức tìm tin về cơ bản vẫn là các biểu thức tìm. Dưới đây ta giới thiệu một số dạng tìm tin chính.

2.4.1. Tìm tin trên các bộ phiếu thủ công

Việc tìm tin trên các mục lục hay bộ phiếu thư mục được thực hiện xuất phát từ các dấu hiệu cần tìm. Tùy theo dấu hiệu cần tìm, người ta sử dụng bộ phiếu tác giả, bộ phiếu chủ đề hay bộ phiếu địa lý.

Việc tìm tin trên các mục lục thông báo hay các bản chỉ dẫn cũng thực hiện tương tự. Việc tìm tin càng nhanh nếu các bản chỉ dẫn chứa nhiều thông tin về các tài liệu cần tra cứu. Tuy nhiên cách tìm này thường tốn thời gian và nhầm lẫn.

Trong cả hai trường hợp trên người ta tìm các từ tiêu để tương ứng với các thuật ngữ có ý nghĩa trong câu hỏi, xem xét tất cả các phiếu tra cứu liên quan, lựa chọn chúng bằng cách loại bỏ các phiếu không tương thích. Người ta thường bắt đầu bằng những tiêu đề cho chỉ dẫn nhiều nhất đến tài liệu, sau đó kiểm tra các dấu hiệu phụ (ngày tháng, ngôn ngữ,...).

Khi tìm tin trên bộ phiếu uniterme, người ta lấy ra các phiếu mang các từ chuẩn đồng nhất với các thuật ngữ của câu hỏi. Tài liệu có số chung ghi trong tất cả các phiếu là tài liệu thích hợp. Từ số này người ta chuyển qua phiếu thư mục tương ứng. Ở đây ghi tất cả các chỉ dẫn thư mục cho phép tìm

tài liệu mà người ta yêu cầu. Như vậy việc tìm phải thực hiện qua ba bước, nhưng các thủ tục so sánh rất đơn giản và nhanh chóng.

Ví dụ: Muốn tìm tài liệu nói về “sâu bệnh của cây lúa ở Việt Nam”, ta lấy ra ba phiếu uniterme tương ứng với các từ khoá SÂU BỆNH, CÂY LÚA, VIỆT NAM. Nếu có tài liệu nào nói về cả ba chủ đề trên thì số ký hiệu kho của nó phải được ghi trên cả ba phiếu uniterme. Ta chỉ cần đối chiếu ba phiếu đó rồi chọn ra số ký hiệu chung.

SÂU BỆNH									
1	2	3	4	5	6	7	8	9	0
71			114				88		
CÂY LÚA									
1	2	3	4	5	6	7	8	9	0
71				25			88		
VIỆT NAM									
1	2	3	4	5	6	7	8	9	0
		43					88		

Hình 18. Ví dụ về tìm tin trên bộ phiếu uniterme

Ở hình 18 ta thấy:

Tài liệu 71 nói về sâu bệnh và cây lúa.

Tài liệu 88 nói về sâu bệnh và cây lúa ở Việt Nam.

Tìm tin trên phiếu lỗ soi người ta cũng làm tương tự. Người ta lấy ra tất cả các phiếu ứng với các từ chuẩn trong biểu thức tìm rồi đặt chồng lên nhau trước một nguồn sáng. Toạ độ của lỗ cho ánh sáng đi qua sẽ là số hiệu của tài liệu được đánh chỉ số bằng tất cả các từ chuẩn có trong biểu thức tìm.

Việc tìm tin trên phiếu lỗ mép được thực hiện bằng cách chọn theo chữ của mã ứng với từ chuẩn ghi trong biểu thức tìm. Người ta thường chia bộ phiếu theo từng lô, mỗi lô khoảng từ 100 đến 150 phiếu và nên bắt đầu bằng từ chuẩn tiêu biểu nhất. Thời gian lựa chọn khá kéo dài, nhưng các phiếu không cần sắp xếp và trên các phiếu đã ghi đầy đủ các thông tin thư mục cần tìm.

2.4.2. Tìm tin trên máy tính điện tử

Việc tìm tin trên máy tính thể hiện những nét khác nhau tùy theo người ta sử dụng bộ phiếu kế tiếp hay bộ phiếu đảo.

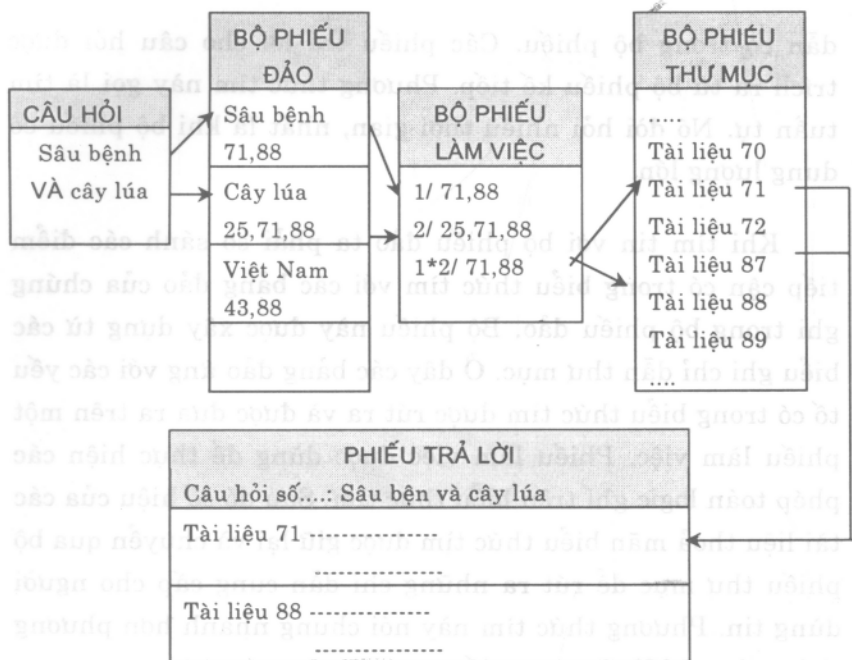
Việc tìm tin với bộ phiếu kế tiếp được thực hiện bằng cách so sánh biểu thức tìm với các điểm tiếp cận có trong các bản đánh chỉ số ghi trong các chỉ dẫn thư mục của từng tài liệu. Các chỉ dẫn này được sắp xếp theo thứ tự nhập của tài liệu. Trong quá trình tìm ta phải đọc lần lượt tất cả các chỉ

dẫn có trong bộ phiếu. Các phiếu trả lời cho câu hỏi được trích ra từ bộ phiếu kế tiếp. Phương thức tìm này gọi là tìm tuần tự. Nó đòi hỏi nhiều thời gian, nhất là khi bộ phiếu có dung lượng lớn.

Khi tìm tin với bộ phiếu đảo ta phải so sánh các điểm tiếp cận có trong biểu thức tìm với các bảng đảo của chúng ghi trong bộ phiếu đảo. Bộ phiếu này được xây dựng từ các biểu ghi chỉ dẫn thư mục. Ở đây các bảng đảo ứng với các yếu tố có trong biểu thức tìm được rút ra và được đưa ra trên một phiếu làm việc. Phiếu làm việc được dùng để thực hiện các phép toán logic ghi trên biểu thức tìm. Sau đó số hiệu của các tài liệu thoả mãn biểu thức tìm được giữ lại và chuyển qua bộ phiếu thư mục để rút ra những chỉ dẫn cung cấp cho người dùng tin. Phương thức tìm này nói chung nhanh hơn phương thức trên, gọi là tìm trực tiếp.

Việc tìm tin bằng hội thoại trên máy tính được thực hiện theo dòng nhờ một thiết bị đầu cuối có màn hình. Nét đặc trưng ở đây là người dùng tin phải có khả năng hội thoại với máy. Nó cho phép người ta có thể điều chỉnh chiến lược tìm trong quá trình tìm tin. Tuy nhiên việc tìm tin càng kéo dài, thời gian sử dụng máy tính càng nhiều thì chi phí càng lớn.

Dưới đây là lược đồ mô tả quá trình tìm tin trên máy tính điện tử với bộ phiếu đảo:



Hình 19. Tìm tin theo bộ phiếu đảo trên MTĐT

2.4.3. Tìm tin trên CD-ROM

Vì mỗi đĩa CD-ROM đều có một phần mềm khai thác ở ngay trên đĩa, nên khi một CSDL được lưu trữ trên CD-ROM thì việc tìm tin được tiến hành như các phương thức tìm đối với các CSDL được lưu trữ trong máy tính điện tử.

Ví dụ, CD-ROM cơ sở dữ liệu PASCAL cung cấp hai phương thức tìm tin:

- Tìm tin trình độ cao (Expert search)
- Tìm tin có trợ giúp (Guided search)

Ngôn ngữ tìm được xây dựng trên cơ sở ba phép toán logic cơ bản AND, OR, NOT. Bạn có thể tìm theo nhiều dấu hiệu tìm kiếm khác nhau, như: nhan đề, tác giả, từ khoá, yếu tố xuất bản, chỉ số phân loại,... PASCAL sử dụng một khung phân loại riêng, rất thuận lợi cho việc sử dụng. Bạn có thể tìm theo ba ngôn ngữ Anh, Pháp và Tây Ban Nha.

Đặc biệt ngày nay, nhờ công nghệ multimedia, các dữ liệu dạng văn bản, âm thanh, hình ảnh tĩnh và động, được lưu trữ trên CD-ROM trong một môi trường tương tác thuận tiện cho người sử dụng. Với giao diện thân thiện, và với công nghệ kết nối siêu văn bản (hypertext link) việc tìm kiếm thông tin có thể chuyển từ điểm truy nhập này sang điểm truy nhập khác một cách thuận tiện và nhanh chóng.

Ví dụ, với CD-ROM bách khoa thư ENCARTA, trên giao diện tìm kiếm bạn gõ vào từ Vietnam, rồi thực hiện tìm. Trên màn hình sẽ hiện ra một bài viết về Việt Nam, với bản đồ, quốc kỳ và quốc ca. Từ đây bạn có thể nghe được bản nhạc Tiến quân ca. Ngay trên trang đầu của bài viết bạn sẽ thấy các từ khoá in màu đậm như Hanoi, Ho Chi Minh City, Red River, Vietnam War, ... Đó chính là những điểm kết nối thông tin siêu văn bản. Từ đây bạn có thể tìm ngay được bài viết về những chủ đề cần tìm bằng cách kích chuột vào thuật ngữ tương ứng. Ví dụ, bạn mở bài viết về Hà Nội chẳng hạn. Từ đây bạn sẽ gặp từ khoá Ho Chi Minh in đậm. Kích chuột vào từ này, bạn sẽ có ngay một bài viết về Chủ tịch Hồ Chí Minh, kèm theo một tấm hình rất đẹp. Quá trình tìm cứ thế tiếp tục, bạn có thể tìm đến không biết bao nhiêu tri thức bách khoa có trong ENCARTA.

2.4.4. Tìm tin on-line

Tìm tin on-line, còn gọi là tìm tin trực tuyến là tìm tin thông qua một mạng máy tính do một cơ quan làm dịch vụ thông tin (máy chủ) đảm nhiệm. Các cơ quan này là đầu mối cung cấp thông tin của các CSDL được kết nối với mạng.

Ví dụ:

DIALOG là cơ quan dịch vụ thông tin của Mỹ, đảm nhiệm cung cấp thông tin cho trên 400 CSDL của nhiều nước trên thế giới. Cơ sở dữ liệu PASCAL của INIST cũng do DIALOG đảm nhận việc cung cấp thông tin.

ESA/IRS là cơ quan dịch vụ thông tin của European Space Agency, đảm nhận cung cấp thông tin cho khoảng 100 CSDL, chủ yếu về các đề tài kỹ thuật.

ORBIT đảm nhận cung cấp thông tin của gần 100 CSDL, chủ yếu về các phát minh sáng chế khoa học và kỹ thuật.

Để sử dụng được dịch vụ tìm tin on-line trên mạng, cần phải có các thiết bị sau:

- Một máy tính cá nhân hay một thiết bị đầu cuối (terminal).
- Một modem, để từ máy tính cá nhân có thể truyền thông tin tới máy chủ và từ đó có thể tiếp cận CSDL mà ta quan tâm..
- Một máy in.

Muốn trở thành khách hàng của hệ thống tìm tin on-line, người dùng tin phải ký hợp đồng với cơ quan dịch vụ thông tin để nhận được một mật khẩu (Password) và phải được

hướng dẫn để biết sử dụng ngôn ngữ hỏi riêng của cơ quan dịch vụ thông tin mà họ đăng ký.

Nói chung các cơ quan này đều cung cấp các tài liệu trợ giúp như: Tập từ vựng các từ khoá, khung phân loại của từng CSDL, sách hướng dẫn hỏi đáp trên các CSDL, sách hướng dẫn sử dụng các thiết bị. Ngoài ra còn có một bộ phận tư vấn giúp người dùng tin có thể sử dụng thuận tiện các dịch vụ của họ.

Tuy nhiên cần lưu ý rằng, người dùng tin cuối cùng thường ít hỏi trực tiếp các CSDL, vì hai lý do. Trước hết các ngôn ngữ hỏi thường khá phức tạp và rất khác nhau, gây trở ngại cho người dùng tin trực tiếp. Mặt khác số các CSDL trên một lĩnh vực thường rất nhiều, người dùng tin khó có thể xác định được CSDL nào có thể cung cấp cho họ câu trả lời tốt nhất.

Vì vậy người dùng tin thường không giao dịch trực tiếp với CSDL mà thông qua đại diện trung gian của mình là các trung tâm thông tin tư liệu. Trong trường hợp này, việc tìm tin được thực hiện gián tiếp bởi các chuyên gia tìm. Đó là các cán bộ tư liệu rất thành thạo trong lĩnh vực tìm tin trên các hệ thống thông tin trực tuyến.

Phương thức tìm tin on-line trên các CSDL mới xuất hiện từ đầu những năm 70 nhưng đã phát triển cực kỳ nhanh chóng và trở thành một phương thức tìm tin thuận tiện, có hiệu quả đáp ứng được nhu cầu thông tin đa dạng và ngày càng phát triển của người dùng tin. Năm 1972 mới có 40 CSDL phục vụ tìm tin on-line thì đến năm 1987 con số này đã lên tới 2800. Sự gia tăng này đánh dấu sự khởi đầu của một

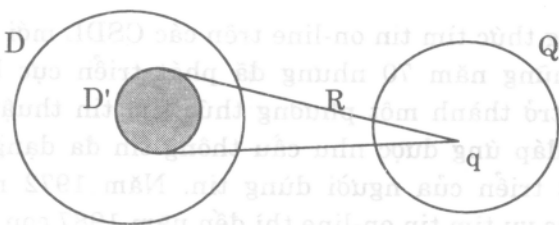
xu hướng lớn là tăng cường sử dụng máy tính điện tử trong việc lưu trữ và tìm kiếm thông tin.

3. MÔ HÌNH HOÁ QUÁ TRÌNH TÌM TIN

Như ta đã biết hoạt động của một đơn vị thông tin được vận hành trên một tập hợp các công đoạn liên tiếp có liên hệ chặt chẽ với nhau gọi là dây chuyền tư liệu. Mở đầu là khâu lựa chọn tài liệu ban đầu để lập nên một kho tài liệu. Tiếp theo là công đoạn xử lý hình thức và nội dung của tài liệu: mô tả thư mục, phân loại, đánh chỉ số, làm tóm tắt. Sản phẩm của công đoạn này được ghi vào bộ nhớ dưới dạng các bộ phiếu khác nhau. Một công đoạn quan trọng của dây chuyền là tra cứu tìm tin và đưa ra câu trả lời cho người dùng tin.

Quá trình tìm tin trình bày ở phần trên có thể mô tả như sau:

Giả sử D là tập hợp các tài liệu, Q là tập hợp các câu hỏi. Với mỗi câu hỏi $q \in Q$, ta gọi R là quan hệ nhờ đó ta ứng tập con $D' \subset D$ với câu hỏi $q \in Q$, sao cho D' là câu trả lời của q . Như vậy việc tìm tin thực chất là một quan hệ R từ D đến Q ứng tập con D' của tập các tài liệu D với câu hỏi q của người dùng tin.



Hình 20. Quan hệ R ứng tập con các tài liệu D' với câu hỏi q

Hiển nhiên phương pháp hiệu quả nhất để tìm tài liệu có chứa những thông tin cần thiết là đọc qua mọi tài liệu $x \in D$ rồi so sánh với yêu cầu q . Nhưng trong thực tế phương pháp đó không thể thực hiện được vì số tài liệu thường là rất lớn. Vì thế người ta không chọn trên tập D các tài liệu mà chọn trên mô hình của tài liệu. Đó là tập F các ảnh y của $x \in D$ cho bởi ánh xạ:

$$I : D \rightarrow F$$

sao cho:

$$\forall y \in F, \exists x \in D: y = I(x)$$

$$\forall x, x' \in D \text{ và } x \neq x' \Rightarrow I(x) \neq I(x')$$

ánh xạ I ở đây thực chất là các khâu xử lý hình thức và nội dung tài liệu. Đó là một song ánh từ D đến F . Ở đây F là tập hợp các chỉ dẫn, là mô hình thu gọn của tài liệu tương ứng, chúng được hệ thống hoá và lưu trữ trên các vật mang tin.

Ánh xạ I ứng mỗi $x \in D$ với $y = I(x)$ được xác định như sau:

$$y = \{ a_x, b_x, m_x, r_x \}$$

trong đó:

a_x là địa chỉ của x

b_x là tập hợp các đặc trưng hình thức của x

m_x là tập hợp các đặc trưng nội dung của x

r_x là tóm tắt của x

Tập hợp các chỉ dẫn trên đây của x còn gọi là mẫu tìm của x . Đó là một thể hiện ngắn gọn các đặc trưng cũng như nội dung cơ bản của tài liệu.

Để thực hiện ánh xạ I người ta phải dùng ngôn ngữ tư liệu. Đó là một ngôn ngữ nhân tạo bảo đảm hai yêu cầu cơ bản: quan hệ ngữ nghĩa là một - một, quan hệ cú pháp là nhất quán.

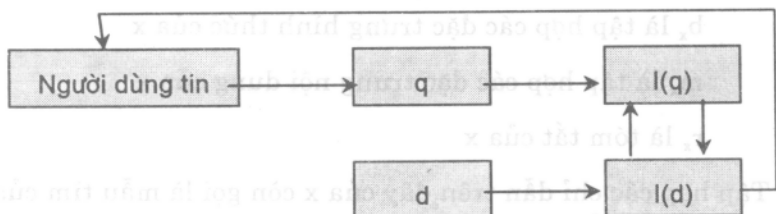
Yêu cầu q của người dùng tin cũng được biểu diễn bằng ngôn ngữ tư liệu dưới dạng một biểu thức tìm. Như vậy biểu thức tìm cũng có thể coi là ảnh q qua ánh xạ I:

$$I: q \rightarrow I(q)$$

Khi đó nhờ ngôn ngữ tư liệu người ta có thể so sánh biểu thức tìm với mẫu tìm để xem tài liệu có phù hợp hay không.

Như vậy thủ tục tìm tin có thể giới hạn trong việc so sánh đơn giản giữa mẫu tìm với biểu thức tìm. Nếu mẫu tìm trùng với biểu thức tìm ở mức độ cần thiết và đầy đủ thì ta nói tài liệu đó đáp ứng được với yêu cầu thông tin. Rõ ràng cách so sánh đó chỉ có thể thực hiện được khi mẫu tìm và biểu thức tìm được diễn tả bằng một ngôn ngữ, đó là ngôn ngữ tư liệu. Việc so sánh giữa mẫu tìm với biểu thức tìm được thực hiện trên các phương tiện tra cứu như: các bộ phiếu, các thư mục, các ấn phẩm tra cứu hay trên máy tính điện tử.

Ta có thể minh hoạ quá trình tìm tin mô tả trên đây bằng mô hình sau:



Hình 21. Mô hình quá trình tìm tin

4. ĐÁNH GIÁ HIỆU QUẢ TÌM TÌN

Hiệu quả của việc tìm tin phụ thuộc vào chất lượng của bộ máy tra cứu tin hay còn gọi là kho tra cứu tin.

Kho tra cứu tin là tập hợp các mẫu tìm. Đó là tập hợp các thông tin cấp hai. Kho tra cứu tin phải được tổ chức thành các mảng để công việc tìm tiến hành được thuận lợi. Từ biểu thức tìm ta có thể xác định xem thông tin cần tìm ở mảng nào. Do đó tiết kiệm được công sức, tránh phải “quét” tất cả kho tra cứu. Tổ chức mảng cũng dựa trên cơ sở của ngôn ngữ tư liệu. Đó cũng là một cách đánh chỉ số. Nhờ tổ chức mảng mà xác suất tìm được nâng lên.

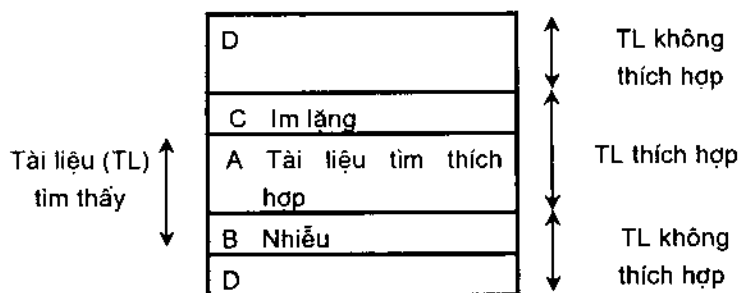
Tìm tin có hai loại: tìm tin dưới dạng hỏi đáp và tìm tài liệu.

Tiêu chuẩn để đánh giá hiệu quả của dịch vụ tìm tin hỏi đáp là đánh giá theo khả năng trả lời câu hỏi, bao gồm các chỉ số sau đây:

- Tỷ lệ câu hỏi có thể trả lời.
- Tỷ lệ câu hỏi tìm nhưng không trả lời được.
- Tỷ lệ câu trả lời đúng, chính xác.

Đối với dịch vụ tìm tin tư liệu có nhiều phương pháp đánh giá. Dưới đây là phương pháp dựa vào tính thích hợp của tài liệu đưa ra, tức là tài liệu đưa ra có thích hợp với yêu cầu của người dùng tin hay không.

Các tham số đánh giá tìm tin tư liệu được trình bày trong hình sau:



Hình 22. Các tham số đánh giá hiệu quả tìm tin

Trong bảng trên C là số tài liệu thích hợp nhưng không tìm thấy, còn gọi là im lặng, B là số tài liệu đưa ra không thích hợp còn gọi là nhiều.

Hiệu quả tìm tin được đánh giá qua các chỉ số sau đây:

Hệ số đầy đủ: Hệ số đầy đủ là tỉ số giữa tài liệu thích hợp tìm ra và tài liệu thích hợp có trong kho:

$$\frac{A}{A + C}$$

Hệ số chính xác: Hệ số chính xác là tỉ số giữa tài liệu thích hợp tìm ra và số tài liệu đã tìm:

$$\frac{A}{A + B}$$

Ta nhận thấy A là tài liệu tìm ra thích hợp càng lớn thì hiệu quả tìm tin càng lớn. Hệ số chính xác $\frac{A}{A + B}$ càng nhỏ khi mô tả nội dung tài liệu không sâu.

Chương 6

CÁC ĐƠN VỊ THÔNG TIN VÀ DỊCH VỤ THÔNG TIN

1. CÁC LOẠI HÌNH ĐƠN VỊ THÔNG TIN VÀ MẠNG LƯỚI THÔNG TIN

Các thiết chế xã hội đã sớm quan tâm đến các tổ chức có chức năng bảo quản và sắp xếp tài liệu, nhằm giúp con người có thể tiếp cận được với những tri thức mà loài người tích lũy được theo thời gian. Các thư viện và cơ quan lưu trữ đã có từ thời cổ đại (thư viện Alexandria của Ai Cập là thư viện lớn nhất thời cổ đại, được xây dựng vào thế kỷ thứ ba trước Công nguyên). Lúc đầu chúng chỉ dành riêng cho tầng lớp trên và những người có học. Về sau cùng với sự tiến bộ của xã hội và giáo dục, chúng được mở rộng cho công chúng. Các hoạt động của các tổ chức này cho đến thế kỷ XX vẫn tập trung chủ yếu vào việc xử lý các tài liệu cấp một.

Trong thời đại ngày nay, số lượng người dùng tin và nhu cầu thông tin gia tăng nhanh chóng, kỹ thuật xử lý thông tin phát triển. Vì vậy cần phải có những cơ quan chuyên môn trong hoạt động thông tin, đi sâu vào những chức năng khác của dây chuyền tư liệu như mô tả nội dung tài liệu, trích dẫn và xử lý dữ liệu, tìm và phổ biến thông tin... nhằm phục vụ

cho những nhóm người dùng tin khác nhau. Chúng có những tên gọi khác nhau như: trung tâm thông tin tư liệu, trung tâm phân tích và tổng hợp tin, ngân hàng dữ liệu v.v...

Các đơn vị thông tin rất đa dạng do sự khác nhau về tên gọi và hoạt động của chúng. Người ta có thể phân loại các đơn vị thông tin theo phạm vi và lĩnh vực hoạt động, theo loại hình tài liệu, theo đối tượng phục vụ, nhưng chủ yếu là phân loại theo chức năng hoạt động của chúng. Từ đó người ta có thể phân biệt ba ngành hoạt động chính:

- Bảo quản và cung cấp tài liệu cấp một (thư viện, lưu trữ).

- Xử lý nội dung tài liệu, phổ biến và thông báo các thông tin và nguồn tin (các trung tâm thông tin tư liệu).

- Trả lời các câu hỏi bằng cách khai thác các thông tin sẵn có, bao gồm cả việc đánh giá và phổ biến chúng (các trung tâm phân tích và tổng hợp tin, NHDL).

Ba nhánh hoạt động này tương ứng với các loại sản phẩm và dịch vụ thông tin mà mức độ xử lý chúng thể hiện ngày càng sâu sắc hơn.

1.1. Các thư viện và cơ quan lưu trữ

1.1.1. Thư viện

Thư viện đơn vị thông tin mà nhiệm vụ chủ yếu của nó là thu thập, bảo quản và cung cấp cho người dùng tin những tài liệu cấp một.

Mặc dù từ *library* (thư viện) có xuất xứ từ thuật ngữ La tinh *liber*, có nghĩa là “sách”, nhưng các tài liệu mà thư viện

bảo quản ngày nay không chỉ có sách mà còn nhiều loại tài liệu khác như: báo, tạp chí, đĩa hát, băng ghi âm, băng ghi hình, đĩa quang CD, phim ảnh v.v...

Các đơn vị này phục vụ cho công chúng rộng rãi thuộc nhiều đối tượng khác nhau. Họ đến để tìm tài liệu mà họ cần và việc mô tả tài liệu ở đây mới chỉ dừng ở xác định tên tài liệu và những chỉ dẫn sơ lược về chủ đề của tài liệu.

Các thư viện ngày nay thường phân chia nhiệm vụ của nó thành hai lĩnh vực, một liên quan đến các hoạt động nội tại của thư viện (dịch vụ kỹ thuật), lĩnh vực kia quan hệ trực tiếp tới người đọc (phục vụ người đọc). Vì vậy hệ thống tổ chức của thư viện thường bao gồm:

- Một ban giám đốc và bộ phận hành chính.

- Các dịch vụ kỹ thuật (bổ sung, biên mục, bảo quản, kho...).

- Các dịch vụ phục vụ người đọc (cho mượn, tra cứu).

Cùng với sự phát triển của nhu cầu thông tin, thư viện cũng dần dần mở rộng hoạt động của mình như: in, sao chụp tài liệu, các dịch vụ chỉ dẫn và định hướng, tìm tin hồi cố...

Thư viện thường được phân loại theo một trong hai cách:

Theo quyền sở hữu và mục đích có:

- Thư viện quốc gia.

- Thư viện tỉnh.

- Thư viện đại chúng (còn gọi là thư viện công cộng).

- Thư viện khoa học.

- Thư viện đại học.
- Thư viện trường học phổ thông.

Theo nội dung chủ đề có:

- Thư viện tổng hợp.
- Thư viện chuyên ngành (khoa học, kỹ thuật, y học, luật, âm nhạc,...).

Mỗi thư viện có một đặc thù riêng tùy theo quy mô tổ chức và đối tượng phục vụ của nó. Hầu hết các thư viện này phục vụ người đọc bằng hình thức cho mượn tài liệu đọc tại chỗ hay cho mượn về trong một thời gian nhất định. Trong thế kỷ XX đã phát triển các thư viện chuyên ngành, có nhiệm vụ cất giữ các ấn phẩm định kỳ về các thông tin trong khoa học, kỹ thuật, y học, thương mại, luật pháp,... Các thư viện này đảm nhiệm thêm các dịch vụ thông tin như: tra cứu, tìm tin, thông tin thư mục và chúng thường được bảo trợ bởi các ngành công nghiệp và các hội nghề nghiệp.

Thư viện quốc gia là một tổ chức của nhà nước có nhiệm vụ thu thập, xử lý, bảo quản và phổ biến tất cả các tài liệu xuất bản trong nước và một phần những tài liệu có giá trị của các nước khác. Hầu hết các thư viện quốc gia đều được nhận (miễn phí) các tài liệu sách, báo, tạp chí xuất bản trong nước theo luật lưu chiểu. Thư viện quốc gia còn có nhiệm vụ biên tập và xuất bản thư mục quốc gia và giữ vai trò chỉ đạo nghiệp vụ đối với các thư viện địa phương.

Các thư viện quốc gia ra đời sớm nhất ở các nước châu Âu như Pháp, Bỉ, Đức vào thế kỷ XV - XVI. Các thư viện quốc gia tiêu biểu hiện nay là Thư viện Quốc hội Mỹ, Thư viện Nhà nước Nga ở Maxcova (trước đây là Thư viện Lênin), Thư viện

Quốc gia Pháp ở Paris, Thư viện Anh ở Luân Đôn. Năm 1995 Thư viện Quốc hội Mỹ bảo quản khoảng 90 triệu đơn vị tài liệu, trong đó có 13 triệu sách thuộc 125 ngôn ngữ khác nhau. Kho sách của Thư viện Nhà nước Nga, với hệ thống phòng đọc gồm 2500 chỗ ngồi, hiện có trên 32 triệu bản thuộc 247 thứ tiếng, trong đó có 91 ngôn ngữ của các dân tộc Liên Xô (cũ). Thư viện Quốc gia Pháp thành lập từ năm 1537, hiện bảo quản trên 15 triệu tài liệu. Kho sách của Thư viện Anh hiện lưu giữ hơn 20 triệu bản, trong đó có một bộ sưu tập toàn diện nhất thế giới về tài liệu sáng chế phát minh. Tiền thân của Thư viện Quốc gia Việt Nam là Thư viện Trung ương Hà Nội thuộc Nhà Lưu trữ và Thư viện Đông Dương, sau này quen gọi là Thư viện Trung ương, được thành lập theo Nghị định ngày 29-11-1917 của Toàn quyền Pháp. Thư viện Quốc gia Việt Nam hiện quản lý vốn tài liệu phong phú, bao gồm khoảng 1.013.000 bản sách, trên 7000 tên báo và tạp chí, trong đó có những sách, báo tạp chí lâu đời nhất ở Việt Nam (theo số liệu năm 1999). Ngoài ra còn khoảng 3000 tài liệu tra cứu bao gồm các từ điển, các bách khoa thư, các sổ tay v.v...

Thư viện đại chúng có nhiệm vụ đáp ứng nhu cầu văn hoá, giải trí và thực tiễn của đông đảo người đọc thuộc đủ các tầng lớp và mọi lứa tuổi. Ngoài các sách báo thông thường, thư viện đại chúng còn có các tài liệu tra cứu, các tài liệu chứa các thông tin về các dịch vụ xã hội, các sách giải trí, đĩa nhạc, băng hình v.v... Ở các thư viện đại chúng thường tổ chức các bộ phận phục vụ theo các đối tượng độc giả như: bộ phận phục vụ thiếu nhi, bộ phận phục vụ người cao tuổi...

Thư viện đại học ưu tiên phục vụ cho hoạt động nghiên cứu, giảng dạy và học tập của các giáo sư, giảng viên và sinh

viên. Vốn tài liệu của nó bao gồm những tài liệu chuyên sâu và thường khá đầy đủ về các ngành khoa học và kỹ thuật được giảng dạy trong nhà trường. Thư viện trường đại học Oxford ở Cambridge (thành lập từ thế kỷ XIV), thư viện trường đại học Harvard ở bang Massachusetts (thành lập năm 1638) còn lớn hơn nhiều thư viện quốc gia cả về tầm cỡ và chất lượng. Ở nhiều trường đại học lớn, ngoài thư viện trung tâm, các khoa còn có thư viện chuyên ngành riêng phục vụ cho công tác đào tạo của mình. Ở thư viện các trường đại học nhiều người có thể tìm và sử dụng những cuốn sách như nhau trong cùng một thời gian như nhau. Vì vậy một tài liệu thường phải nhân thành nhiều bản mới đáp ứng được yêu cầu của người sử dụng.

Các thư viện quốc gia, thư viện đại chúng, thư viện đại học ít nhiều đáp ứng yêu cầu của công chúng khá rộng rãi. Bên cạnh đó còn có các thư viện chuyên ngành có nhiệm vụ thu thập và cung cấp các tài liệu chuyên ngành cho một lớp các đối tượng đặc biệt, đó là các nhà chuyên môn. Thư viện chuyên ngành thường được tổ chức gắn liền với các cơ quan ngành, các viện nghiên cứu thuộc các lĩnh vực, các bệnh viện, bảo tàng, v.v... Các thư viện chuyên ngành thường hoạt động theo một chương trình kế hoạch chặt chẽ và vốn tài liệu thường được quan tâm xây dựng rất công phu.

Từ đầu những năm 70, việc ứng dụng công nghệ thông tin, đặc biệt là máy tính điện tử vào thư viện đã đem lại những biến đổi sâu sắc trong hoạt động của các thư viện truyền thống. Ở các thư viện này người ta thường xây dựng các thư mục công cộng truy nhập trực tuyến, gọi tắt là OPAC (On-line Public Access Catalog). Đó là các CSDL thư mục được khai thác trên mạng, giúp người sử dụng có thể truy

nhập các thông tin thư mục một cách trực tiếp mà không cần sự hỗ trợ trung gian của nhân viên thư viện.

1.1.2. Cơ quan lưu trữ

Tài liệu lưu trữ là một loại tài liệu chuyên dạng, bao gồm:

- Tài liệu của các cơ quan quyền lực nhà nước ở trung ương và địa phương, trong đó có: các tài liệu quản lý hành chính, tài liệu khoa học kỹ thuật, tài liệu phim ảnh, ghi âm.
- Tài liệu của chính quyền qua các thời đại, các thời kỳ.
- Tài liệu của các nhân vật hoạt động nổi tiếng như các nhà hoạt động chính trị, các nhà văn, nhà thơ, nhà khoa học, v.v...

Có thể nói tài liệu lưu trữ là những ký lục sinh động của toàn xã hội, trải qua các thời kỳ lịch sử. Đó là di sản của dân tộc, có giá trị đặc biệt phục vụ nghiên cứu khoa học, lịch sử và công tác thực tiễn. Việc bảo quản và sử dụng các tài liệu này phải tuân theo những quy tắc chặt chẽ, tùy theo mức độ quan trọng và phạm vi phổ biến của tài liệu.

Các cơ quan lưu trữ là tổ chức có trách nhiệm tiếp nhận, bảo quản và cung cấp cho người dùng tin có trình độ cao những tài liệu lưu trữ về quản lý hành chính của quốc gia, của các địa phương hoặc của các ngành, các đơn vị cơ quan xí nghiệp. Hoạt động của các cơ quan này tuân theo những quy tắc quản lý chặt chẽ.

Cơ quan lưu trữ cũng có nhiều loại:

- Cơ quan lưu trữ quốc gia.
- Cơ quan lưu trữ của các địa phương.

- Cơ quan lưu trữ của các ngành như nội vụ, quốc phòng, ngoại giao.

- Cơ quan lưu trữ của các cơ quan, xí nghiệp.

Cơ quan lưu trữ quốc gia bảo quản các tài liệu về quản lý hành chính của các triều đại phong kiến, của chính quyền qua các thời kỳ, của các cá nhân, gia đình, dòng họ nổi tiếng, tài liệu về các sự kiện lịch sử có ý nghĩa chính trị xã hội v.v...

Lưu trữ của các ngành như: công nghiệp, nông nghiệp, y tế,... thường chứa đựng nhiều thông tin khoa học và kỹ thuật có giá trị.

Lưu trữ của các cơ quan có trách nhiệm về tất cả các tài liệu của cơ quan đó.

1.2. Thư viện điện tử

Các thư viện truyền thống từ lâu đã là một phần của xã hội, nó mang thông tin và tri thức đến mọi người. Tuy nhiên ngày nay khi các máy tính cá nhân và mạng Internet đã đã làm thay đổi cách thức giao lưu của con người, khi mà từ máy tính cá nhân người ta có thể với tới các thông tin lưu trữ ở khắp nơi trên thế giới, thì người ta thấy rằng có thể xây dựng những thư viện có khả năng cung cấp thông tin tốt hơn so với các thư viện truyền thống. Đó là lý do cơ bản của sự ra đời một loại hình thư viện mới: thư viện điện tử (Electronic Library), còn gọi là thư viện số (Digital Library).

Vậy thư viện điện tử là gì ?

Thư viện điện tử có thể coi là một kho thông tin số hoá, được cấu trúc sao cho dễ dàng truy cập thông qua các mạng máy tính hay các mạng viễn thông quốc tế.

Có thể nói thư viện điện tử là một hệ thống thông tin tự động hoá mà ở đó người ta có thể thu thập, xử lý, lưu trữ, tìm kiếm và phổ biến các tài liệu dưới dạng số hoá thông qua các phương tiện của công nghệ thông tin và truyền thông.

Lợi ích của thư viện điện tử là:

- Với một máy tính cá nhân kết nối mạng, người sử dụng có thể truy cập thông tin từ xa. Người sử dụng không cần đến thư viện vẫn tìm được thông tin. Người ta nói: “Thư viện điện tử mang thông tin đến tận bàn làm việc của bạn”.

- Với thư viện điện tử, sức mạnh của máy tính được dùng để tìm kiếm thông tin. Trong hầu hết các trường hợp, tìm tin bằng máy tính bao giờ cũng nhanh hơn, tiện lợi hơn phương pháp thủ công, và đặc biệt có thể chuyển dễ dàng từ nguồn thông tin này sang nguồn thông tin khác.

- Với thư viện điện tử, thông tin luôn sẵn sàng, bởi vì của thư viện điện tử không bao giờ đóng.

- Với thư viện điện tử các nguồn thông tin có thể chia sẻ, người sử dụng không chỉ truy cập, sao chép các nguồn thông tin nội tại của thư viện mà có thể với tới nhiều nguồn thông tin từ bên ngoài nhờ các dịch vụ thông tin liên kết.

- Thông tin trong thư viện điện tử dễ bổ sung, cập nhật nên thư viện điện tử luôn bảo đảm có những thông tin mới nhất.

- Với thư viện điện tử, các dạng thông tin mới như thông tin điện tử, thông tin đa phương tiện trở thành hữu dụng, vì thư viện điện tử có khả năng tích hợp thông tin số hoá.

Các thành phần cơ bản của thư viện điện tử

Thư viện điện tử được xây dựng trên nền tảng của một thư viện truyền thống nên nó phải tuân thủ những tiêu chuẩn nghiệp vụ căn bản của thư viện, tuy nhiên có sự điều chỉnh một số giá trị cũ sao cho phù hợp việc ứng dụng công nghệ mới.

Hạt nhân của thư viện điện tử là nguồn tài liệu số hoá. Trong đó có một bộ phận là tài liệu hiện hữu của thư viện được số hoá (có chọn lọc), nhưng chủ yếu là các bộ sưu tập tài liệu số mới được xây dựng hoặc sưu tầm.

Thư viện điện tử hoạt động trên nền giao diện Web của môi trường mạng Internet, nên nguồn tài liệu của thư viện điện tử thường được trình bày định dạng bằng ngôn ngữ đánh dấu siêu văn bản HTML. Ở đó các tài liệu được thiết kế đặc biệt để có thể kết nối với các tập tin khác thông qua các điểm kết nối siêu văn bản (hypertext link points). Nền tảng công nghệ của thư viện điện tử chính là Internet và Worl Wide Web.

Thư viện điện tử được quản lý bởi một phần mềm tích hợp quản trị thư viện. Phần mềm này bao gồm nhiều phân hệ chức năng và tuân thủ các chuẩn quốc tế về nghiệp vụ thư viện cũng như các chuẩn về công nghệ thông tin và truyền thông. hiện đại, như: khổ mẫu biên mục đọc máy MARC, khổ mẫu trao đổi dữ liệu thư mục ISO.2709, giao thức tìm kiếm liên thư viện Z.39.50, giao thức truyền thông Internet TCP/IP, làm việc trên môi trường Web và hỗ trợ chuẩn định dạng HTML, XML.

Cấu trúc của thư viện điện tử

Các thư viện điện tử đều được thiết kế trên một trang Web, mà ở đó ngoài các mục giới thiệu chung như giới thiệu về cơ quan, về thư viện, về các công cụ trợ giúp thì phần chủ yếu là nội dung, tức là giới thiệu tới các nguồn tài nguyên thông tin.

Tài nguyên thông tin này thường bao gồm:

- Các thông tin chuyên đề
- Các CSDL.
- Các nguồn tài nguyên thông tin liên kết trên mạng.

Các thông tin chuyên đề được thể hiện dưới dạng một danh mục các chủ đề. Từ mỗi chủ đề này ta có thể vào những mục, tiểu mục với các thông tin là những bài viết đề cập đến những vấn đề mà người dùng tin quan tâm. Các xuất bản phẩm điện tử (tạp chí, bản tin, kỷ yếu,...) thường được tổ chức sắp xếp theo kiểu này. Những nội dung thông tin này có mối liên kết nhiều chiều với nhau theo kiểu kết nối siêu văn bản, tạo thuận lợi cho người sử dụng có thể tiếp cận dễ dàng.

Các CSDL bao gồm các CSDL nội sinh của thư viện và cả những CSDL nhập từ bên ngoài. Các CSDL này được sắp xếp trong một danh mục theo vần chữ cái. Để tìm tin trong CSDL nào đó, người sử dụng chỉ việc kích chuột vào tên CSDL tương ứng. Một giao diện tìm kiếm sẽ hiện ra và người sử dụng có thể thực hiện các thao tác tìm trên đó. Người dùng tin có thể tiếp cận các CSDL này theo nhiều mức độ: từ thư mục tới toàn văn.

Các nguồn tài nguyên thông tin liên kết trên mạng là các nguồn thông tin có được trong sự hợp tác với các cơ quan

thông tin, các thư viện điện tử khác. Chúng được tích hợp vào hệ thống và được khai thác trong một thể thống nhất. Để truy cập tới các thông tin này cần có sự hỗ trợ của các công cụ tìm kiếm (search engine).

Tóm lại, cấu trúc của thư viện điện tử là cấu trúc của một trang Web có liên kết đến các nguồn thông tin số hoá, trong đó quan trọng nhất là các CSDL toàn văn. Các nguồn tài liệu số hoá này được sưu tầm và được tổ chức theo một cơ chế thống nhất sao cho có thể dễ dàng truy cập, sao chép trên các mạng thông tin viễn thông.

Cần nhấn mạnh rằng thư viện điện tử hoạt động trên giao diện Web và trong môi trường Internet, nhưng một Website không thể là một thư viện điện tử vì những thông tin trong đó, tuy khá phong phú nhưng thiếu đặc điểm *được sưu tầm* và *được tổ chức* như những thông tin trong thư viện điện tử.

Phần cốt lõi của thư viện điện tử là kho tài liệu số hoá. Vì vậy xây dựng kho tài liệu số hoá được coi là công việc quan trọng hàng đầu trong xây dựng thư viện điện tử. Công việc này đòi hỏi phải có đầu tư lớn về công sức và tài chính.

Hiện nay nhiều Thư viện của các trường đại học ở Mỹ hoạt động theo mô hình Thư viện công Internet (Internet Public Library – IPL). Thực chất đó là các thư viện điện tử hoạt động trên môi trường Internet với các đặc trưng nêu trên.

1.3. Các trung tâm thông tin tư liệu

Các trung tâm thông tin tư liệu là những đơn vị thông tin mà nhiệm vụ chủ yếu là xử lý nội dung tài liệu và xác định với độ chính xác cao những thông tin hữu ích cho người dùng

tin, giúp họ tìm đến những tài liệu cấp một tương ứng, thông báo cho họ những chỉ dẫn này và trả lời câu hỏi của họ.

Về mặt lý thuyết các đơn vị này chỉ xây dựng những bộ phiếu với những mô tả thư mục và mô tả nội dung tài liệu, không làm chức năng bảo quản tài liệu cấp một. Nhưng trong thực tế, nhiều khi chúng cũng đảm nhiệm dịch vụ của một thư viện, đặc biệt đối với những đơn vị thông tin chuyên ngành.

Tuy nhiên cần nhấn mạnh rằng các đơn vị thông tin không phải là cơ quan bảo quản tài liệu, mà là cơ quan khai thác tài liệu. Các trung tâm thông tin tư liệu xem xét và đánh giá tất cả các nguồn tài liệu cấp một, cấp hai, bao gồm cả những tài liệu mà thư viện không quan tâm như các tài liệu không công bố, các catalog công nghiệp, các tài liệu thương mại và nhiều loại tài liệu khác. Các sản phẩm của các đơn vị này có thể dùng để trả lời các câu hỏi cụ thể của các nhà khoa học, các nhà nghiên cứu và sản xuất hoặc để phổ biến dưới dạng các ấn phẩm thông tin như các chỉ dẫn, tóm tắt, các báo cáo về các thông tin chuyên ngành.

Về mặt tổ chức, ngoài bộ phận lãnh đạo và hành chính, một trung tâm thông tin tư liệu thường phải có ba phân ban sau:

- Phân ban thư viện.
- Phân ban biên tập tư liệu, thông tin.
- Phân ban xuất bản và phát hành.

Các dịch vụ mà một trung tâm thông tin tư liệu phải bảo đảm là:

- Biên tập các thư mục khoa học chuyên ngành, các bản mô tả nội dung tài liệu, các tổng luận khoa học, sản xuất các CSDL.

- Tổ chức các dịch vụ phổ biến thông tin: cung cấp tài liệu, trả lời các câu hỏi, phổ biến thông tin có chọn lọc.

- Biên dịch các tài liệu khoa học.

- Xuất bản các ấn phẩm thông tin.

Các trung tâm thông tin tư liệu có nhiều loại, tùy theo lĩnh vực chuyên môn, đối tượng phục vụ và quy mô của chúng. Người ta phân biệt:

- Trung tâm thông tin tư liệu quốc gia.

- Trung tâm thông tin tư liệu của các địa phương.

- Trung tâm thông tin tư liệu của các đơn vị cơ quan, xí nghiệp, trường đại học, viện nghiên cứu,...

- Trung tâm thông tin tư liệu chuyên ngành (nông nghiệp, y tế, dầu khí, mỏ địa chất v.v...).

Các trung tâm thông tin tư liệu quốc gia là những tổ chức của nhà nước, bao quát tất cả các lĩnh vực chuyên môn mà nó đảm nhận và phục vụ tất cả các đối tượng người dùng tin. Nó có trách nhiệm thu thập và bảo quản tất cả các tài liệu và các nguồn thông tin sản xuất ở trong nước và cả những tài liệu và thông tin của các nước khác trên tất cả các lĩnh vực của tri thức. Khi hoạt động đạt đến một quy mô nào đó, các trung tâm này thường quản lý bên trong nó một mạng lưới các đơn vị thông tin chuyên ngành hoặc chuyên dạng.

Ở Việt Nam, Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia được thành lập ngày 24/9/1990 trên cơ sở

hợp nhất hai đơn vị là Thư viện Khoa học và Kỹ thuật Trung ương (ra đời năm 1960) và Viện Thông tin Khoa học và Kỹ thuật Trung ương (1972), với hai chức năng chủ yếu:

- Là trung tâm đầu ngành thông tin và thư viện của toàn mạng quốc gia về lĩnh vực khoa học, công nghệ và môi trường.

- Thực hiện chức năng quản lý nhà nước về thông tin và tư liệu khoa học và công nghệ.

Trung tâm có nhiệm vụ:

- Thu thập và phát triển các nguồn thông tin và tư liệu về khoa học công nghệ và môi trường trong và ngoài nước. Xây dựng vốn thông tin tư liệu đáp ứng các nhu cầu phát triển kinh tế - xã hội, khoa học và công nghệ của đất nước.

- Xử lý, phân tích - tổng hợp các nguồn thông tin tư liệu trong và ngoài nước nhằm tạo ra các sản phẩm và dịch vụ thông tin đáp ứng nhu cầu của người dùng tin và bạn đọc.

- Tổ chức, cung cấp thông tin tư liệu cho các yêu cầu thẩm định, đánh giá các đề tài nghiên cứu, các dự án đầu tư phát triển kinh tế xã hội.

- Thực hiện công tác nghiên cứu - triển khai trong lĩnh vực thông tin - tư liệu. Áp dụng công nghệ thông tin mới.

- Hướng dẫn và chỉ đạo thống nhất nghiệp vụ cho các ngành, các địa phương về thông tin - thư viện khoa học và công nghệ.

- Tổ chức đào tạo và nâng cao trình độ chuyên môn và nghiệp vụ cho cán bộ về công tác thông tin và tư liệu.

Trung tâm có một vốn tài liệu khá phong phú, bao gồm:

- Kho tài liệu của Thư viện Khoa học và Kỹ thuật Trung ương với trên 350 000 cuốn sách, trong đó 97% là sách khoa học - kỹ thuật nước ngoài, 5000 tên tạp chí, trong đó có 1800 tạp chí đang đặt mua thường xuyên.

- Hơn 10 cơ sở dữ liệu thư mục về khoa học, công nghệ và môi trường với khoảng 400 000 biểu ghi (do trung tâm xây dựng, sử dụng phần mềm CDS/ISIS).

- Các CSDL của nước ngoài ghi trên CD-ROM và trên đĩa mềm.

- Các ấn phẩm thông tin do trung tâm biên tập và xuất bản.

Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia hoạt động theo mô hình thông tin - thư viện. Trung tâm đã chú trọng tạo lập một hệ thống sản phẩm và dịch vụ khá phong phú, ngày càng đáp ứng được nhu cầu của người dùng tin, góp phần thúc đẩy sự phát triển của đất nước. Với mong muốn bắt kịp công nghệ tiên tiến của thế giới, từ cuối những năm 80 trung tâm đã từng bước tin học hoá các hoạt động của mình và mới đây Mạng thông tin Khoa học và Công nghệ Việt Nam - VISTA được xây dựng và đưa vào hoạt động đã đánh dấu một bước phát triển mới về chất trong hoạt động thông tin tư liệu của trung tâm.

Gần đây Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia đã đổi tên thành Trung tâm Thông tin Khoa học và Công nghệ Quốc gia với tên giao dịch quốc tế là NACESTI (National Centre for Scientific and Technological Information).

1.4. Ngân hàng dữ liệu

Ngân hàng dữ liệu, còn gọi là ngân hàng thông tin, là một hệ thống thông tin tự động hoá, khai thác trực tiếp, được cấu trúc thành CSDL bao quát một lĩnh vực nào đó của tri thức.

NHDL có các thành phần sau:

- Một CSDL hoặc một tập hợp các CSDL khác nhau, trong đó có thể gồm các CSDL tự xây dựng lấy hoặc bổ sung từ bên ngoài.

- Tổ hợp các phương tiện xây dựng, bảo trì và khai thác CSDL, bao gồm: máy tính điện tử và các thiết bị ngoại vi; các thiết bị truyền và nhận dữ liệu từ xa; các thiết bị sao chụp; các phần mềm quản trị các CSDL; các ngôn ngữ giao tiếp giữa người và máy, các thủ tục và phương pháp; các nhân viên quản trị, vận hành và khai thác CSDL.

Trong NHDL các CSDL vừa là bộ phận cấu thành, vừa là sản phẩm của NHDL. Với tư cách là một sản phẩm của NHDL, các CSDL được đưa vào các mạng lưới thông tin để sử dụng chung theo chế độ cung cấp thông tin trực tuyến hoặc trao đổi giữa các cơ quan thông tin với nhau.

Các CSDL chính là các kho thông tin đã được tin học hoá. Muốn khai thác và sử dụng chúng thì phải đưa chúng vào các NHDL. Các kho thông tin này thường phải rất phong phú mới phục vụ được các loại khách hàng với nhiều yêu cầu đa dạng khác nhau. Thông tin ở đây trở thành hàng hoá và được khai thác trực tiếp, không cần phải trở lại các tài liệu cấp một và các nguồn thông tin khác.

Có thể nói khi các CSDL được đưa vào mạng lưới thông tin khai thác trực tuyến và từ xa, thông qua các cơ quan dịch

vụ thông tin thì chúng trở thành NHDL. Như vậy các NHDL do các cơ quan thông tin xây dựng và được các cơ quan dịch vụ thông tin khai thác.

Năm 1980 trên thế giới mới có 400 NHDL, do 221 cơ quan sản xuất và được khai thác bởi 59 cơ quan dịch vụ thông tin. Đến năm 1986 con số đó đã là 3169 NHDL, do 1914 cơ quan sản xuất và được khai thác bởi 486 cơ quan dịch vụ thông tin.

Việc xây dựng và vận hành các NHDL liên quan chặt chẽ đến các hoạt động của các cơ quan thông tin, từ khâu thu thập, xử lý tiền máy đến các khâu phục vụ thông tin cho người dùng tin. Vì vậy, việc xây dựng và khai thác các CSDL và NHDL được xem như quá trình xây dựng và khai thác các hệ thống thông tin tự động hoá.

1.5. Mạng thông tin

Cùng với sự phát triển của công nghệ thông tin, nhu cầu giao lưu, hợp tác và trao đổi thông tin dẫn đến việc hình thành các mạng thông tin.

Mạng thông tin là một tập hợp các cơ quan thông tin được tổ chức liên hệ với nhau, ở đó người ta có thể thực hiện việc trao đổi thông tin dưới những hình thức khác nhau một cách đều đặn và có tổ chức.

Để đạt được mục tiêu là chia sẻ trách nhiệm và tăng thêm nguồn tin, các thành viên của mạng phải thiết lập những quy ước, xác định phương thức chung về yêu cầu và kỹ thuật sử dụng.

Để duy trì và phát triển một mạng thông tin hoạt động có hiệu quả, các thành viên của mạng phải có ý thức chung về quyền lợi cũng như về trách nhiệm, cùng vượt qua khó khăn góp phần xây dựng mạng. Ngoài ra các phương tiện kỹ thuật của mạng phải phù hợp với mọi đơn vị thành viên.

Hoạt động của một mạng thông tin có thể đi từ hợp tác đơn giản từng phần đến việc mở rộng ra tất cả các lĩnh vực hoạt động của công tác tư liệu, với sự tham gia đầy đủ của các thành viên trong một hệ thống thông tin duy nhất.

Có nhiều loại mạng thông tin. Người ta phân biệt có các loại sau:

- Mạng thông tin trên một lãnh thổ, ở mức độ một tỉnh, một vùng hay một quốc gia, cho phép thoả mãn các yêu cầu thông tin của các đối tượng người dùng tin trên lãnh thổ đó.

- Mạng thông tin chuyên trong một lĩnh vực chuyên môn hay một ngành hoạt động. Ở đó các đơn vị thành viên kết hợp để hỗ trợ lẫn nhau, điều hoà trách nhiệm và sản phẩm.

- Mạng thông tin chuyên theo chức năng hoạt động của công tác tư liệu như: bổ sung, biên mục, cho mượn, đánh chỉ số, làm tóm tắt, tìm kiếm, hỏi đáp...

- Mạng thông tin chuyên phục vụ cho một đối tượng người dùng tin riêng biệt. Ví dụ: các xí nghiệp nhỏ, các nhà sản xuất cà phê...

Những đặc trưng trên đây có thể kết hợp trong cùng một hệ thống và các hệ thống khác nhau có thể kết hợp hoạt động với nhau.

Sự phát triển của các phương tiện và kỹ thuật truyền thông hiện đại đã tạo điều kiện cho việc phát triển các mạng thông tin tự động hoá truyền tin từ xa mà người ta gọi là mạng thông tin viễn thông. Các thông tin ở đây được lưu trữ trong máy tính điện tử và được tiếp cận nhờ các mạng truyền thông. Người dùng tin có thể dùng các thiết bị đầu cuối khác nhau như minitel, các máy vi tính hoặc máy télétip để liên lạc với máy tính trung tâm của NHDL. Các tín hiệu nhị phân của máy tính và của đầu cuối phải được biến đổi thành các tín hiệu có thể truyền đi trên các tuyến truyền thông theo hai chiều nhờ một thiết bị gọi là modem (modulator - demodulator).

Có thể nói mạng thông tin máy tính là cơ sở hạ tầng của truyền thông, là yếu tố thiết yếu cấu thành hệ thống thông tin trực tuyến (on-line).

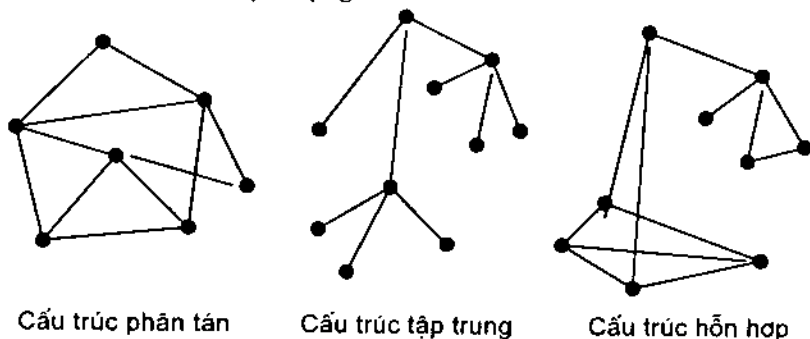
Cấu trúc của mạng rất quan trọng, nó ảnh hưởng đến cách truyền và giao lưu thông tin. Có thể có các dạng cấu trúc sau đây:

- Cấu trúc phân tán: Tất cả các đơn vị đều giao lưu với nhau. Các kênh liên hệ nhiều và ngắn, nhưng khó quản lý. Đó là trường hợp mạng lưới cho mượn của các thư viện.

- Cấu trúc tập trung: Các đơn vị giao lưu với nhau thông qua một trung tâm, hình thành một mô hình phân cấp. Chẳng hạn mạng lưới thư viện quốc gia và thư viện các địa phương.

- Cấu trúc hỗn hợp: là sự phối hợp cả hai cấu trúc trên.

Các đồ thị cho trong Hình 23 dưới đây cho ta hình ảnh về cấu trúc của các loại mạng lưới.



Hình 23. Cấu trúc mạng

Ví dụ:

Ở Pháp có mạng liên lạc viễn thông transpac, phát triển vào bậc nhất trong thế giới phương Tây. Nhờ mạng lưới transpac, người dùng tin ở Pháp có thể với tới thông tin ở bất kỳ NHDL nào, ở bất cứ nơi nào trong nước.

Các nước trong khối cộng đồng chung châu Âu có mạng thông tin trực tuyến Euronet, thành lập năm 1975, mở rộng phục vụ công chúng vào đầu năm 1979. Mạng lưới này tích nhập các dịch vụ thông tin trực tuyến của khoảng 100 NHDL của các nước thành viên trong một mạng lưới chung, cùng hợp tác và chia sẻ trách nhiệm, cho phép người dùng tin ở bất cứ nơi nào trong cộng đồng, thông qua màn hình có thể tiếp cận trực tiếp tới các thông tin khoa học, kỹ thuật, kinh tế xã hội và nhiều lĩnh vực khác.

Mạng thông tin khoa học và công nghệ Việt Nam-VISTA (Vietnam Information Network for Science and Technology Advance) là mạng thông tin máy tính diện rộng, do Trung tâm Thông tin Tư liệu Khoa học và Công nghệ Quốc gia xây dựng, quản lý và phát triển nhằm phục vụ cho các đối tượng làm công tác khoa học và công nghệ, đồng thời phổ biến kiến thức cho quảng đại quần chúng nhân dân, nâng cao dân trí trong lĩnh vực khoa học và công nghệ. Các nguồn tin của VISTA khá phong phú, bao gồm các CSDL thư mục về các tài liệu khoa học và công nghệ, các CSDL dữ kiện, các xuất bản phẩm điện tử. Mạng thông tin VISTA được thiết kế trên nền hệ điều hành Windows NT 4.0, theo mô hình vùng đơn (Single Domain), tức là các máy chủ trong mạng hoạt động như một hệ thống duy nhất. Giao thức mạng của VISTA là giao thức chuẩn của Internet TCP/IP. Hệ quản trị dữ liệu được xây dựng theo kiến trúc khách/chủ.

Các dịch vụ của mạng VISTA bao gồm:

- Tra cứu thông tin thư mục trong các CSDL khoa học và công nghệ trong và ngoài nước của mạng VISTA.

- Tra cứu các tin tức, các tổng quan, các kết quả nghiên cứu, các ấn phẩm khoa học và công nghệ trong và ngoài nước và sử dụng các dịch vụ thư viện của trung tâm.

- Qua mạng VISTA người sử dụng có thể nhận được tài liệu gốc có ở trung tâm hoặc trên thế giới.

- Sử dụng các dịch vụ Internet qua mạng VISTA như World Wide Web, FTP, E-mail và Telnet.

Hiện nay có hàng trăm cơ quan tổ chức kết nối và khai thác các dịch vụ của VISTA.

1.6. Mạng thông tin toàn cầu Internet

Ngày nay Internet như một ma lực đang thu hút từng cư dân trên trái đất lên siêu xa lộ thông tin. Các cá nhân có thể tìm thấy trên Internet đủ các loại thông tin, từ giá chứng khoán đến thực đơn ở một khách sạn. Sinh viên tìm thấy trên Internet lời giải các bài toán khó. Các nhà khoa học tìm thấy các kết quả nghiên cứu khoa học của các đồng nghiệp ở bất cứ nơi nào trên trái đất nếu họ đang nối mạng Internet. Nhiều công ty có thể sử dụng Internet để tìm kiếm các thông tin về đối thủ cạnh tranh, về khách hàng và tìm kiếm cơ hội kinh doanh...

Theo con số thống kê mới nhất, hiện nay có khoảng 200 triệu người đang sử dụng Internet. Tín tức thế giới, thư tín cá nhân, những yêu cầu về khoa học, giáo dục, văn hoá, nghệ thuật và sản xuất, kinh doanh luân chuyển qua mạng thông tin toàn cầu Internet, đã và đang kết nối mọi người ở khắp các miền trên thế giới lại với nhau.

Vậy Internet là gì?

Internet là hệ thống liên mạng máy tính có quy mô toàn cầu, ở đó các máy tính có thể liên lạc trực tiếp với nhau.

Internet sử dụng một hệ thống địa chỉ chung và một họ giao thức truyền thông gọi là TCP/IP (Transmission Control Protocol/Internet Protocol). TCP/IP là bộ giao thức chuẩn dùng để thực hiện việc truyền dữ liệu giữa hai máy tính kết nối với nhau trên mạng Internet. Có thể nói TCP/IP là “ngôn ngữ” để các máy tính nói chuyện với nhau. TCP/IP được đề xuất và phát triển bởi người sáng lập ra ARPAnet là Robert Kahn và nhà nghiên cứu Vinton Cerf (Đại học Stanford) vào năm 1974.

Internet được cấu trúc để có thể kết nối thông qua nhiều phương tiện truyền thông khác nhau, bao gồm đường điện thoại, tuyến cáp quang, hệ thống vệ tinh hoặc qua các mạng cục bộ (LAN).

Để truy nhập Internet, mỗi máy tính phải đăng ký một địa chỉ xác định và phải thông qua một nhà cung cấp dịch vụ Internet (Internet Service Provider - ISP). Khi hoà mạng Internet, mỗi máy tính gọi là một host. Khi mạng LAN sử dụng công nghệ Internet thì trở thành Intranet.

Tiền thân của Internet là ARPAnet. Đó một mạng máy tính có độ tin cậy cao, phục vụ cho các hoạt động của quốc phòng và an ninh, được xây dựng vào năm 1969, theo một dự án của Cục các dự án nghiên cứu tiên tiến (Advanced Research Projects Agency - ARPA) thuộc Bộ quốc phòng Mỹ. Giao thức truyền thông dùng trong ARPANET lúc đó được đặt tên là NCP (Network Control Protocol) và đến năm 1983 thì được thay thế bởi TCP/IP.

ARPAnet nhanh chóng mở rộng thêm các nút mới và trở thành một mạng quốc gia. Sau đó Bộ quốc phòng Mỹ quyết định tách phần "quân sự" trong ARPAnet ra thành mạng Milnet. Năm 1987, trên cơ sở của mạng ARPAnet, dưới sự bảo trợ của Ủy ban khoa học quốc gia Mỹ (National Science Foundation - NSF), mạng NSFnet ra đời với tốc độ đường truyền nhanh hơn, phục vụ cho nghiên cứu khoa học.

Sự xuất hiện của mạng xương sống NSFnet và các mạng vùng đã thúc đẩy mạnh mẽ sự tăng trưởng của Internet. Một xa lộ thông tin mới hình thành với sự tham gia của nhiều trường đại học, viện nghiên cứu, các tổ chức chính phủ và các giới kinh doanh. Về mặt địa lý, Internet nhanh chóng vượt ra

khỏi phạm vi nước Mỹ và trở thành một liên mạng thông tin toàn cầu. Cho đến đầu năm 1995 đã có hơn 50 000 mạng máy tính và 5 triệu máy tính được kết nối với nhau thông qua mạng Internet.

Cùng với sự phát triển của Internet, các phần mềm Internet cũng lần lượt ra đời, giúp người tham gia có thể sử dụng các dịch vụ Internet. Hai phần mềm Internet nổi tiếng nhất hiện nay là Netscape Navigator của hãng Netscape và Internet Explorer của hãng Microsoft.

Một vấn đề đặt ra là: ai quản lý Internet? Thực tế là không có cơ quan quản lý tối cao cho toàn bộ mạng Internet trải khắp hành tinh như hiện nay. Tuy nhiên có một tổ chức có vai trò điều phối các hoạt động của Internet là Hiệp hội Internet (Internet Society) viết tắt là ISOC, có trụ sở đặt tại bang Virginia. Đây là một tổ chức phi lợi nhuận, tập hợp các thành viên chuyên nghiệp nhằm hỗ trợ sự phát triển kỹ thuật và thúc đẩy việc sử dụng Internet trên toàn thế giới. Bên dưới ISOC có các trung tâm thông tin mạng ở các khu vực viết tắt là NIC (Network Information Center) đảm nhận cung cấp các dịch vụ thông tin, dịch vụ cơ sở dữ liệu và thư mục, dịch vụ đăng ký. NIC của khu vực châu Á - Thái Bình Dương gọi tắt là APNIC có trụ sở tại Tokyo, Nhật Bản. Hiện tại APNIC chịu trách nhiệm điều hành và phân phối các địa chỉ cho các khách hàng ở Việt Nam.

Bạn có thể làm gì với Internet ?

Bạn có thể dùng Internet liên lạc với bạn bè, gia đình, đồng nghiệp hoặc tham gia các hội thảo có tính toàn cầu, tham gia các trò chơi, hướng dẫn nghiên cứu khoa học, sao chép các chương trình máy tính, tra cứu thông tin trong các

cơ sở dữ liệu, theo dõi các chương trình thể thao, âm nhạc mới nhất...

Tất cả các khả năng này có thể được tiến hành thông qua các nguồn tin và các dịch vụ có sẵn trên mạng Internet, với sự hỗ trợ của nhiều công cụ khác nhau như Telnet, Gopher, FTP, WAIS, World Wide Web,... Đó là các chương trình mà hầu hết hoạt động theo phương thức khách/chủ (client/server), trong đó khách hàng đề xuất yêu cầu và người phục vụ đáp ứng yêu cầu đó.

Các khả năng của Internet là vô hạn. Chúng ta kể ra ở đây một số khả năng cơ bản.

1- Trao đổi thư điện tử và tham gia diễn đàn thông tin

Bạn có thể trao đổi thư từ với bất cứ ai trên mạng Internet thông qua dịch vụ thư điện tử (electronic-mail) hay gọi một cách phổ biến là e-mail. E-mail là dịch vụ đơn giản nhất nhưng có số người tham gia đông đảo nhất trên Internet hiện nay.

Để có thể sử dụng dịch vụ E-mail, cần phải có một tài khoản trong hệ thống. Mỗi tài khoản ứng với một người sử dụng trong hệ thống thư. Đó chính là địa chỉ thư điện tử. Mỗi địa chỉ E-mail có dạng: userid@domain, trong đó domain là tên máy tính chủ chứa hệ thống thư và userid là tên tài khoản trong hệ thống này. Ví dụ: Anh@FPT.VN.

Muốn trao đổi thư điện tử với ai bạn phải biết địa chỉ thư điện tử của người mà bạn cần giao tiếp. Sau đó phải truy nhập một phần mềm ứng dụng điều khiển các công việc soạn thảo, chuyển phát và tiếp nhận thư.

Để bảo đảm tính bảo mật, mỗi tài khoản này cần phải có một mật khẩu riêng để không cho người khác thâm nhập vào tài khoản của mình.

Bạn cũng có thể tham gia vào các hội thảo điện tử có tính toàn cầu thông qua dịch vụ nhóm tin (News Groups). Dịch vụ nhóm tin có thể coi là một diễn đàn thông tin công cộng mà trên đó người ta trao đổi và thảo luận về những chủ đề mà họ quan tâm. Mỗi chủ đề là một nhóm tin. Có nhiều nhóm tin khác nhau, mỗi nhóm quan tâm đến một chủ đề như: hội họa, âm nhạc, thể thao, công nghệ máy tính, v.v... Hiện tại có hơn 28 000 diễn đàn trao đổi thông tin trên Internet, trung bình mỗi ngày có hơn 200 000 thông tin được đưa lên mạng.

Dữ liệu của nhóm tin được quản lý bởi chương trình NewsServer. Bạn sẽ sử dụng phần mềm NewsGroups client để đọc thông tin trên diễn đàn. Thông tin được lưu trữ tại các server và được gọi là news site.

Mỗi nhóm tin được đặt một tên. Tên của các nhóm tin được cấu trúc theo kiểu phân cấp. Tên gồm nhiều phần, mỗi phần được phân cách nhau bằng một dấu chấm. Phần đầu là chủ đề của nhóm tin. Phần sau là những đề mục chi tiết hơn.

Với dịch vụ nhóm tin, người sử dụng có thể nhận được thông tin mà mình quan tâm của nhiều người từ khắp nơi, đồng thời có thể gửi thông tin của mình đi cho những người có cùng mối quan tâm này.

2- Truyền tệp (FTP - File Transfer Protocol)

Bạn có thể gửi/nhận thông tin tới/từ các máy tính khác đặt ở xa nhờ giao thức truyền tệp (File Transfer Protocol).

FTP là dịch vụ cho phép sao chép thông tin (theo cả hai chiều) từ các kho dữ liệu trên Internet với máy tính cá nhân của bạn.

Khi máy tính của bạn đã hoà mạng Internet, nhờ FTP bạn có thể ghép nối với bất kỳ máy chủ nào trong hệ thống, xem các tệp trong các thư mục của máy chủ, thực hiện các tệp đó (đọc nội dung, chạy chương trình...). Nếu thấy cần, bạn có thể sao chép các tệp đó về máy tính của mình (download). Chính bằng cách này bạn có thể tiếp nhận được các tài liệu điện tử (sách, báo điện tử) có rất nhiều trên các nguồn tin của Internet. Ngược lại bạn cũng có thể truyền các tệp từ máy tính của bạn về máy chủ (nếu được phép).

3- Truy nhập máy chủ (Telnet)

Telnet là một dịch vụ cơ bản của Internet, là thủ tục để kết nối máy tính của bạn với máy tính khác. Máy tính khác này thường là máy chủ lớn (Mainframe hoặc Mini) đặt từ xa. Khi đó máy tính của bạn sẽ trở thành một terminal làm việc của máy chủ đó. Các terminal có thể sử dụng hoàn toàn các thiết bị và tài nguyên của máy tính lớn như chạy chương trình trên máy tính đó hoặc truy nhập các CSDL và nhiều dịch vụ thông tin khác. Điều này rất có ý nghĩa vì máy tính lớn đã và sẽ còn đóng vai trò quan trọng trong nghiên cứu và áp dụng các thành tựu khoa học mới.

Những máy tính chủ cho phép người sử dụng truy nhập tới và chạy các chương trình ứng dụng hay tra cứu các thông tin lưu trữ trên các CSDL của chúng được gọi là các trạm đăng ký khai thác công cộng.

4- Tra cứu tìm tin trên các CSDL và NHDL

Như chúng ta đã biết các cơ quan dịch vụ thông tin như DIALOG, QUESTEL, DATA STAR, ORBIT... đảm nhiệm việc cung cấp thông tin của hàng trăm NHDL trên thế giới cho người dùng tin. Ngày nay hầu như tất cả các cơ quan dịch vụ thông tin lớn nói trên đều có thể tiếp cận qua mạng Internet.

Chính lệnh Telnet cho phép chuyển máy tính cá nhân của người sử dụng thành terminal. Việc kết nối được thiết lập bằng cách cho tên hay địa chỉ IP của server trên Internet. Khi đó bạn có thể tra cứu thông tin trên các NHDL thông qua các cơ quan dịch vụ thông tin mà bạn đã đăng ký truy nhập.

Trong công tác thông tin tư liệu ta biết rằng có thể tra cứu tìm tin trên các thư mục công cộng truy nhập trực tuyến, viết tắt là OPAC (On-line Public Access). Đó là các thư mục điện tử hoá (CSDL thư mục) được thiết kế sao cho người sử dụng có thể truy nhập trực tiếp mà không cần sự hỗ trợ trung gian của nhân viên thư viện. Trong danh mục các nguồn tin của Internet xuất bản ở Mỹ năm 1991 đã giới thiệu hàng trăm thư mục OPAC của các thư viện tổng hợp của Mỹ, Anh và một số nước ở châu Âu, châu Á.

Hiện nay có tới trên hai triệu máy tính đã đăng ký cung cấp thông tin trên mạng Internet. Sự gia tăng quá nhiều các nguồn thông tin trên Internet đã gây lúng túng cho người sử dụng do tính phân tán của các nguồn thông tin này. Vì thế người ta phải thiết kế các công cụ quản lý thông tin dựa trên mô hình khách/chủ, cho phép người sử dụng có thể tra cứu trực tiếp các nguồn tin mà không cần sự trợ giúp trung gian. Ngày nay ta có thể thực hiện tìm tin phân tán trên Internet qua các dịch vụ sau:

xuất bản tài liệu mà không có sự can thiệp của các đơn vị thông tin.

Phổ biến thông tin là lý do tồn tại và là điều quan tâm chủ yếu của đơn vị thông tin. Việc xây dựng và duy trì vốn tài liệu là phương tiện chủ yếu để đạt được mục tiêu trên.

Việc phổ biến thông tin có thể thực hiện trên:

- Các tài liệu cấp một hay các bản sao của chúng.
- Các bảng tra cứu của tài liệu dưới dạng các sản phẩm thông tin cấp hai, các CSDL thư mục.
- Các thông tin chứa trong tài liệu và được trình bày dưới dạng các sản phẩm thông tin cấp ba.
- Các nguồn thông tin khác nhau.

Như vậy, phổ biến thông tin là thuật ngữ chung để chỉ tất cả các dịch vụ nhằm cung cấp cho người dùng tin những tài liệu cấp một, những bản tra cứu hay nguồn của chúng, những thông tin có thể khai thác trực tiếp, được trình bày dưới dạng phù hợp với yêu cầu của người dùng tin.

Các dạng phổ biến thông tin khác nhau tùy theo phương tiện truyền thông được sử dụng, tính định kỳ, phạm vi và tính chất yêu cầu đề xuất bởi người dùng tin.

Hình thức phổ biến thông tin thường gặp nhất là phổ biến thông tin dưới dạng văn bản. Các văn bản này có thể ghi trên các tài liệu thông thường hoặc lưu trữ trong bộ nhớ của MTĐT và được phổ biến thông qua thiết bị in và thiết bị đọc.

Tuy nhiên việc phổ biến thông tin bằng lời giữa các cá nhân hoặc giữa các nhóm người (như hội nghị, tham quan, lên lớp...) cũng giữ một vị trí quan trọng. Người ta cũng phổ

khác. Nhờ đó mà từ một từ ngữ được chọn trong văn bản, ta có thể truy nhập đến các tài liệu khác chứa các thông tin bổ sung. Các tài liệu này có thể là văn bản, hình ảnh, âm thanh hay hỗn hợp của chúng. Các từ ngữ có các liên kết này chính là các từ khoá và được in đậm nét trên màn hình để phân biệt chúng với các từ ngữ khác.

Khả năng tự động kết nối này gọi là kết nối siêu văn bản (hypertext links), làm cho Web trở nên trong suốt đối với các kho dữ liệu và làm cho nó trở thành một dịch vụ thông tin hấp dẫn nhất và phát triển nhanh nhất hiện nay. Hiện nay có trên 200 bộ phục vụ WWW trên Internet, cung cấp một lượng thông tin phong phú, đa dạng mà không dễ dàng gì truy nhập được bằng các công cụ mạng khác.

Trên đây chỉ là một số dịch vụ cơ bản của Internet. Có thể nói khả năng của Internet là vô hạn. Tương lai Internet sẽ đóng vai trò rất quan trọng trong giáo dục từ xa. Rồi đây các thư viện trên khắp thế giới sẽ sớm được đưa lên tuyến của mạng Internet. Và khi đó chỉ trong vài phút bạn có thể truy nhập vào danh mục các tài liệu của Thư viện Đại học Oxford (UK), thư viện Quốc hội Mỹ hay bất kỳ thư viện của một trường đại học nào trên thế giới.

2. CÁC DỊCH VỤ PHỔ BIẾN THÔNG TIN

2.1. Khái niệm chung về phổ biến thông tin

Phổ biến thông tin là chuyển giao cho người dùng tin những thông tin mà họ cần hoặc giúp cho họ có khả năng tiếp cận các thông tin đó. Đó là kết quả của việc xử lý thông tin và là công đoạn cuối cùng của dây chuyền thông tin. Tuy nhiên một vài dạng phổ biến thông tin có thể thực hiện ngay khi

xuất bản tài liệu mà không có sự can thiệp của các đơn vị thông tin.

Phổ biến thông tin là lý do tồn tại và là điều quan tâm chủ yếu của đơn vị thông tin. Việc xây dựng và duy trì vốn tài liệu là phương tiện chủ yếu để đạt được mục tiêu trên.

Việc phổ biến thông tin có thể thực hiện trên:

- Các tài liệu cấp một hay các bản sao của chúng.
- Các bảng tra cứu của tài liệu dưới dạng các sản phẩm thông tin cấp hai, các CSDL thư mục.
- Các thông tin chứa trong tài liệu và được trình bày dưới dạng các sản phẩm thông tin cấp ba.
- Các nguồn thông tin khác nhau.

Như vậy, phổ biến thông tin là thuật ngữ chung để chỉ tất cả các dịch vụ nhằm cung cấp cho người dùng tin những tài liệu cấp một, những bản tra cứu hay nguồn của chúng, những thông tin có thể khai thác trực tiếp, được trình bày dưới dạng phù hợp với yêu cầu của người dùng tin.

Các dạng phổ biến thông tin khác nhau tùy theo phương tiện truyền thông được sử dụng, tính định kỳ, phạm vi và tính chất yêu cầu đề xuất bởi người dùng tin.

Hình thức phổ biến thông tin thường gặp nhất là phổ biến thông tin dưới dạng văn bản. Các văn bản này có thể ghi trên các tài liệu thông thường hoặc lưu trữ trong bộ nhớ của MTĐT và được phổ biến thông qua thiết bị in và thiết bị đọc.

Tuy nhiên việc phổ biến thông tin bằng lời giữa các cá nhân hoặc giữa các nhóm người (như hội nghị, tham quan, lên lớp...) cũng giữ một vị trí quan trọng. Người ta cũng phổ

biến thông tin bằng các phương tiện nghe nhìn (phim ảnh, truyền thanh, truyền hình) hoặc tổ chức các hình thức triển lãm, ở đó người ta huy động tất cả các phương tiện: phân phát tài liệu, tiếp xúc cá nhân, giới thiệu sản phẩm...

Việc phổ biến thông tin có thể đột xuất hay thường xuyên. Yêu cầu đột xuất về thông tin thường được thoả mãn bằng việc tìm tin đúng chỗ, do người dùng tin trực tiếp tiến hành hay do cán bộ thông tin thực hiện. Yêu cầu thông tin thường xuyên được thoả mãn bằng các công đoạn kế tiếp do đơn vị thông tin thực hiện, nhằm tìm và cung cấp thông tin định kỳ cho người dùng tin, dưới dạng các tài liệu cấp hai hay cấp ba.

Việc phổ biến thông tin có thể thụ động khi phải giải quyết yêu cầu theo đơn đặt hàng của người dùng tin. Nhưng đơn vị thông tin cần phải đi trước người dùng tin, chủ động cung cấp cho họ những sản phẩm thông tin có ích, trên cơ sở phân tích các yêu cầu của họ và phát hiện những vấn đề thông tin nảy sinh cần giải quyết. Đó là phổ biến thông tin chủ động.

Những dịch vụ phổ biến thông tin mà một đơn vị thông tin thường phải đảm nhiệm là:

- Các dịch vụ cung cấp tài liệu cấp một.
- Dịch vụ dịch tài liệu.
- Các dịch vụ phổ biến thông tin cấp hai, cấp ba.
- Các dịch vụ tìm tin on-line, tìm tin trên CD-ROM.
- Phổ biến thông tin có chọn lọc.

Những trở ngại cho việc phổ biến thông tin có thể là:

- Những trở ngại liên quan đến tính đa dạng trong yêu cầu của người dùng tin, sự mơ hồ trong các yêu cầu của họ, sự không hiểu biết của họ đối với hoạt động và khả năng thực tế của đơn vị thông tin.

- Những trở ngại về mặt tài chính, làm cho đơn vị thông tin không có đầy đủ điều kiện và phương tiện để thoả mãn yêu cầu của người dùng tin với các sản phẩm thông tin có chất lượng.

- Những trở ngại về tổ chức liên quan tới vị trí của các đơn vị thông tin: xa các nguồn thông tin, không bổ sung được các thông tin quý hiếm.

- Những trở ngại về kỹ thuật như thiếu các trang thiết bị cần thiết, thiếu các cán bộ chuyên môn giỏi, xử lý tài liệu chậm làm ảnh hưởng đến chất lượng và hiệu quả phục vụ.

- Những trở ngại do yếu kém về quản lý gây ra.

2.2. Các dịch vụ cung cấp tài liệu cấp một

Việc phổ biến tài liệu cấp một có thể tiến hành theo nhiều hình thức khác nhau:

- Cho mượn đọc tại chỗ ở một đơn vị thông tin.
- Cho mượn đọc ở nhà theo một thời gian quy định.
- Bán cho người dùng tin tài liệu ở dạng tài liệu gốc hay bản sao chụp của tài liệu gốc.

Việc cho mượn đọc tại chỗ có thể thực hiện bằng hai cách:

- Cho mượn ở phòng đọc có kiểm soát, thông qua hệ thống thẻ đọc. Ở đây người dùng tin tra cứu hệ thống mục lục rồi ghi vào phiếu yêu cầu các chỉ dẫn thư mục của tài liệu cần

tìm, sau đó đưa cho thủ thư. Thủ thư tìm tài liệu trong kho và cung cấp cho người dùng tin đọc tại chỗ ở phòng đọc. Họ phải trả lại tài liệu vào cuối buổi phục vụ mỗi ngày. Hình thức này thường áp dụng đối với các kho tài liệu đặc biệt, các tài liệu quý hiếm.

- Tra cứu tự do là hình thức mà ở đây người dùng tin có thể tự tìm tài liệu trên giá, tra cứu tại chỗ rồi lại đặt tài liệu về vị trí cũ của nó. Hình thức này thường áp dụng cho các tài liệu tra cứu.

Chế độ cho mượn đọc tại nhà thường phục vụ cho các cán bộ nghiên cứu, nhưng loại trừ các tài liệu quý hiếm và các tài liệu tra cứu mà người đọc thường xuyên sử dụng. Các đơn vị thông tin cần ấn định số lượng và thời gian mỗi lần cho mượn. Một người không thể mượn quá lâu một tài liệu có thể cần cho người khác.

Hai hình thức trên chính là cách phục vụ truyền thống của các thư viện. Hình thức bán tài liệu chủ yếu do các trung tâm thông tin tư liệu thực hiện.

Khi bạn muốn có văn bản của một tài liệu mà bạn đã có chỉ dẫn về tài liệu đó, đơn vị thông tin sẽ cung cấp cho bạn bản sao của tài liệu gốc được cất giữ trong kho. Trong trường hợp không có tài liệu trong kho, đơn vị thông tin sẽ tìm cách thoả mãn yêu cầu của bạn bằng cách sử dụng các tài liệu của các đơn vị khác trong mạng lưới.

Hệ thống dịch vụ cung cấp tài liệu cấp một thường được xây dựng dựa trên các yếu tố sau:

- Các mục lục liên hợp.
- Hệ thống cho mượn giữa các thư viện.

- Các trung tâm thông tin tư liệu, có nhiệm vụ thu thập và phân phối các tài liệu khoa học kỹ thuật thuộc đủ các loại cho người dùng tin.

Hệ thống cho mượn giữa các thư viện

Hệ thống cho mượn giữa các thư viện là một hệ thống trong đó một thư viện không có một tài liệu có thể mượn tài liệu đó ở một thư viện khác để cung cấp cho người dùng tin, theo hai hình thức: cho mượn tại chỗ hay cho mượn về nhà.

Hệ thống này được xây dựng dựa trên sự thoả thuận chi tiết giữa các thư viện thành viên và tuân theo những phương thức đã chuẩn hoá trong các thủ tục cho mượn tài liệu như: phiếu đăng ký mượn, điều kiện mượn, cách kiểm tra, cách gửi tài liệu đi,... Hệ thống cho mượn giữa các thư viện thường được thiết lập ở trình độ quốc gia và nhất thiết phải xây dựng mục lục liên hợp để các đơn vị thành viên biết được kết cấu và định vị của vốn tài liệu. Mục lục liên hợp này được in thành một tập duy nhất, cho phép mở rộng đáng kể khả năng phục vụ người dùng tin, hợp lý hoá và chia sẻ trách nhiệm giữa các thư viện thành viên trong việc bổ sung tài liệu.

Ví dụ: Hệ thống cho mượn giữa các thư viện của Pháp, hoạt động tự động hoá hoàn toàn, bao gồm 200 đơn vị thành viên. Mục lục liên hợp quốc gia của Pháp về các ấn phẩm định kỳ (CNN - Catalogue collectif national des publication en série) hoạt động với 400 000 tên ấn phẩm định kỳ.

Cung cấp các bản sao chụp của tài liệu và tài liệu vi dạng

Việc cung cấp các bản sao chụp của tài liệu là một phương tiện thuận lợi cung cấp tài liệu cấp một mà không phải lấy đi

tài liệu của kho tư liệu. Tuy nhiên nó thường chỉ áp dụng với các tài liệu có dung lượng nhỏ và bị khống chế bởi luật bản quyền tác giả. Ngày nay các đơn vị thông tin thường tổ chức một bộ phận riêng, chuyên làm dịch vụ sao chụp tài liệu theo yêu cầu của người dùng tin.

Việc cung cấp các tài liệu vi dạng (microformes) cũng là một phương tiện cung cấp tài liệu cấp một ngày càng được sử dụng nhiều, nhất là đối với các vi phiếu (microfiches). Vi phiếu là một bản phim có kích thước 105×135 mm, có thể chứa bản chụp của một tài liệu dày hàng trăm trang. Ngày nay nhiều cơ quan xuất bản sản xuất tài liệu đồng thời dưới dạng thông thường và dạng vi phiếu. Giá của các tài liệu này tương đối hạ và cước phí vận chuyển hầu như không đáng kể (chỉ bằng tiền gửi một bức thư). Tuy nhiên để sử dụng cá nhân các tài liệu này đòi hỏi cần phải có thiết bị đọc, ngoài ra cũng bị khống chế bởi luật bản quyền như đối với các tài liệu sao chụp.

Cung cấp tài liệu từ xa

Việc cung cấp tài liệu từ xa là một phương thức tương đối mới, chỉ áp dụng trong một số điều kiện và đối với một số loại tài liệu. Các phương thức truyền thường được sử dụng là: truyền điện thư (facsimile transmission, viết tắt là fax), truyền văn bản bằng tín hiệu truyền hình cáp (teletext).

Với công nghệ thông tin hiện đại, ngày nay các nhu cầu về cung cấp tài liệu của người dùng tin được xử lý một cách tự động và bản sao của tài liệu được gửi tới tận nhà người dùng tin trong thời gian ngắn nhất bằng đường bưu điện hoặc bằng fax. Người dùng tin phải trả tiền bản sao tài liệu và cước phí chuyển giao tài liệu.

Ví dụ: Dịch vụ cung cấp tài liệu của INIST thực hiện hoàn toàn tự động. Ngay sau khi nhận được yêu cầu của người dùng tin, yêu cầu đó được đưa vào xử lý trong máy tính. Tài liệu cất giữ trong kho sẽ được sao chụp và gửi tới người dùng tin bằng đường bưu điện (theo dịch vụ thông thường trong vòng 48 giờ, theo dịch vụ nhanh trong vòng 24 giờ) hoặc bằng fax. Trong trường hợp tài liệu không có trong kho, theo sự thoả thuận của người dùng tin, tài liệu sẽ được tìm trong mạng lưới gồm một trăm thư viện thành viên ở châu Âu. Kho tư liệu của INIST về khoa học, công nghệ, y học, khoa học xã hội và nhân văn cung cấp hàng năm gần 600000 bản sao chụp tài liệu cho hơn 6000 khách hàng ở Pháp và thế giới.

Vấn đề luật bản quyền

Việc cung cấp tài liệu phải tuân theo luật bản quyền tác giả. Đó là luật bảo vệ quyền lợi vật chất và tinh thần của tác giả đối với tác phẩm của họ. Việc sao chụp một vài bản của một tác phẩm, dùng cho việc nghiên cứu, học tập của cá nhân, không vì mục đích vụ lợi là được phép. Nhưng nếu sao chụp tài liệu cho sử dụng tập thể hay vì mục đích thương mại thì phải được tác giả cho phép theo luật bản quyền và phải trả một số tiền nhất định theo quy định của luật.

Để bảo đảm, các đơn vị thông tin thường phải yêu cầu khách hàng ký cam kết tuân theo luật bản quyền khi sử dụng bản sao của tài liệu do đơn vị thông tin cung cấp.

Các tài liệu được bảo đảm bởi luật bản quyền phải có một ghi chú riêng và năm bắt đầu của luật, theo sau ký hiệu C là chữ đầu của từ "Copyright".

Các tài liệu sắp xuất bản (preprints) và in với số lượng nhỏ là phương tiện giao lưu thông tin trực tiếp của tác giả. Các tác giả thường dùng các bản này tặng bạn bè, đồng nghiệp để thông báo và xin ý kiến. Các đơn vị thông tin có thể quan hệ trực tiếp với tác giả để xin một bản, bổ sung vào vốn tư liệu của mình. Đó cũng là một nguồn tài liệu cung cấp cho người dùng tin.

Để không ngừng nâng cao hiệu quả và khả năng của dịch vụ cung cấp tài liệu, đơn vị thông tin phải thường xuyên bổ sung vốn tài liệu và mở rộng sự hợp tác với các đơn vị thông tin khác trong việc chia sẻ nguồn lực thông tin.

2.3. Dịch vụ dịch tài liệu

Dịch tài liệu là một phương tiện chủ yếu phân phối tài liệu cấp một khi một bộ phận quan trọng của tài liệu được người dùng tin quan tâm lại xuất bản bằng ngôn ngữ mà họ không đọc được.

Một số tạp chí khoa học xuất bản bằng những ngôn ngữ rất khó, lại ít phổ biến như tiếng Nga và tiếng Nhật, được dịch toàn bộ sang tiếng Anh và được xuất bản đồng thời với bản gốc. Người dùng tin có thể sử dụng thuận tiện các bản dịch phù hợp với mình.

Tuy nhiên, không thể tránh được trường hợp phải cung cấp các bản dịch khi người dùng tin yêu cầu. Các bản dịch này thường được thực hiện bởi các nhà dịch thuật chuyên nghiệp, cộng tác với đơn vị thông tin. Các đơn vị thông tin lớn thường tổ chức một bộ phận làm dịch vụ dịch tài liệu riêng, phục vụ yêu cầu dịch tài liệu cho mọi đối tượng khách hàng.

Việc dịch tài liệu có thể thực hiện theo nhiều hình thức khác nhau. Người ta có thể dịch cả tài liệu hay chỉ một phần nào đó của tài liệu, tùy theo yêu cầu của người dùng tin.

Trước khi tiến hành dịch một tài liệu, người ta phải kiểm tra xem bản dịch đó đã có hay chưa, nhất là trong trường hợp dịch toàn bộ tài liệu. Trong trường hợp đã có bản dịch, người ta chỉ cần tìm bản dịch và cung cấp cho khách hàng. Muốn vậy phải có quan hệ chặt chẽ với các trung tâm dịch thuật trong nước và quốc tế. Cơ sở dữ liệu của các bản dịch quốc tế WTI (World Translation Index) do INIST và Trung tâm dịch thuật quốc tế ITC (International Translations Centre) cộng tác sản xuất cho phép xác định các bản dịch của các tài liệu đã có. Từ năm 1979, WTI đã tập hợp được hơn 300 000 bản dịch các tài liệu khoa học kỹ thuật của thế giới ra tiếng các nước phương Tây.

2.4. Các dịch vụ phổ biến thông tin cấp hai

Các dịch vụ phổ biến thông tin cấp hai có thể thực hiện dưới nhiều dạng khác nhau, phụ thuộc vào nội dung, phương thức, thời gian và mục tiêu của chúng.

Dịch vụ tra cứu tin

Dịch vụ tra cứu tin giúp người dùng tin sử dụng tốt nhất nguồn tài liệu hiện có để thỏa mãn nhu cầu thông tin của họ. Các ấn phẩm thư mục, các mục lục, các bảng chỉ dẫn và tóm tắt, các CSDL thư mục và CSDL đa phương tiện và các mục lục liên hợp do các đơn vị thông tin hợp tác biên soạn đã giúp người dùng tin mở rộng phạm vi các nguồn tài liệu và thông tin mà họ cần khai thác. Việc tìm tài liệu theo chủ đề là một phương pháp tìm có hiệu quả mà công nghệ thông tin đem

lại. Nó cho phép người dùng tin có thể tìm tài liệu theo nội dung đến mức độ chi tiết cần thiết. Điều này phụ thuộc trước hết vào chất lượng của việc xử lý nội dung tài liệu (phân loại, đánh chỉ số, làm tóm tắt) khi xây dựng các CSDL. Ngày nay khả năng tìm tin đa dạng và linh hoạt trên các hệ thống thông tin dựa trên máy tính điện tử đã mở ra khả năng to lớn giúp người sử dụng tra cứu các nguồn thông tin rất phong phú và đa dạng.

Dịch vụ thông tin định hướng

Dịch vụ thông tin định hướng có nhiệm vụ chỉ cho người dùng tin không phải là tài liệu hay những thông tin trả lời cho câu hỏi của họ, mà giúp tìm đến các nguồn thông tin. Đó là các ấn phẩm thông tin cấp hai, các đơn vị thông tin, các tổ chức nghề nghiệp và nghiên cứu, các nhà chuyên môn v.v... mà ở đó người ta có thể xác định được các nguồn tin. Ở đây người ta thường sử dụng các mục lục và các bộ phiếu về nguồn tài liệu. Việc cập nhật các bộ phiếu về nguồn tài liệu là yêu cầu bắt buộc.

Các dịch vụ thông tin hiện tại

Các dịch vụ thông tin hiện tại (Current Awareness Services - CAS) có nhiệm vụ giúp người dùng tin nắm bắt được những thông tin mới nhất, đặc biệt trong những lĩnh vực mà họ quan tâm. Công cụ sử dụng ở đây là các loại ấn phẩm định kỳ khác nhau như các ấn phẩm thư mục, các tạp chí tóm tắt và các CSDL thư mục. Chúng lưu ý người dùng tin tới những khía cạnh mới của tri thức và cho phép họ nắm bắt được tình hình phát triển của lĩnh vực tri thức mà họ quan tâm, giúp cho họ duy trì năng lực khoa học kỹ thuật và năng lực quản lý của mình.

Các ấn phẩm thư mục được xuất bản đều kỳ hàng tháng, dựa trên các bản mô tả thư mục của tài liệu mà đơn vị thông tin nhận được trước đó. Hầu hết các thông báo thư mục đều bao gồm các bản tóm tắt và bản chỉ mục gồm các từ khoá. Chúng giúp cho người dùng tin có một ý niệm chính xác về tài liệu và giúp họ chọn ra trong số rất lớn các tài liệu những tài liệu mà họ cần.

Các tạp chí tóm tắt tập hợp các mô tả thư mục cùng với các bản tóm tắt của các công trình nghiên cứu, đăng tải trên các tạp chí khoa học khác nhau. Chúng được xuất bản ngày càng nhiều ở các trung tâm thông tin lớn, bảo đảm việc tìm kiếm tài liệu trong lĩnh vực mà người ta quan tâm. Việc xuất bản tạp chí tóm tắt trở thành một trong những hoạt động chủ yếu trong dịch vụ phổ biến thông tin của một đơn vị thông tin.

2.5. Dịch vụ tìm tin on-line

Một cách tiếp cận thông tin mới

Tất cả các dạng tìm tin truyền thống đều dựa trên hai yếu tố sau đây:

- Một mặt là các “kho thông tin” (các thư viện, các trung tâm thông tin tư liệu...) ở đó lưu trữ các kết quả nghiên cứu dưới dạng các công trình hoặc các bài báo in trong các tạp chí khoa học.

- Mặt khác là các “chỉ dẫn”, nó cho phép ta tìm đến các tài liệu cùng với các mô tả nội dung của chúng. Đó là các thư mục, các hệ thống mục lục, các tạp chí tóm tắt.

Những phương thức xử lý thông tin truyền thống trên đây không đáp ứng được với sự gia tăng nhanh chóng của tài liệu trên thế giới do hiện tượng bùng nổ thông tin gây ra.

Từ giữa những năm 60 người ta bắt đầu áp dụng tin học trong công tác thông tin tư liệu. Các MTĐT tỏ ra rất thích hợp với việc lưu trữ và tìm kiếm nhanh thông tin, đồng thời khả năng bộ nhớ không ngừng mở rộng cho phép đáp ứng với sự gia tăng của tài liệu.

Các nhà sản xuất thông tin, các nhà quản lý các trung tâm thông tin tư liệu đã tận dụng các khả năng này của tin học và từ đầu những năm 70 họ bắt đầu xây dựng những kho thông tin tự động hoá, lưu trữ và khai thác thông tin với dung lượng lớn. Các CSDL thư mục và dữ kiện đầu tiên ra đời.

Từ những năm 70, sự phát triển của kỹ thuật viễn thông cho phép vận chuyển trên mạng lưới điện thoại các thông tin lưu trữ trên máy tính đã đem lại cho con người khả năng tiếp cận thông tin từ xa, theo tuyến (on-line) tới các kho tư liệu đã tin học hóa.

Sự kết hợp giữa tin học và viễn thông đã mở ra triển vọng mới cho việc tìm tài liệu. Từ nay, bằng phương tiện của mạng lưới truyền thông, các kho thông tin có thể khai thác từ những địa điểm xa hàng nghìn cây số.

Các đổi mới kỹ thuật đã kích thích nhu cầu và mở rộng công chúng phục vụ đối với các kho thông tin đã được tin học hoá: thông tin đã tìm thấy một thị trường và trở thành hàng hoá. Bên cạnh các cơ quan sản xuất thông tin, nhiều cơ quan chuyên làm dịch vụ cung ứng thông tin ra đời. Các thông tin được truyền đi trên mạng lưới viễn thông. Tất cả tạo thành

một hệ thống thông tin tự động hoá với khả năng truy nhập hoàn toàn mới, gọi là hệ thống thông tin trực tuyến (on-line). Ở đó người dùng tin có thể tiếp cận tất cả các CSDL mà cơ quan dịch vụ thông tin đảm nhiệm phân phối.

Mô tả kỹ thuật

Trong hệ thống thông tin on-line, các thông tin được lưu trữ trong bộ nhớ của MTĐT và được tiếp cận thông qua một hay nhiều mạng viễn thông. Trên các mạng này người dùng tin thiết lập một “điểm tiếp cận thông tin”, đó là một thiết bị đầu cuối dùng để hỏi, nó cho phép người dùng tin lựa chọn ra các thông tin và các nguồn tin mà họ cần tìm, đồng thời nhận được các thông tin đó.

Các thiết bị của hệ thống bao gồm:

- Máy tính chủ: cho phép tìm các thông tin nhờ một tập hợp các chương trình (phần mềm). Các dữ liệu được lưu trữ trên các đĩa từ hoặc băng từ.

- Mạng viễn thông: dành cho việc truyền dữ liệu. Ví dụ ở Pháp là mạng TRANSPAC.

- Modem: dùng để biến đổi các tín hiệu nhị phân sản sinh bởi máy tính thành các tín hiệu điện có thể truyền đi trên các tuyến viễn thông và ngược lại.

- Bàn phím: cho phép chuyển các lệnh tới máy tính.

- Màn hình: cho phép theo dõi các đối thoại với máy tính.

Tập hợp các modem, bàn phím, màn hình tạo thành một thiết bị đầu cuối dùng để hỏi. Trong trường hợp của videotex, ba thành phần này được tích nhập trong một thiết bị gọi là Minitel. Videotex là một hệ thống thông tin tương tác có thể

nối nhiều CSDL với những máy ghi nhận sóng truyền hình qua đường dây điện thoại hay qua dây cáp truyền hình. .

- Máy in: cho phép in các lệnh, các câu trả lời của máy tính và các kết quả tìm tin.

Dịch vụ tìm tin on-line

Tìm tin on-line là tìm tin trên các CSDL thông qua một cơ quan dịch vụ thông tin, dựa trên một mạng truyền dữ liệu, xuất phát từ một đầu cuối.

Đối với người dùng tin, để thực hiện tìm tin on-line cần phải:

- Ký hợp đồng với cơ quan dịch vụ thông tin để nhận được một mật khẩu (password). Ví dụ: CSDL PASCAL do các cơ quan dịch vụ thông tin QUESTEL, DIALOG và ESA - IRS đảm nhận.

- Sử dụng một trong các thiết bị sau: một đầu cuối của máy tính, một máy tính có trang bị modem hay một Minitel.

- Biết sử dụng ngôn ngữ hỏi riêng của từng cơ quan dịch vụ.

Tìm tin on-line thông qua cơ quan dịch vụ thông tin là phương thức tìm tin cho phép ta nhận thông tin nhanh trên tất cả các lĩnh vực và với giá rẻ hơn so với phương pháp tìm tin thủ công. Tuy nhiên chất lượng của dịch vụ tìm tin on-line phụ thuộc vào nhiều yếu tố:

- Ở người sản xuất thông tin là tính chọn lọc của các ấn phẩm thông tin, chất lượng của việc đánh chỉ số, chất lượng của việc ghi thông tin trên các đĩa từ, chất lượng của các tài liệu trợ giúp cho người dùng tin...

- Ở cơ quan dịch vụ thông tin là dung lượng và hiệu năng của máy tính chủ, hiệu năng của phần mềm, số lượng và chất lượng của các yếu tố tìm tin, chất lượng đào tạo hướng dẫn người dùng tin,...

- Ở mạng chuyển tải thông tin là khả năng vận hành tốt của mạng truyền dữ liệu, bảo đảm truyền liên tục (không gián đoạn), tính đơn giản của các thủ tục truy nhập mạng (số điện thoại, mật khẩu...).

- Ở người dùng tin là sự hiểu biết về lĩnh vực được hỏi, sự hiểu biết về nội dung và cấu trúc của NHDL, khả năng sử dụng ngôn ngữ hỏi, khả năng sửa đổi các yếu tố của câu hỏi tùy theo các câu trả lời của máy tính.

Các thông tin mà các NHDL cung cấp có thể là:

- Toàn bộ văn bản (các văn bản về luật, các bài báo...).

- Các thông tin dữ kiện khai thác trực tiếp (các dữ liệu thống kê, các hằng số lý hoá, các thông tin về đo lường tiêu chuẩn,...).

- Các tra cứu thư mục, các bản tóm tắt của tài liệu, các chỉ dẫn tới tài liệu gốc hoặc các bản sao của tài liệu gốc.

2.6. Phổ biến thông tin trên CD-ROM

CD-ROM là vật mang tin với dung lượng rất lớn (khoảng 600MB tức là 300 000 trang in hay 1 500 đĩa mềm). Người dùng tin không thể xoá cũng như ghi lại trên đĩa. CD-ROM rất thích hợp với việc lưu trữ các tài liệu văn bản, đồ thị, âm thanh và các dữ liệu tin học. Vì vậy CD-ROM được coi như là một xuất bản phẩm để phổ biến các CSDL lớn, các sách tra cứu như từ điển, bách khoa toàn thư, các phần mềm... Các

CD-ROM đầu tiên ra đời vào năm 1986 và đến cuối năm 1987 đã có 80 tên sản phẩm CD-ROM của Mỹ được bán trên thị trường với 50 000 đầu đọc. Việc cài đặt các CSDL lớn trên CD-ROM cho phép người dùng tin có thể tiếp cận thông tin tại chỗ, đã tạo ra một sự cạnh tranh với các dịch vụ thông tin on-line do các cơ quan dịch vụ thông tin tiến hành, nhất là đối với các CSDL hay được khai thác và sử dụng.

Nhờ các phần mềm khai thác dữ liệu có sẵn trên đĩa, các CD-ROM cơ sở dữ liệu cho phép người dùng tin có thể tìm ngay được các thông tin và in các kết quả ra đĩa hoặc ra giấy mà không gặp một trở ngại nào về thời gian.

Các CD-ROM cơ sở dữ liệu thường sử dụng hai hình thức tra cứu: tra cứu có trợ giúp và tra cứu kiểu chuyên gia và có thể tìm kiếm qua nhiều dấu hiệu như: lĩnh vực, từ vựng, tác giả, cơ quan, ngôn ngữ, các yếu tố xuất bản, loại hình tài liệu, ký hiệu phân loại.

Ngày nay với công nghệ đa phương tiện (multimedia và hypermedia), các thông tin lưu trữ trên CD-ROM có thể thể hiện dưới nhiều dạng khác nhau: văn bản, âm thanh, màu sắc, hình ảnh tĩnh và động. Đồng thời giữa chúng có các điểm kết nối dạng trang web, giúp ta có thể truy nhập thông tin từ lĩnh vực này sang lĩnh vực khác một cách nhanh chóng và thuận lợi.

Để giúp bạn thực hiện các dịch vụ tìm tin trên, các đơn vị thông tin phải cung cấp các tài liệu trợ giúp như:

- Sách hướng dẫn sử dụng CD – ROM.
- Sách hướng dẫn hỏi đáp trên các CSDL.
- Tập từ vựng các từ khoá.

- Các khung phân loại được sử dụng cho CSDL.

2.7. Tìm tin trên Internet

Một phát triển mới nhất trong dịch vụ tìm tin hiện nay là tìm tin trên Internet. Người cán bộ thông tin ngày nay cần phải biết hướng dẫn khách hàng sao cho họ có thể sử dụng có hiệu quả Internet và truy cập tới nguồn thông tin khổng lồ trên Internet.

Có ba cách tìm tin trên Internet:

1. *Tìm theo địa chỉ của nguồn tin - URL.* Nếu bạn biết địa chỉ của nguồn tin - URL (Uniform Resource Locator), bạn có thể gõ trực tiếp địa chỉ mà bạn cần tìm vào dòng Address.

Ví dụ: <http://www.paris.gov/>

Với địa chỉ này bạn sẽ có một trang Web giới thiệu thành phố Paris với nhiều chuyên mục: nghệ thuật, kinh tế, xã hội, du lịch, các địa chỉ e-mail mà bạn quan tâm ở Paris....

Ví dụ: URL của trang chủ mạng VISTA của Trung tâm Thông tin - Tư liệu Khoa học và Công nghệ Quốc gia là:

<http://www.vista.gov.vn/>

2. *Tìm tin trên Website.* Nếu bạn đã tìm được một website, bạn có thể đi từ trang này sang trang khác bằng cách kích vào các điểm kết nối.

Ví dụ: Các thông tin về đất nước và con người Việt Nam, về các điểm du lịch, các lễ hội truyền thống. Ở Việt Nam có thể tìm trên Website của Tổng cục Du lịch Việt Nam

<http://www.vietnamtourist.com>

Ví dụ: Các thông tin về khổ mẫu MARC có thể tìm trên Website của Thư viện Quốc hội Mỹ

<http://lcweb.loc.gov/marc>

3. *Sử dụng máy tìm - search engine.* Nếu không biết địa chỉ, bạn cần sử dụng một máy tìm - search engine. Một số máy tìm nổi tiếng là:

Alta Vista (<http://www.altavista.digital.com>)

Yahoo (<http://www.yahoo.com>)

HotBot (<http://www.hotbot.com>)

Infoseek (<http://www.infoseek.com>)

Lycos (<http://www.lycos.com>)

Để tìm thông tin trên máy tìm bạn có thể thực hiện theo hai cách sau:

- Gõ vào một từ khoá hoặc một từ.

Ví dụ: MARC

Máy sẽ liệt kê các địa chỉ mà ở đó bạn có thể tìm được những thông tin hữu ích về MARC.

- Chọn một trong các chủ đề chung như Sport, Education,... rồi thu hẹp phạm vi tìm thông qua các chủ đề con.

Một số máy tìm cho phép bạn sử dụng các toán tử logic AND, OR, NOT để lập các biểu thức tìm, cho phép bạn tìm được các thông tin theo yêu cầu.

2.8. Các dịch vụ phổ biến thông tin cấp ba

Các tài liệu cấp ba là kết quả tổng hợp của nhiều công đoạn khác nhau. Đó là tập hợp các thông tin được cung cấp bởi nhiều nguồn, được phân tích, đánh giá cô đọng và rút ra những dữ liệu. Chúng có thể được biên tập để trả lời một câu hỏi riêng (tổng quan, tổng luận, thực trạng của một vấn đề...) hoặc để xuất bản đều kỳ hàng năm (tạp chí hàng năm). Nó có lợi ích kép là cô đọng thông tin hiện có về một chủ đề đồng thời cung cấp những thông tin có chất lượng cao.

Dịch vụ phổ biến thông tin cấp ba được thực hiện dưới nhiều hình thức khác nhau. Ta có thể kể ở đây một vài hình thức:

- Các dịch vụ thông tin liên kết thông qua môi giới trung gian giữa các dịch vụ thông tin và người dùng tin. Ở đây các chuyên gia của các lĩnh vực tìm đến người dùng tin và xác định những khó khăn về thông tin của họ, rồi đặt quan hệ với các dịch vụ thông tin thích hợp hay khai thác những dịch vụ này theo chủ đích của họ. Các dịch vụ này là phương tiện kích thích sự giao lưu thông tin giữa đơn vị thông tin và người dùng tin. Các cán bộ phụ trách công việc này phải có kiến thức về kỹ thuật và có hiểu biết về nghiệp vụ thông tin.

- Các dịch vụ thông tin nhanh, cung cấp dưới dạng cô đọng những thông tin quan trọng hiện hành, hoặc để giúp người dùng tin nắm bắt được những thông tin đang diễn ra trong lĩnh vực của họ, hoặc để trả lời cho một yêu cầu thường xuyên của họ.

- Phổ biến thông tin qua tiếp xúc trực tiếp là hình thức mà các đơn vị thông tin thường xem nhẹ. Tuy nhiên người dùng tin thường tìm thấy ở đây những thông tin mà họ cần

tìm và nó cung cấp cho họ những chỉ dẫn chắc chắn, do khả năng đối thoại, đánh giá, lựa chọn mà dịch vụ này đem lại. Người ta có thể tổ chức các cuộc tiếp xúc giữa người dùng tin và cán bộ hay cơ quan thông tin hoạt động trong lĩnh vực mà họ quan tâm hay tổ chức các hội nghị thông tin, các hội thảo khoa học.

Những hoạt động như vậy đem lại lợi ích rõ rệt là thúc đẩy các hoạt động thông tin, làm cho quá trình chuyển giao thông tin trở nên năng động. Nó cho phép bổ khuyết các nguồn thông tin bị thiếu.

2.9. Phổ biến thông tin có chọn lọc

Phổ biến thông tin có chọn lọc (SDI- Selective Dissemination of Information) là một phương thức chủ động nhằm cung cấp cho người dùng tin những thông tin mới, phù hợp với yêu cầu thường xuyên đã được xác định và đăng ký trước của họ.

Nói cách khác, SDI là dịch vụ cung cấp cho người dùng tin những thông tin cập nhật theo một chủ đề nhất định. Những thông tin này được chọn ra từ tất cả các bản mô tả nội dung tài liệu mà cơ quan thông tin mới nhận được trong thời gian đó.

Lợi ích của SDI là nó giúp cho các nhà khoa học tiết kiệm được rất nhiều thời gian trong việc tìm kiếm thông tin. Họ không còn phải mất thời gian để đọc tất cả các tạp chí tóm tắt, mà số này thường là rất nhiều. Họ cũng không phải mất công chọn tài liệu mà họ cần tìm. Dịch vụ này cũng bảo đảm sự liên hệ thường xuyên giữa người dùng tin và cơ quan thông tin.

Hiệu quả của SDI phụ thuộc vào hai yếu tố: chất lượng của các bản tóm tắt và tính tương thích của profil của người dùng tin.

Vậy profil là gì?

Profil của người dùng tin là một tập hợp các từ khoá có cấu trúc phù hợp với hệ thống mô tả các chủ đề mà người dùng tin quan tâm. Nói cách khác profil của người dùng tin là một biểu thức tìm, thể hiện những thông tin mà họ yêu cầu cung cấp một cách đều đặn, do dịch vụ phổ biến thông tin có chọn lọc thực hiện. Đó là một mô tả những thông tin mà họ quan tâm trong một giai đoạn nhất định.

Có profil cá nhân, tương ứng với yêu cầu của một người dùng tin. Có profil tập thể, tương ứng với yêu cầu của một nhóm người dùng tin hoạt động trên một lĩnh vực.

Với mỗi một CSDL thư mục, một profil phải được thiết lập theo tổ chức và ngôn ngữ riêng của CSDL đó. Về cơ bản, việc xây dựng một profil không khác việc xây dựng một biểu thức tìm nhưng đòi hỏi phải cẩn thận hơn, vì profil được sử dụng lặp đi lặp lại nhiều lần trong một thời gian khá dài. Một profil quá hẹp có thể dẫn đến không tìm được câu trả lời nhưng nếu quá rộng thì lại tìm được quá nhiều tài liệu tương thích.

Khi đã có profil, thủ tục tìm tin được tiến hành như sau: Những từ khoá của profil được đối chiếu với những từ khoá có trong mô tả nội dung tài liệu và nếu chúng phù hợp nhau thì đó là tài liệu cần lựa chọn. Các bản tóm tắt tương ứng sẽ được gửi đến cho người dùng tin. Thông thường có kèm theo một phiếu kiểm tra để trên đó người dùng tin ghi rõ tài liệu có

phù hợp hay không và nếu phù hợp thì họ cần bản sao hay tài liệu gốc.

Biên soạn profil là một việc làm công phu, đòi hỏi sự tham gia của các chuyên gia thông tin giỏi và cần có sự đóng góp của người dùng tin. Nó thường được tiến hành theo nhiều bước và được kiểm tra cập nhật đều đặn.

Đối với một số lượng người dùng tin hạn chế, người ta có thể thực hiện SDI bằng phương thức thủ công. Tuy nhiên hệ thống này thường được thực hiện một cách tự động bằng máy tính điện tử với sự tham gia của các CSDL thư mục lưu trữ trong máy tính.

Dịch vụ SDI hoạt động trên cơ sở các đơn đặt hàng của người dùng tin. Các đơn đặt hàng này được xây dựng trên cơ sở các CSDL thư mục được sử dụng và danh mục các từ khoá có trong các profil. Mỗi khi CSDL cập nhật thêm những tài liệu mới, theo định kỳ hàng tuần hay hàng tháng các thông tin chọn ra sẽ được gửi cho người dùng tin. Dịch vụ SDI có thể bảo đảm phục vụ cho từng cá nhân hay các nhóm người dùng tin. Hiện nay đây là phương thức tốt nhất trong dịch vụ thông tin hiện tại.

3. NGƯỜI DÙNG TIN VÀ NHU CẦU THÔNG TIN

Vai trò của người dùng tin

Người dùng tin là yếu tố cơ bản của mọi hệ thống thông tin. Đó là đối tượng phục vụ của công tác thông tin tư liệu. Người dùng tin vừa là khách hàng của các dịch vụ thông tin, đồng thời họ cũng là người sản sinh ra thông tin mới.

Người dùng tin giữ vai trò quan trọng trong các hệ thống thông tin. Họ như là yếu tố tương tác hai chiều với các đơn vị thông tin. Vai trò đó thể hiện trên các mặt sau:

- Người dùng tin luôn là cơ sở để định hướng các hoạt động của đơn vị thông tin.

- Người dùng tin tham gia vào hầu hết các công đoạn của dây chuyền thông tin. Họ biết các nguồn thông tin và có thể thông báo hoặc đánh giá các nguồn thông tin đó. Chính sách bổ sung phụ thuộc vào yêu cầu của người dùng tin. Họ có thể tham gia vào việc hiệu chỉnh các hoạt động kỹ thuật như xây dựng ngôn ngữ tư liệu, xác định cấu trúc của các bộ phiếu. Họ tham gia vào việc mô tả nội dung, hình thành chiến lược tìm tin và đánh giá các kết quả tìm được.

- Người dùng tin cũng tham gia sản sinh ra thông tin mới, tham gia vào các dòng thông tin bằng tiếp xúc cá nhân.

Tóm lại, người dùng tin là yếu tố thiết yếu của một trung tâm thông tin. Họ là yếu tố năng động trong hoạt động của một trung tâm thông tin.

Những trở ngại trong việc đối thoại giữa cán bộ thông tin và người dùng tin thường gây ra do nhiều người dùng tin không nhận thức được hoạt động thông tin là một tập hợp các nhiệm vụ đặc biệt, đòi hỏi phải có kỹ thuật chuyên môn và có tổ chức. Họ không coi trọng cán bộ thông tin và các cơ quan thông tin. Và lại các cán bộ thông tin, về phần mình lại có xu hướng thu vào các loại hoạt động kỹ thuật và không chú ý đến việc phổ biến thông tin và nhu cầu của người dùng tin. Để khắc phục tình trạng đó cần tránh những quan điểm đơn giản lệch lạc từ hai phía. Người cán bộ thông tin phải nhận thức được rằng mục đích của hoạt động thông tin là phục vụ

người dùng tin và cần có thái độ hợp tác với họ. Về phần người dùng tin phải nhận thức được sự cần thiết của việc chuyển giao thông tin, coi trọng các cán bộ thông tin và phải được đào tạo để có thể thích ứng với các kỹ thuật thông tin.

Các nhóm người dùng tin và nhu cầu thông tin của họ

Người ta phân biệt ba nhóm người dùng tin, tùy theo hoạt động xã hội mà họ tham gia. Đó là:

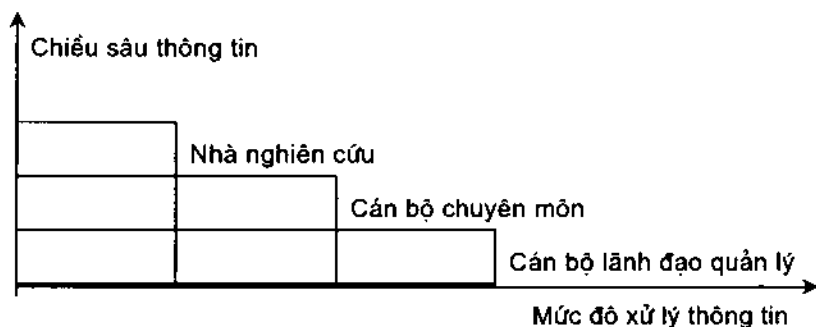
- Các nhà nghiên cứu trong khoa học cơ bản và khoa học ứng dụng, các giáo sư, các sinh viên.
- Các cán bộ chuyên môn, các nhà kỹ thuật và thực hành, các nhà doanh nghiệp hoạt động trong các lĩnh vực khác nhau của kỹ thuật, sản xuất và kinh doanh.
- Các nhà lãnh đạo, các cán bộ quản lý, các cán bộ kế hoạch có trách nhiệm điều hành các hoạt động khoa học kỹ thuật, kinh tế xã hội của một đơn vị, một địa phương hay một quốc gia.

Các nhu cầu thông tin của người dùng tin thường nảy sinh khi họ cần nắm bắt được những kết quả của một lĩnh vực mà họ quan tâm, khi họ cần nắm bắt được các thông tin dữ kiện, những số liệu và phương pháp cần cho công việc của họ.

Các nhu cầu thông tin này thay đổi tùy theo bản chất công việc và nhiệm vụ mà người dùng tin phải hoàn thành. Các nhà lãnh đạo quản lý cần những thông tin xác thực, tổng hợp, cô đọng và đã được phân tích một cách có hệ thống, các thông tin chiến lược có tính dự báo để giúp họ ra được những quyết định đúng đắn. Các cán bộ nghiên cứu cần những

thông tin đầy đủ về những chuyên ngành hẹp, những vấn đề còn đang đặt ra, những thông tin định hướng nghiên cứu trong khoa học. Còn các cán bộ chuyên môn quan tâm đến các thông tin dữ kiện, các số liệu và phương pháp cụ thể liên quan đến các ngành mà họ quan tâm.

Quan hệ giữa chiều sâu của thông tin, mức độ xử lý thông tin và các nhóm người dùng tin trên đây thể hiện ở biểu đồ sau:



Hình 24. Quan hệ giữa chiều sâu thông tin - Mức độ xử lý thông tin - Người dùng tin

Yêu cầu thông tin là dạng thể hiện cụ thể nhu cầu thông tin của người dùng tin. Một khi đã nhận được một yêu cầu thông tin, người cán bộ thông tin phải tìm cách thoả mãn nó. Muốn vậy phải biết nội dung chủ đề, khối lượng thông tin và thời hạn mà người dùng tin yêu cầu.

Cách thức tìm kiếm và chuyển giao thông tin hay thái độ cư xử đối với thông tin, còn gọi là tập quán thông tin của người dùng tin, phụ thuộc trước hết vào việc đào tạo hướng dẫn để người dùng tin biết sử dụng các sản phẩm và các dịch

vụ thông tin. Ngoài ra điều kiện làm việc, trình độ chuyên môn và kinh nghiệm công tác cũng tác động đến tập quán thông tin của người dùng tin.

Quan hệ giữa người dùng tin và đơn vị thông tin phụ thuộc vào yêu cầu và tập quán thông tin của người dùng tin, khả năng phục vụ và chính sách đối với người dùng tin của cơ quan thông tin. Các cán bộ thông tin và cơ quan thông tin phải tạo mối quan hệ mật thiết với người dùng tin vì đó là đối tượng và thước đo hiệu quả hoạt động của họ.

Nghiên cứu người dùng tin

Để nâng cao hiệu quả công tác thông tin, việc nghiên cứu người dùng tin là cần thiết. Việc nghiên cứu người dùng tin nhằm vào các mục tiêu sau:

- Phân tích yêu cầu của người dùng tin, tức là nội dung và các loại hình tài liệu mà người dùng tin cần. Nó cho phép xác định các sản phẩm và dịch vụ thông tin đáp ứng với yêu cầu của người dùng tin.

- Phân tích thói quen và tập quán thông tin của người dùng tin, nó giúp ta hiểu để thoả mãn yêu cầu của người dùng tin, đồng thời xác định được phải hướng dẫn đào tạo đối với người dùng tin.

Những nghiên cứu này cung cấp cơ sở cần thiết cho việc định hướng hoạt động của một đơn vị thông tin, phục vụ cho một đối tượng người dùng tin xác định. Việc nghiên cứu này có thể tiến hành trước khi thành lập đơn vị thông tin và cả ngay trong quá trình hoạt động của nó.

Phương pháp nghiên cứu người dùng tin là phương pháp nghiên cứu tâm lý xã hội. Người ta có thể sử dụng các phương pháp sau:

- Phương pháp phỏng vấn.
- Phương pháp điều tra bằng cách chuẩn bị một hệ thống các câu hỏi.
- Phương pháp quan sát trực tiếp các tập quán thông tin.
- Phân tích các biểu ghi, các mẫu thống kê về tình hình phục vụ và yêu cầu của người dùng tin.
- Tổ chức các buổi họp mặt gặp gỡ với người dùng tin.

Sự phát triển của một đơn vị thông tin đặt ra nhiệm vụ là phải làm cho người dùng tin tiềm năng hiểu về đơn vị thông tin và những lợi ích mà họ có thể nhận được, đồng thời phải làm cho họ tham gia và ủng hộ hoạt động của đơn vị thông tin. Người ta có thể dùng nhiều phương pháp khác nhau như: tổ chức tham quan, tiếp xúc cá nhân, tuyên truyền trên các phương tiện thông tin đại chúng, cung cấp hoặc giới thiệu các dịch vụ thông tin thử nghiệm v.v...

Đào tạo người dùng tin

Việc đào tạo người dùng tin nhằm giúp họ hiểu được cơ chế tổ chức của công tác thông tin tư liệu và biết sử dụng các nguồn thông tin hiện có. Người ta phân biệt nhiều mức độ đào tạo khác nhau:

- Mức độ cảm thụ: Cung cấp những hiểu biết chung.
- Mức độ định hướng: Hướng dẫn một cách tóm tắt các nguồn và cách sử dụng chúng.

- Đào tạo sơ cấp: Tổ chức các lớp ngắn hạn, cung cấp cho người dùng tin những hiểu biết cần thiết về các dịch vụ và các phương tiện chuyển giao thông tin hiện đại.

- Đào tạo chuyên sâu: Cung cấp cho người dùng tin là cộng tác viên những hiểu biết chuyên sâu về một lĩnh vực hoạt động của đơn vị thông tin để họ có thể tham gia vào một dịch vụ thông tin cụ thể trong trường hợp cần thiết. Ví dụ các cộng tác viên là các nhà khoa học tham gia biên tập các tóm tắt, tổng luận...

Việc đào tạo người dùng tin thường là trách nhiệm không thể thiếu của một cơ quan thông tin hay một tổ chức quốc gia hay quốc tế. Nó có thể coi là một bộ phận của các chương trình giáo dục.

Chương 7

CÁC HỆ THỐNG THÔNG TIN

1. MỞ ĐẦU

Phương tiện truyền bá đầu tiên cho quá trình xử lý có chủ đích đối với thông tin là hệ thống thông tin. Hệ thống này được cấu trúc để thu thập, tổ chức, lưu trữ và hiển thị thông tin ở tất cả các dạng của nó (dữ liệu thô, dữ liệu đã được giải thích, tri thức) và các hình thức thể hiện (văn bản, hình ảnh, âm thanh). Về nguyên tắc mọi hệ thống lưu trữ và chỉ dẫn như danh bạ điện thoại, các danh mục tra cứu tin... đều có thể coi là hệ thống thông tin. Tuy nhiên ở đây ta đề cập tới hệ thống thông tin điện tử, được xây dựng trên công nghệ thông tin hiện đại.

Hệ thống thông tin điện tử là hiện tượng của nửa sau của thế kỷ XX. Sự phát triển của nó gắn liền với những tiến bộ của hai công nghệ cơ bản: mạch tích hợp và truyền thông số.

Mạch tích hợp là các chip silic chứa các transistor dùng để lưu trữ và xử lý thông tin. Ngày nay một chip có khả năng lưu trữ 125 000 000 bytes, tương đương với 14 500 trang sách. Tốc độ xử lý của các chip vi xử lý đạt tới hàng ngàn triệu phép toán trong một giây. Để các hệ thống thông tin có thể khai thác được các tiến bộ kỹ thuật trên, đòi hỏi phải có

sự phát triển tương ứng trong lĩnh vực phần mềm. Đóng góp lớn nhất của nó là đã mở ra khả năng sử dụng máy tính cho các cá nhân. Các phần mềm quản trị CSDL ra đời và phát triển đã làm cho các CSDL trở thành yếu tố trung tâm của mọi hệ thống thông tin.

Lợi ích của máy tính tăng lên rất nhiều khi chúng có thể kết nối với các máy tính khác để chia sẻ thông tin và xử lý chúng. Từ đó xuất hiện các mạng cục bộ, mạng diện rộng.

Người ta phân biệt các loại hệ thống thông tin theo mục đích của nó: hỗ trợ cho các hoạt động quản lý điều hành hay hỗ trợ cho hoạt động tác nghiệp và dịch vụ.

2. QUẢN LÝ VÀ THÔNG TIN TRONG QUẢN LÝ

2.1. Tổ chức và quản lý

Tổ chức là gì?

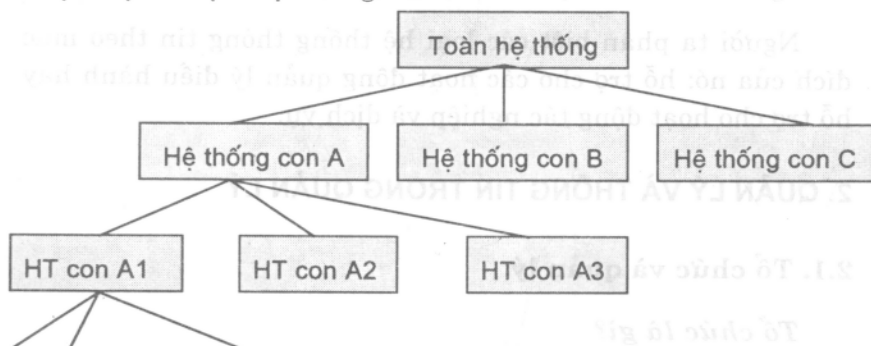
Tổ chức là một tổ hợp các thành phần có quan hệ chặt chẽ với nhau, cùng hỗ trợ và thúc đẩy lẫn nhau nhằm thực hiện nhiệm vụ chung. Tổ chức bao giờ cũng hoạt động trong một môi trường nhất định và chịu tác động của môi trường đó.

Tổ chức được tạo thành bởi một số yếu tố sau:

- Con người làm việc ở đó.
- Các phương thức mà họ sử dụng.
- Các thiết bị mà họ vận hành.
- Các chất liệu mà họ sử dụng.
- Ngân sách cần thiết để chi phí cho các hoạt động.

Các thông tin được sản sinh và sử dụng để điều khiển hoạt động của tổ chức được coi là một nguồn lực của tổ chức. Nguồn lực này được tạo ra và khai thác để hỗ trợ cho việc thực hiện mục tiêu của tổ chức.

Trên quan điểm hệ thống, hoạt động của một tổ chức có thể coi như hoạt động của một hệ thống có cấp bậc các hệ thống con, mà mỗi bộ phận có trách nhiệm đóng góp vào việc thực hiện nhiệm vụ chung của tổ chức.



Hình 25. Hệ thống cấp bậc của tổ chức

Trong tổ chức, mỗi hệ thống con đều có chức năng nhiệm vụ riêng theo sự phân công của tổ chức. Hệ thống thu thập, xử lý và khai thác dữ liệu gọi là hệ thống thông tin. Đó là một thành phần quan trọng trong các nguồn lực của tổ chức, nó là sợi dây liên hệ giữa các thành viên của tổ chức và là nguồn động lực thúc đẩy việc tổ chức và sử dụng các nguồn lực khác nhằm hoàn thành mục tiêu của tổ chức.

Người ta phân biệt các loại tổ chức sau:

- Tổ chức hành chính, sự nghiệp: các cơ quan quản lý hành chính, quản lý sự nghiệp...

- Tổ chức thương mại: các hãng, các công ty, các tổ chức kinh doanh.

- Tổ chức sản xuất: các nhà máy, xí nghiệp,...

- Tổ chức dịch vụ: các cơ quan dịch vụ công cộng.

Quản lý là gì?

Quản lý là một hoạt động thiết yếu của con người nhằm thiết kế và duy trì một môi trường làm việc bên trong và bên ngoài của tổ chức, sao cho nó bảo đảm sự phối hợp những nỗ lực của các cá nhân, các bộ phận để có thể hoàn thành các nhiệm vụ và các mục tiêu đã định, trên cơ sở sử dụng tốt nhất các nguồn tài nguyên. Các tài nguyên ở đây bao gồm: con người, tiền của, vật chất, năng lượng, không gian, thời gian, v.v...

Trên phương diện thực hành thì cách quản lý là một nghệ thuật. Còn kiến thức có tổ chức về quản lý, làm cơ sở cho hoạt động quản lý, có thể coi là một khoa học.

Quá trình quản lý có thể được xác định như một loạt các hoạt động định hướng theo mục tiêu, trong đó có các hành động cơ bản là: xác định mục tiêu, lập kế hoạch để thực hiện mục tiêu, tổ chức và kiểm tra việc thực hiện kế hoạch đó.

Quản lý hiện đại cũng là một tinh thần, một thái độ làm việc nhằm hướng tới tính hiệu quả và hợp lý. Quản lý chỉ có hiệu quả khi nó trở thành công việc của mọi thành viên trong tổ chức, ở đó mỗi người có vai trò không thể thay thế được và mỗi người đều phải biết công việc và chịu trách nhiệm về công việc của mình.

Hoạt động quản lý bao trùm lên tất cả các hoạt động của một tổ chức, cũng như tất cả các yếu tố vật chất và con người tạo thành tổ chức đó. Một nhà quản lý phải lập kế hoạch cho các hoạt động của tổ chức mà họ phụ trách, tổ chức bố trí nhân sự, chỉ đạo và điều hành các hoạt động, kiểm tra bằng cách đánh giá các thông tin phản hồi và tiến hành điều chỉnh khi cần thiết. Vì vậy trong bất kỳ tổ chức nào hoạt động quản lý cũng bao gồm các chức năng chủ yếu sau đây:

- **Lập kế hoạch:** bao gồm việc xác lập mục tiêu và lập các kế hoạch dài hạn, ngắn hạn để thực hiện mục tiêu. Trong đó bao gồm các công việc xác định mục tiêu trước mắt và lâu dài, hoạch định chính sách, chiến lược, xác định các phương pháp và tiêu chuẩn. Lập kế hoạch cũng đòi hỏi phải biết nắm bắt và phân tích thời cơ, lựa chọn giải pháp và thiết kế các chương trình để hoàn thành mục tiêu đề ra.

- **Tổ chức:** bao gồm việc xây dựng cơ cấu của tổ chức, xác định các nhiệm vụ phải hoàn thành của các cá nhân và các bộ phận trong tổ chức, phân công trách nhiệm và quyền hạn cho họ.

- **Điều khiển:** đó là sự lãnh đạo thông qua sự truyền đạt các thông báo, chỉ thị và huy động các nguồn lực của tổ chức.

- **Kiểm tra:** bao gồm việc nhận xét, đánh giá sự thực hiện của tổ chức và tác động của môi trường xung quanh và điều chỉnh các kế hoạch và hoạt động của tổ chức khi cần thiết.



Hình 26. Các chức năng của quản lý

Hoạt động quản lý phải dựa trên các cơ sở sau:

- Xác định mục tiêu thực tế.
- Có thông tin đầy đủ, chính xác, kịp thời.
- Định rõ các tiêu chuẩn của nhiệm vụ phải thực hiện.
- Nghiên cứu các điều kiện để bảo đảm cho nhiệm vụ được hoàn thành.
- Kiểm tra, đánh giá các kết quả thu được.

Các nhà quản lý ở mọi cấp, mọi nơi đều có nhiệm vụ cơ bản là thiết kế và duy trì môi trường cho phép các cá nhân làm việc để hoàn thành các nhiệm vụ và các mục tiêu đã định.

Nhiệm vụ quan trọng của quản lý là ra các quyết định. Chất lượng của quyết định phụ thuộc vào chất lượng của các thông tin, các số liệu và dữ kiện được cung cấp.

Hệ thống thông tin đóng vai trò trợ giúp các nhà quản lý trong việc ra các quyết định và thực hiện các chức năng của quản lý.

Các mức độ quản lý

Căn cứ vào mức độ quan trọng và đối tượng của quản lý, người ta thấy có ba mức độ của quản lý là:

- Quản lý mức chiến lược: là mức cao nhất của công tác quản lý, có tác động đến toàn bộ tổ chức. Việc quản lý ở đây nhằm xác định các mục tiêu chiến lược và đường lối chính sách để thực hiện các mục tiêu đó.

- Quản lý mức chiến thuật. Đây là công tác quản lý ở mức trung gian, có quy mô một đơn vị. Quản lý chiến thuật là xác định các chức năng cụ thể cần làm để thực hiện các mục tiêu và đường lối đã đề ra trong mức quản lý chiến lược. Nó định hướng vào các nhiệm vụ và các quá trình.

- Quản lý mức tác nghiệp: là quản lý các công việc hàng ngày, nhằm thực hiện các nhiệm vụ cụ thể, với quy mô phòng làm việc, tổ công tác hoặc một phân xưởng sản xuất.

Các khái niệm quan trọng đối với vấn đề quyết định là quyết định có cấu trúc và quyết định không có cấu trúc. Vấn đề quyết định được gọi là không có cấu trúc nếu nó không có quy trình rõ ràng khi làm ra quyết định và các thông tin liên quan đến các yếu tố cơ bản cần thiết để xem xét trong quá trình ra quyết định không thể xác định được trước một cách thường xuyên.

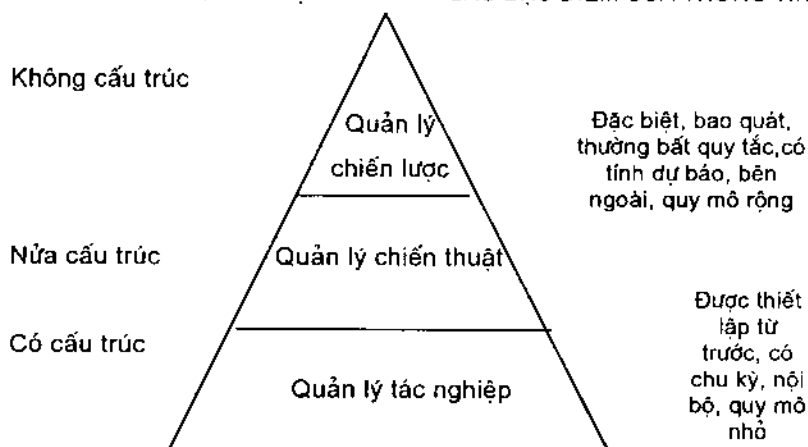
Ngược lại, vấn đề quyết định được gọi là có cấu trúc nếu nó có quy trình rõ ràng khi làm ra quyết định và các thông tin cần thiết để xem xét trong quá trình ra quyết định có thể được xác định trước.

Người ta nhận thấy các vấn đề quyết định tại mức tác nghiệp thường là quyết định có cấu trúc, còn các vấn đề quyết

định tại mức chiến thuật hoặc chiến lược thường là các quyết định nửa cấu trúc hoặc không có cấu trúc.

CẤU TRÚC CỦA QUYẾT ĐỊNH

CÁC ĐẶC ĐIỂM CỦA THÔNG TIN



Hình 27. Các mức độ của quản lý và cấu trúc của quyết định

Hình 27 minh họa mối quan hệ giữa các mức độ quản lý và cấu trúc của quyết định. Kèm theo đó là các đặc điểm của thông tin.

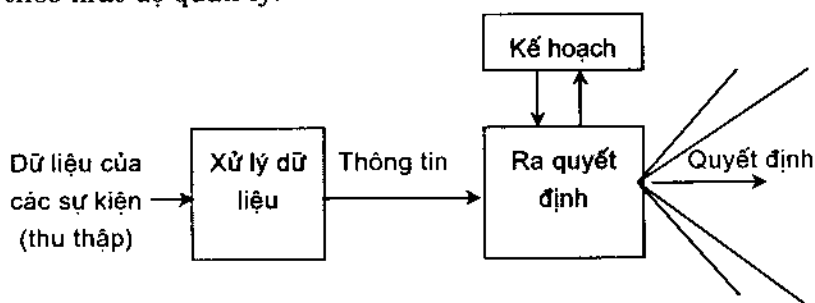
2.2. Vấn đề ra quyết định và nhu cầu thông tin

Trách nhiệm đặt ra đối với các nhà quản lý là phải tổ chức, hướng dẫn, huy động và kiểm tra các nguồn lực để đem lại lợi ích cho tổ chức. Để kiểm tra, ta cần phải có các thông tin về những gì đã xảy ra (nguyên tắc thông tin phản hồi) và phải có một kế hoạch đã chuẩn bị trước về những việc phải làm. Ta đối chiếu thông tin trên các sự kiện không bảo đảm kế hoạch và quyết định các việc phải làm để hoàn thành kế hoạch.

Để nhận được thông tin trên các sự kiện xảy ra trong quá trình quản lý, ta phải thu thập các dữ liệu, xử lý chúng để nhận được các thông tin rồi sử dụng các thông tin đó để đối chiếu với kế hoạch và khi đó một quyết định được hình thành.

Các nhà quản lý đòi hỏi phải có thông tin để hỗ trợ họ trong việc ra quyết định. Quá trình ra quyết định có thể mô tả bằng sơ đồ ở hình 28.

Nhu cầu thông tin đối với các nhà quản lý khác nhau tùy theo mức độ quản lý.



Hình 28. Quá trình ra quyết định

Các nhà quản lý chiến lược đòi hỏi thông tin có tính tổng hợp, ít chi tiết, những thông tin có tính dự báo, có quy mô rộng và thường không được xác định trước. Trong nhiều trường hợp đó là những thông tin đặc biệt, những số liệu, những biểu đồ nói lên tình trạng cơ bản và xu hướng phát triển của tổ chức. Thông tin chiến lược là sự hoà trộn của những thông tin sinh ra từ nội bộ và những thông tin nhận được từ bên ngoài.

Các nhà quản lý chiến thuật và tác nghiệp đòi hỏi thông tin có tính chi tiết, những thông tin được quy định trước, có tính định kỳ và với quy mô nhỏ. Những thông tin ở mức quản

lý này là những thông tin được sản sinh từ những nguồn bên trong, nó phản ánh tình trạng hiện tại của tổ chức. Trong nhiều trường hợp, đó là những thông tin được rút ra từ các báo cáo tác nghiệp hoặc từ việc kiểm tra hàng ngày các hoạt động của tổ chức.

Các dòng thông tin trong tổ chức

Các thông tin được thu thập từ các nguồn bên trong và bên ngoài tổ chức. Chúng có thể phân chia theo các lĩnh vực cơ bản sau:

Thông tin kinh tế, xã hội.

Thông tin về dân cư, lao động.

Thông tin về tài nguyên, đất đai.

Thông tin về khoa học, công nghệ.

Thông tin về kế hoạch, đầu tư.

Thông tin về tài chính, ngân sách.

Thông tin về công chức, viên chức nhà nước.

Thông tin về pháp luật.

Thông tin về văn hoá nghệ thuật

Thông tin về các doanh nghiệp hoạt động trên địa bàn.

Thông tin về các dự án trọng điểm trên địa bàn.

v.v...

Các thông tin này khi đưa vào hệ thống phải được trải qua quá trình tiền xử lý, phân loại của các chuyên gia có kinh nghiệm trong công việc để nâng cao giá trị sử dụng của thông

tin. Ở đây chất lượng thông tin sẽ phụ thuộc vào kiến thức chuyên gia của các chuyên viên xử lý. Các thông tin thuộc lĩnh vực thông tin khác nhau sẽ có cấu trúc dữ liệu khác nhau và do đó sẽ được tổ chức, quản lý và khai thác trên các cơ sở dữ liệu khác nhau, thông qua một cơ chế thống nhất, giúp cho người sử dụng tìm tới thông tin một cách nhanh chóng và thuận tiện.

Hệ thống các cơ sở dữ liệu chính là cái lõi thông tin của toàn bộ hệ thống tổ chức. Các thông tin này ngày càng trở nên phong phú do được bổ sung thường xuyên trong quá trình sử dụng.

3. CÁC HỆ THỐNG THÔNG TIN

3.1. Khái niệm hệ thống

Hệ thống có thể được định nghĩa như một tập hợp các phần tử có liên hệ với nhau để tạo thành một tổng thể chung. Ta có thể tìm thấy nhiều ví dụ về hệ thống trong vật lý, sinh học, trong công nghệ hiện đại và trong xã hội loài người. Ví dụ hệ thống mặt trời và các hành tinh, hệ thống cơ thể con người, hệ thống lọc dầu, hệ thống các tổ chức kinh doanh. Tuy nhiên định nghĩa này quá tổng quát. Ta dùng ở đây định nghĩa hẹp hơn, phù hợp với nhu cầu mô tả hệ thống thông tin.

Hệ thống là một tập hợp các phần tử (các thành phần) có liên hệ với nhau, hoạt động để hướng tới mục đích chung theo cách tiếp nhận các yếu tố vào, sinh ra các yếu tố ra, trong một quá trình xử lý có tổ chức.

Như vậy hệ thống có ba thành phần hay chức năng cơ bản có quan hệ tương tác với nhau:

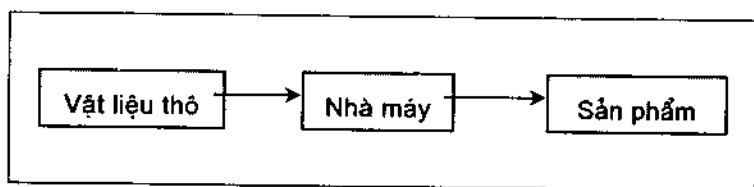
- Đầu vào: bao gồm các yếu tố được thu thập vào hệ thống để xử lý. Ví dụ: vật liệu thô, năng lượng, dữ liệu, nguồn nhân lực bảo đảm cho các quá trình xử lý.

- Xử lý: bao gồm các quá trình xử lý, chế biến để biến các yếu tố vào thành các yếu tố ra. Ví dụ các quá trình sản xuất, quá trình hô hấp của con người, quá trình xử lý dữ liệu hay các tính toán số học.

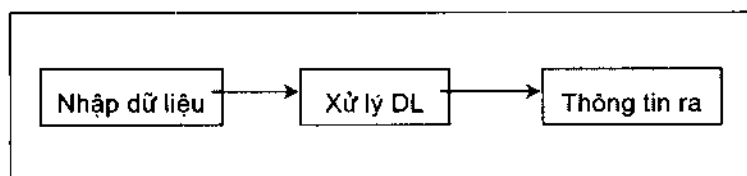
- Đầu ra: bao gồm các phần tử được tạo ra từ quá trình xử lý cho mục đích cuối cùng. Ví dụ các sản phẩm của quá trình sản xuất, các sản phẩm thông tin cần chuyển tới người dùng tin, các kết quả tính toán.

Ví dụ: Một hệ thống sản xuất tiếp nhận vật liệu thô như là đầu vào và sản xuất ra các sản phẩm cuối cùng như là đầu ra.

Một hệ thống thông tin có thể coi như là một hệ thống tiếp nhận các nguồn dữ liệu như là đầu vào và xử lý chúng thành các sản phẩm thông tin như là đầu ra.



Hình 29. Hệ thống sản xuất



Hình 30. Hệ thống thông tin

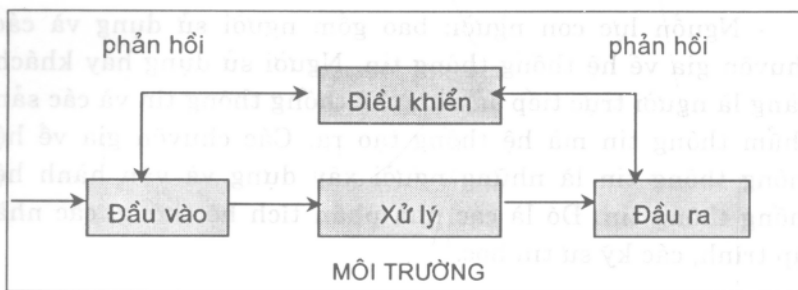
Ngoài ra người ta còn đưa thêm hai thành phần liên quan đến điều khiển hệ thống là phản hồi và điều khiển. Một hệ thống với các thành phần phản hồi và điều khiển đôi khi được gọi là hệ thống điều khiển, đó là hệ thống tự kiểm tra, tự điều chỉnh.

- Phản hồi (hay còn gọi là liên hệ ngược) là dữ liệu về sự hoàn thành nhiệm vụ của hệ thống. Ví dụ: dữ liệu về doanh thu là phản hồi cho các nhân viên bán hàng.

- Điều khiển là giám sát và đánh giá các thông tin phản hồi để xem hệ thống có hoạt động đúng hướng nhằm đạt tới mục tiêu hay không. Khi đó chức năng điều khiển phải thực hiện những điều chỉnh cần thiết đối với đầu vào và quá trình xử lý để bảo đảm đầu ra thích hợp.

Khi xem xét hệ thống người ta còn có thể đề cập đến các yếu tố và các khái niệm khác liên quan đến hệ thống như:

- Môi trường mà hệ thống tồn tại.
- Hệ thống con của hệ thống.
- Hệ thống mở, hệ thống có trao đổi với môi trường...



Hình 31. Các khối của hệ thống

3.2. Hệ thống thông tin

Hệ thống thông tin là hệ thống sử dụng nguồn lực con người và công nghệ thông tin để tiếp nhận các nguồn dữ liệu như yếu tố đầu vào và xử lý chúng thành các sản phẩm thông tin là các yếu tố đầu ra.

Công nghệ thông tin bao gồm phần cứng và phần mềm dùng để xây dựng và khai thác hệ thống thông tin. Phần cứng là các thiết bị tham gia vào quá trình xử lý thông tin như: máy tính, các phương tiện lưu trữ và truyền dữ liệu. Phần mềm là các chương trình máy tính, bao gồm các hệ điều hành, các chương trình ứng dụng và các thủ tục dành cho người sử dụng.

Xử lý thông tin bao gồm các hoạt động: tiếp nhận, truyền, xử lý, lưu trữ, tìm kiếm và hiển thị thông tin. Nhờ các hoạt động xử lý thông tin, các nguồn dữ liệu được thu thập và chế biến thành các sản phẩm thông tin cung cấp cho người sử dụng.

Như vậy bốn thành phần cơ bản cũng là bốn nguồn tài nguyên của hệ thống thông tin là:

- Nguồn lực con người: bao gồm người sử dụng và các chuyên gia về hệ thống thông tin. Người sử dụng hay khách hàng là người trực tiếp sử dụng hệ thống thông tin và các sản phẩm thông tin mà hệ thống tạo ra. Các chuyên gia về hệ thống thông tin là những người xây dựng và vận hành hệ thống thông tin. Đó là các nhà phân tích hệ thống, các nhà lập trình, các kỹ sư tin học.

- Phần cứng: bao gồm tất cả các thiết bị và phương tiện kỹ thuật dùng để xử lý thông tin. Trong đó chủ yếu là máy tính, các thiết bị ngoại vi dùng để lưu trữ và vào ra dữ liệu, mạng lưới viễn thông dùng để truyền dữ liệu.

- Phần mềm: bao gồm các chương trình máy tính: các phần mềm hệ thống, các phần mềm chuyên dụng và các thủ tục dành cho người sử dụng.

- Nguồn dữ liệu. Dữ liệu là vật liệu thô của hệ thống thông tin. Dữ liệu có thể ở nhiều dạng khác nhau, có dữ liệu văn bản, dữ liệu bằng hình ảnh, âm thanh.

Các nguồn dữ liệu của hệ thống thông tin được tổ chức thành:

- Các cơ sở dữ liệu, tổ chức và lưu giữ các dữ liệu đã được xử lý.

- Các cơ sở mô hình, lưu giữ các mô hình khái niệm, mô hình logic, mô hình toán học diễn đạt các mối quan hệ, các quy trình tính toán, các kỹ thuật phân tích.

- Các cơ sở tri thức, lưu giữ các tri thức ở các dạng khác nhau như các sự kiện, các quy tắc suy diễn về các đối tượng khác nhau.

Các hoạt động xử lý thông tin trong hệ thống thông tin bao gồm:

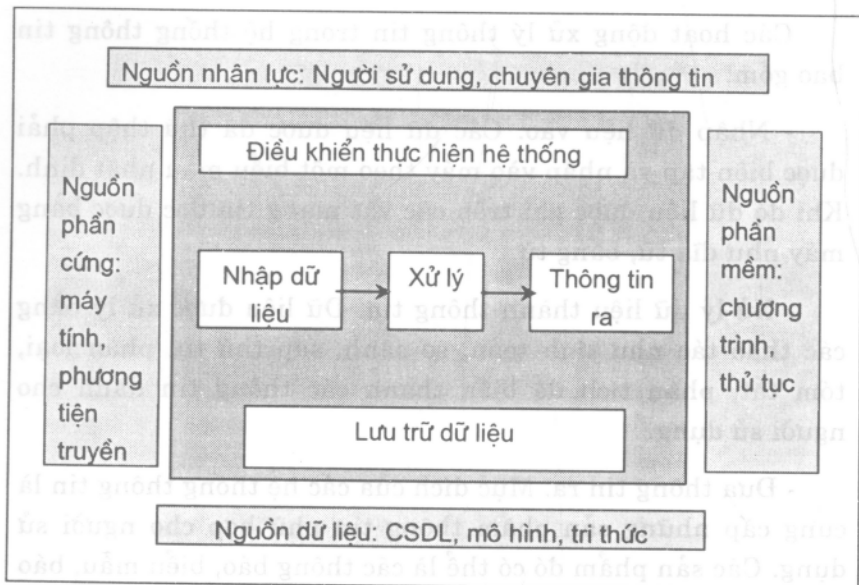
- Nhập dữ liệu vào. Các dữ liệu được đã thu thập phải được biên tập và nhập vào máy theo một biểu mẫu nhất định. Khi đó dữ liệu được ghi trên các vật mang tin đọc được bằng máy như đĩa từ, băng từ.

- Xử lý dữ liệu thành thông tin. Dữ liệu được xử lý bằng các thao tác như tính toán, so sánh, sắp thứ tự, phân loại, tóm tắt, phân tích để biến thành các thông tin dành cho người sử dụng.

- Đưa thông tin ra. Mục đích của các hệ thống thông tin là cung cấp những sản phẩm thông tin phù hợp cho người sử dụng. Các sản phẩm đó có thể là các thông báo, biểu mẫu, báo cáo, danh sách, đồ thị, hình ảnh hiển thị trên màn hình hoặc in ra trên giấy.

- Lưu trữ các nguồn dữ liệu. Lưu trữ là một hoạt động cơ bản của hệ thống thông tin, trong đó các dữ liệu và thông tin được giữ lại theo một cách tổ chức nào đó để sử dụng sau này. Các dữ liệu thường được tổ chức và lưu trữ dưới dạng các trường, các biểu ghi, các tệp và các cơ sở dữ liệu.

- Kiểm tra hoạt động của hệ thống. Hệ thống thông tin phải tạo ra các thông tin phản hồi về các quá trình vào, ra, xử lý và lưu trữ dữ liệu để có thể đánh giá và điều chỉnh hiệu quả hoạt động của hệ thống.



Hình 32. Mô hình cơ bản của hệ thống thông tin

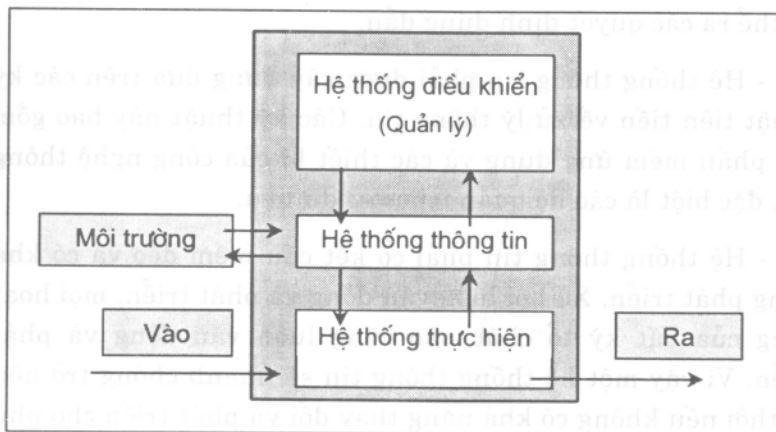
3.3. Vai trò của hệ thống thông tin

Trong bất kỳ một tổ chức nào cũng có thể xác định ba hệ thống:

- Hệ thống điều khiển: có nhiệm vụ ra các quyết định.
- Hệ thống thực hiện: hoạt động nhằm thực hiện các quyết định của hệ thống điều khiển
- Hệ thống thông tin: thực hiện sự liên hệ giữa hai hệ thống trên, bảo đảm cho tổ chức hoạt động đạt được các mục tiêu đã đặt ra.

Như vậy hệ thống thông tin là một yếu tố cấu thành của một tổ chức. Nhiệm vụ cơ bản của hệ thống thông tin trong tổ

chức là hỗ trợ cho việc ra các quyết định và giúp nhà quản lý thực hiện các chức năng của quản lý. Hệ thống thông tin cung cấp các thông tin và dữ liệu cần thiết, giúp nhà quản lý lập kế hoạch, tổ chức bố trí nhân sự với nguồn lực con người đã có, thực hiện các chức năng chỉ đạo và kiểm tra các hoạt động của tổ chức. Ngoài ra nó còn trang bị cho các nhà quản lý các phương pháp và kỹ thuật mới trong xử lý, phân tích và đánh giá. Vì vậy hệ thống thông tin trở thành một thành phần cơ bản của một tổ chức, giữ vai trò quan trọng trong thành công của hoạt động quản lý điều hành của một tổ chức.



Hình 33. Ba hệ thống trong một tổ chức

Ngày nay hệ thống thông tin trở thành một lĩnh vực nghiên cứu của khoa học quản trị và kinh doanh. Bất kỳ nhà quản lý kinh doanh nào cũng phải có những hiểu biết cơ bản về hệ thống thông tin cũng như những hiểu biết trong lĩnh vực chuyên môn của họ.

Các yêu cầu của một hệ thống thông tin

Để thực hiện được chức năng của hệ thống thông tin trong tổ chức, hệ thống thông tin phải bảo đảm các yêu cầu sau đây:

- Hệ thống thông tin phải được thiết kế xây dựng phù hợp với hệ thống tổ chức, phục vụ cho nhiệm vụ tổng thể của tổ chức và phục vụ cho các nhiệm vụ khác nhau của tổ chức.

- Việc xây dựng hệ thống thông tin phải đạt mục đích là hỗ trợ cho việc ra các quyết định, bảo đảm cung cấp thông tin đầy đủ, chính xác, kịp thời cho các cấp lãnh đạo quản lý để họ có thể ra các quyết định đúng đắn.

- Hệ thống thông tin phải được xây dựng dựa trên các kỹ thuật tiên tiến về xử lý thông tin. Các kỹ thuật này bao gồm các phần mềm ứng dụng và các thiết bị của công nghệ thông tin, đặc biệt là các hệ quản trị cơ sở dữ liệu.

- Hệ thống thông tin phải có kết cấu mềm dẻo và có khả năng phát triển. Xã hội luôn vận động và phát triển, mọi hoạt động của bất kỳ tổ chức nào cũng luôn vận động và phát triển. Vì vậy một hệ thống thông tin sẽ nhanh chóng trở nên lỗi thời nếu không có khả năng thay đổi và phát triển cho phù hợp với sự biến đổi và phát triển của thực tế.

4. MỘT SỐ HỆ THỐNG THÔNG TIN

Vai trò của hệ thống thông tin trong tổ chức không ngừng được mở rộng trong mấy chục năm gần đây.

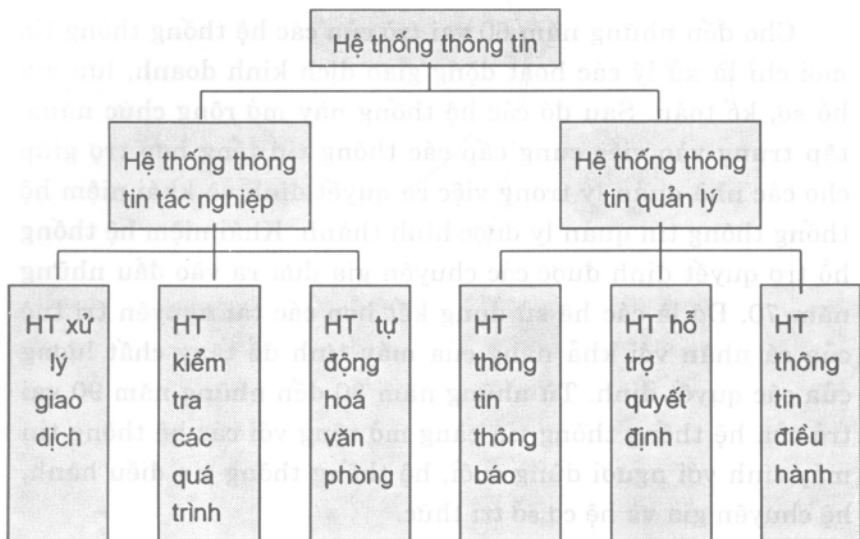
Cho đến những năm 60 vai trò của các hệ thống thông tin mới chỉ là xử lý các hoạt động giao dịch kinh doanh, lưu giữ hồ sơ, kế toán. Sau đó các hệ thống này mở rộng chức năng, tập trung vào việc cung cấp các thông tin tổng hợp trợ giúp cho các nhà quản lý trong việc ra quyết định và khái niệm hệ thống thông tin quản lý được hình thành. Khái niệm hệ thống hỗ trợ quyết định được các chuyên gia đưa ra vào đầu những năm 70. Đó là các hệ sử dụng kết hợp các tài nguyên trí tuệ của cá nhân với khả năng của máy tính để tăng chất lượng của các quyết định. Từ những năm 80 đến những năm 90 vai trò của hệ thống thông tin càng mở rộng với các hệ thông tin máy tính với người dùng cuối, hệ thống thông tin điều hành, hệ chuyên gia và hệ cơ sở tri thức.

Các hệ thống thông tin đóng vai trò quan trọng trong việc hỗ trợ cho các hoạt động tác nghiệp và quản lý trong các doanh nghiệp và các tổ chức khác. Do đó các hệ thống thông tin có thể chia thành hai loại là hệ thống thông tin tác nghiệp và hệ thống thông tin quản lý.

Hệ thống thông tin tác nghiệp lại chia thành: Hệ thống xử lý giao dịch, hệ thống kiểm tra các quá trình, hệ thống tự động hoá văn phòng.

Trong hệ thống thông tin quản lý lại có: Hệ thống thông tin thông báo, hệ thống hỗ trợ quyết định, hệ thống thông tin điều hành.

Hình dưới đây minh hoạ sự phân loại trên của các hệ thống thông tin:



Hình 34. Phân loại các hệ thống thông tin

4.1. Các hệ thống thông tin tác nghiệp

Hệ thống thông tin tác nghiệp (OIS - Operations Information Systems) gắn liền với việc xử lý các hoạt động tác nghiệp của một chức năng nghiệp vụ xác định. Nó nhằm nâng cao hiệu quả hoạt động có tính thường xuyên trong một tổ chức.

Hệ thống thông tin tác nghiệp chịu trách nhiệm xử lý các hoạt động giao dịch, công việc điều khiển các quá trình hoặc các hệ thống tự động hoá văn phòng.

Hệ thống xử lý tác nghiệp có các đặc trưng sau:

- Khối lượng công việc giao dịch nhiều.

- Các quy trình để xử lý giao dịch là rõ ràng, chặt chẽ, có thể mô tả một cách chi tiết.

- Ít có trường hợp ngoại lệ.

Do quy trình hoạt động của các hệ thống xử lý tác nghiệp có quy chuẩn rõ ràng nên có thể xây dựng các quy trình xử lý, trong đó xác định rõ máy tính có thể tham gia vào những khâu nào của quy trình này. Các chương trình máy tính phải được xây dựng để:

- Cập nhật, kiểm tra và quản lý các dữ liệu đưa vào máy tính.

- Thực hiện các xử lý dữ liệu cần thiết theo quy trình.

- Lưu trữ dữ liệu.

- Hiển thị các thông tin cần thiết theo yêu cầu.

Đầu ra định kỳ của hệ thống thường là các báo cáo đã được định dạng sẵn.

Hệ thống xử lý tác nghiệp là không linh hoạt, không thể điều tiết được việc xử lý dữ liệu hay nhu cầu thông tin khi chưa được xây dựng trước trong hệ thống.

Hệ thống xử lý giao dịch (TPS - Transaction Processing Systems) là ví dụ tiêu biểu về hệ thống thông tin tác nghiệp. Đó là hệ thống thông tin xử lý các dữ liệu thu được từ các việc xảy ra hàng ngày trong các hoạt động giao dịch của một tổ chức. Đó là các hoạt động như mua vào, bán ra, gửi tiền ở ngân hàng, rút tiền ra, trả tiền, thanh toán v.v...

Hệ thống xử lý giao dịch hỗ trợ hoạt động từng ngày của một tổ chức bằng cách duy trì các ghi chép chi tiết về các kết

quả của hoạt động giao dịch trong ngày. Hệ thống này giúp đơn vị nắm chắc tình hình để chỉ đạo các hoạt động của nó. Vì ở vị trí trung tâm của các hoạt động của một đơn vị nên TPS phải thực hiện công việc của mình một cách nhanh chóng và chắc chắn.

Có thể kể một vài hệ thống loại này như:

- Hệ thống thanh toán tài vụ.
- Hệ thống gửi tiền qua bưu điện.
- Hệ thống thanh toán ngân hàng.
- Hệ thống xử lý hoạt động giao dịch bán hàng.
- Hệ thống các dịch vụ bảo hiểm.
- Hệ thống dịch vụ vận tải hàng hoá.
- Hệ thống dịch vụ bán vé máy bay, bán vé tàu hoả.
- Hệ thống dịch vụ cho mượn ở thư viện v.v...

Phần lớn các hệ thống này đều hợp tác chặt chẽ với các mạng lưới thông tin viễn thông.

Hệ thống xử lý giao dịch thu thập và xử lý các dữ liệu về các hoạt động giao dịch, cập nhật vào CSDL và tạo ra các sản phẩm thông tin để sử dụng ở trong và ngoài hệ thống. Quá trình xử lý là một chu trình bao gồm năm bước:

- (1) Nhập dữ liệu.
- (2) Thực hiện các hoạt động giao dịch.
- (3) Xử lý các tệp dữ liệu và các CSDL.

(4) Tạo ra các tài liệu và báo cáo.

(5) Tiến hành các hoạt động thẩm tra.

Các quy trình xử lý giao dịch thường được gọi là các quy trình hoạt động chuẩn. Phần lớn quy trình giao dịch chuẩn gắn liền với mức hoạt động hàng ngày của tổ chức. Các quy trình này được mô phỏng bằng các chương trình trên máy tính để kiểm tra, quản lý các dữ liệu đưa vào máy tính, xử lý các chi tiết, lưu trữ và hiển thị các thông tin cần thiết đáp ứng các nhu cầu.

Ví dụ, các quy trình xử lý các hoạt động giao dịch ở ngân hàng là các quy trình chuẩn để xử lý các dịch vụ gửi và rút tiền, thanh toán séc và một số dịch vụ khác. Từ quy trình chuẩn người ta xây dựng chương trình máy tính quản lý việc gửi, rút tiền tự động. Hệ thống gửi, rút tiền tự động là một TPS gồm nhiều thiết bị đầu cuối đặt ở những nơi cần thiết, cho phép khách hàng thực hiện dịch vụ gửi và rút tiền theo tài khoản mà không cần giao dịch trực tiếp tại văn phòng của ngân hàng.

Hệ thống xử lý giao dịch bán hàng cũng là một ví dụ về TPS. Hệ thống này xử lý các hàng hoá bán ra, mua vào và hàng hoá tồn kho. Kết quả của các quá trình này được cập nhật trong các CSDL về khách hàng, CSDL về hàng bán ra và CSDL về hàng tồn kho. Hệ thống này còn sản sinh ra các sản phẩm thông tin khác nhau dành cho các sử dụng nội bộ và bên ngoài hệ thống như các đánh giá về khách hàng, về tài chính, doanh thu, kiểm tra lãi suất, các dạng thuế, các đơn đặt hàng...

Hệ thống điều khiển các quá trình (PCS - Process Control Systems) là hệ thống sử dụng máy tính điện tử để ra các

quyết định điều chỉnh các quá trình sản xuất một cách tự động. Ví dụ: các hệ thống lọc dầu, các dây chuyền lắp ráp tự động của các nhà máy thường dùng các hệ thống này. Ở đây hệ thống kiểm tra các quá trình vật lý, thu thập và xử lý các dữ liệu được phát hiện bởi các biến cảm và thực hiện các điều chỉnh cần thiết đối với quá trình.

Hệ tự động hoá hoạt động văn phòng (OAS - Office Automation Systems) cũng là một hệ thống xử lý tác nghiệp. Ở đây máy tính được sử dụng để thực hiện các chức năng của hoạt động văn phòng như:

- *Xử lý văn bản.* Người quản lý sử dụng các phần mềm soạn thảo, xử lý văn bản để biên soạn, in ấn các tài liệu văn bản như thư từ, công văn, báo cáo, bảng biểu. Đó là ứng dụng phổ biến nhất của tin học văn phòng.

- *Thư tín điện tử.* Dùng thư tín điện tử có thể liên lạc với nhiều nơi trên trái đất, nhanh và rẻ hơn thư bưu điện. Qua dịch vụ này, không những văn bản mà các hình ảnh, âm thanh... cũng được truyền.

- *Bảng tính điện tử.* Sử dụng các bảng tính điện tử để lập các bảng biểu thống kê, tính toán và quản trị dữ liệu. Các dữ liệu được thể hiện không chỉ trong các bản dữ liệu mà còn dưới dạng biểu đồ, bảng biểu liên quan đến dữ liệu đó.

- *Facsimile (Fax).* Dùng một máy fax bạn có thể truyền cả một văn bản in trên giấy đến máy fax ở nơi nhận thông qua tuyến điện thoại hoặc các tuyến truyền thông khác. Đó chính là một phương thức truyền thông cho phép sao chép tài liệu từ xa.

Ngoài các ứng dụng trên nhiều đơn vị còn sử dụng các *phần mềm quản lý văn phòng*, đáp ứng các yêu cầu thông dụng của các cơ quan quản lý hành chính như: quản lý lịch công tác của lãnh đạo và các bộ phận, quản lý các công việc thường xuyên hoặc bất thường của cơ quan, lập danh mục các thông báo nội bộ, các thư điện tử và các công văn đi/đến, v.v...

Bộ chương trình tin học văn phòng nổi tiếng đang được sử dụng rộng rãi ở nước ta hiện nay là Microsoft Office 97, hoặc Microsoft Office 2000 bao gồm hai chương trình cơ bản:

- Hệ soạn thảo văn bản WinWord.
- Hệ xử lý bảng tính điện tử Excel.

4.2. Các hệ thống thông tin quản lý

Hệ thống thông tin quản lý (MIS - Management Information Systems) có mục đích cung cấp thông tin trợ giúp các nhà lãnh đạo, các nhà quản lý trong việc ra quyết định và quản lý công việc trong tổ chức.

Cung cấp thông tin và hỗ trợ cho việc ra quyết định ở tất cả các cấp độ quản lý, từ quản lý chiến lược, quản lý chiến thuật đến quản lý tác nghiệp là nhiệm vụ của hệ thống thông tin quản lý.

Hệ thống thông tin quản lý bao gồm các cơ sở dữ liệu, các luồng thông tin và được quy định các chức năng để thực hiện mục tiêu chung. Hệ thống này hỗ trợ nhiều chức năng xử lý dữ liệu trong giao dịch và lưu trữ, thích ứng được với những thay đổi của quy trình xử lý thông tin, cung cấp đầy đủ thông tin để các nhà quản lý sử dụng trong quá trình ra quyết định và điều hành hoạt động của tổ chức.

So với hệ thống thông tin tác nghiệp, hệ thống thông tin quản lý mềm dẻo hơn, có nhiều chức năng xử lý dữ liệu hơn.

Có ba loại hệ thống thông tin quản lý chính là: Hệ thống thông tin thông báo, hệ thống hỗ trợ quyết định và hệ thống thông tin điều hành.

4.2.1. Hệ thống thông tin báo cáo - IRS

Hệ thống thông tin báo cáo (IRS - Information Reporting Systems) là dạng chung nhất của hệ thống thông tin quản lý. Nó cung cấp cho nhà quản lý các sản phẩm thông tin hỗ trợ việc ra quyết định hàng ngày của họ. Nội dung của các sản phẩm thông tin này đã được nhà quản lý chỉ ra từ trước. Đó là những báo cáo cung cấp thông tin mà họ cần.

Báo cáo là một hình thức thông tin văn bản trình bày về một vấn đề, một sự việc, tình hình hoặc hiện trạng... với một đối tượng nhất định trong tổ chức. Đó là một hình thức thông tin chủ yếu trong các hình thức thông tin phục vụ cho công tác chỉ đạo và quản lý điều hành. Báo cáo cũng là một khâu quan trọng trong hệ thống giao tiếp của tổ chức, là nội dung liên hệ giữa cấp dưới và cấp trên, giữa cơ quan và các thành viên trong tổ chức. Báo cáo đầy đủ, chính xác và kịp thời sẽ giúp cấp trên nắm được thông tin của cấp dưới để có thể hoạch định chủ chương, ra quyết định được chuẩn xác.

Báo cáo được phân thành nhiều loại, có báo cáo định kỳ, báo cáo bất thường, báo cáo nhanh, báo cáo chuyên đề, báo cáo hội nghị, báo cáo sơ kết, tổng kết, tổng hợp, Các báo cáo này thường phải tuân theo một hình thức và thủ tục nhất định.

Các hệ thống thông tin báo cáo tìm các thông tin về các hoạt động nội bộ từ các CSDL được cập nhật bởi hệ thống xử lý các hoạt động giao dịch. Chúng cũng có thể nhận dữ liệu về môi trường xung quanh từ các nguồn bên ngoài.

Hệ thống phải cung cấp cho nhà quản lý những thông tin theo yêu cầu, những thông tin mang tính định kỳ hoặc theo một danh mục đã xác định từ trước. Ngoài ra còn cung cấp những thông tin đặc biệt mang tính ngoại lệ, những bản báo cáo theo yêu cầu và những câu trả lời tức thì cho những câu hỏi. Ví dụ, người quản lý bán hàng có thể nhận được: câu trả lời tức thời về tình hình bán một sản phẩm nào đó, các báo cáo hàng tuần đánh giá các kết quả bán được của một nhân viên hay một cửa hàng.

Các chương trình ứng dụng và các phần mềm quản trị CSDL của IRS sẽ cho phép nhà quản lý tiếp cận tới các CSDL hợp thành của tổ chức và cả những CSDL bên ngoài khi cần thiết.

4.2.2. Hệ thống thông tin hỗ trợ quyết định (DSS)

Hệ thống thông tin hỗ trợ quyết định (DSS- Decision Support Systems) được xác định như hệ thống dựa trên sự tương tác với máy tính, giúp nhà quản lý sử dụng các mô hình và dữ liệu trong các CSDL chuyên ngành để hỗ trợ cho việc ra quyết định của họ.

Hệ thống thông tin hỗ trợ quyết định giúp người quản lý tìm được tiến trình hành động tốt nhất để trả lời các câu hỏi dạng "Nếu ... thì ..." do hệ thống đưa ra.

DSS làm đơn giản quá trình ra quyết định chứ không trực tiếp ra các quyết định. Đó là hệ hỗ trợ việc ra các quyết định,

vì nó sử dụng các công cụ, mô hình, dữ liệu và các tài nguyên giúp nhà quản lý hiểu, phân tích, đánh giá và giải quyết vấn đề đang đặt ra. Các hệ hỗ trợ quyết định sử dụng kết hợp trí tuệ của các cá nhân với khả năng của máy tính để nâng cao chất lượng của các quyết định.

Thành phần chủ yếu của DSS là:

- Phần cứng, bao gồm các trạm làm việc bằng máy tính. Nó có thể sử dụng như một cơ sở độc lập, nhưng thường được liên kết trong một mạng máy tính để có thể khai thác các nguồn dữ liệu và mô hình của các hệ DSS khác.

- Phần mềm, bao gồm các hệ quản trị CSDL, hệ quản trị các mô hình và giao diện người sử dụng.

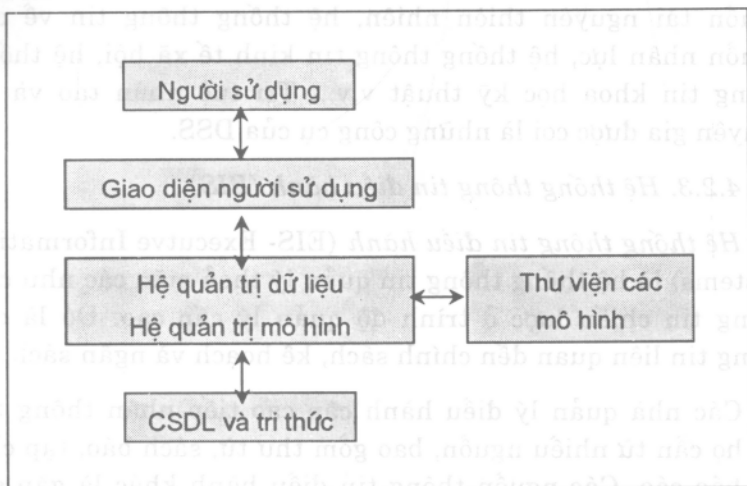
- Cơ sở dữ liệu và tri thức, chứa các thông tin và dữ liệu rút ra từ các CSDL của tổ chức, các CSDL bên ngoài và CSDL của cá nhân nhà quản lý. Nó cũng chứa các dữ liệu và thông tin tổng hợp cần thiết cho nhà quản lý.

- Cơ sở mô hình, bao gồm một thư viện các mô hình toán học và các kỹ thuật phân tích được lưu trữ dưới dạng các chương trình.

Trong DSS, giao diện thân thiện và thông minh với người sử dụng là yêu cầu tối thiểu. Không những người ta dùng các bảng để biểu diễn dữ liệu mà dùng cả các đồ họa và thông tin đa phương tiện. Ngày nay, những thiết bị tin học cho phép thực hiện liên hệ người - máy đẹp và hiệu quả.

DSS thường được xây dựng dựa trên hệ thống thông tin tác nghiệp và hệ thống thông tin thông báo tồn tại trong tổ chức. Dữ liệu trong DSS thường là những thông tin mang

tính định hướng, quan tâm đến tổng thể hơn là các sự kiện đơn lẻ.



Hình 35. Các khối chính của hệ hỗ trợ quyết định

Hệ thống thông tin hỗ trợ quyết định cung cấp cho nhà quản lý các mô hình phân tích để mô phỏng các vấn đề trong thực tiễn, cách tìm kiếm dữ liệu, các khả năng biểu diễn thông tin và thường nhấn mạnh việc kết xuất thông tin bằng hình ảnh. Từ đó các nhà quản lý sử dụng khả năng tương tác người - máy của hệ thống để tạo ra thông tin mà họ cần để ra các quyết định, khi mà vấn đề đặt ra là không có cấu trúc hoặc nửa cấu trúc.

DSS không tự làm quyết định cụ thể giúp con người trong công tác quản lý mà chỉ hỗ trợ việc tính toán các phương án để nhà quản lý lựa chọn và đưa ra quyết định cuối cùng. Các phương án do DSS đưa ra ứng với các tình huống mà nhà quản lý đặt ra trong các câu hỏi dạng “nếu - thì”.

Ví dụ về DSS là: các hệ thống thông tin kế hoạch hoá quốc gia và các vùng lãnh thổ, hệ thống thông tin về các nguồn tài nguyên thiên nhiên, hệ thống thông tin về các nguồn nhân lực, hệ thống thông tin kinh tế xã hội, hệ thống thông tin khoa học kỹ thuật v.v... Trí tuệ nhân tạo và hệ chuyên gia được coi là những công cụ của DSS.

4.2.3. Hệ thống thông tin điều hành (EIS)

Hệ thống thông tin điều hành (EIS- Executve Information Systems) là hệ thống thông tin quản lý thoả mãn các nhu cầu thông tin chiến lược ở trình độ quản lý cấp cao. Đó là các thông tin liên quan đến chính sách, kế hoạch và ngân sách.

Các nhà quản lý điều hành cấp cao tiếp nhận thông tin mà họ cần từ nhiều nguồn, bao gồm thư từ, sách báo, tạp chí, các báo cáo. Các nguồn thông tin điều hành khác là gặp gỡ, trao đổi và các hoạt động xã hội. Như vậy nhiều thông tin của hoạt động điều hành cấp cao xuất phát từ những nguồn không phải do máy tính cung cấp.

Hệ thống thông tin điều hành dựa trên máy tính là hệ thống kết hợp nhiều nét của hệ thống thông tin thông báo và hệ hỗ trợ quyết định. Tuy nhiên mục đích chính của EIS là cung cấp một cách nhanh chóng và thuận lợi cho các nhà quản lý cấp cao các thông tin có chọn lọc về các yếu tố mang tính giải pháp, giúp họ hoàn thành mục tiêu chiến lược của tổ chức.

EIS hoạt động dựa vào các phần mềm quản trị CSDL và phần mềm quản lý mạng viễn thông, cho phép tiếp cận nhanh chóng và thuận lợi các CSDL bên trong, các CSDL bên ngoài và các CSDL quản lý đặc thù của hệ thống.

Như vậy EIS gồm các thành phần sau:

- Phần cứng: một mạng máy tính bao gồm các trạm điều hành (executive workstations)
- Phần mềm: phần mềm EIS, phần mềm quản trị CSDL, phần mềm quản lý mạng truyền thông.
- Các nguồn dữ liệu: Các CSDL tác nghiệp nội bộ, các CSDL bên ngoài, các CSDL quản lý đặc thù.

EIS cung cấp những thông tin về hiện trạng của tổ chức và có những công cụ giúp người sử dụng có thể tìm đến những thông tin chi tiết hơn. Vì vậy EIS thường sử dụng công nghệ siêu văn bản (hypertext technology) cho phép người sử dụng dễ dàng tìm đến những thông tin liên quan từ một tệp dữ liệu dữ liệu đã có. Ngoài khả năng tìm các thông tin chi tiết, EIS còn cung cấp các công cụ phân tích định hướng và các báo cáo ngoại lệ nhằm giúp nhà quản lý nhanh chóng phát hiện ra các yếu tố chủ đạo chi phối hoạt động của tổ chức.

EIS sử dụng giao diện đồ hoạ với người sử dụng. Do đó EIS dễ vận hành và dễ hiểu. Trong EIS các biểu diễn đồ thị được sử dụng rộng rãi để trình bày các thông tin. EIS cung cấp nhanh các thông tin chi tiết dựa trên các văn bản, số liệu hay đồ thị về hiện trạng của tổ chức và định hướng cho những giải pháp của các hoạt động điều hành ở cấp cao.

EIS có các đặc trưng cơ bản sau:

- Giao diện đồ hoạ đơn giản, dễ hiểu và dễ sử dụng
- Có khả năng tiếp cận trên diện rộng các nguồn dữ liệu bên trong và bên ngoài tổ chức.
- Có khả năng tìm đến những thông tin chi tiết hơn.

- Các thông tin được biểu diễn dưới dạng đồ thị, bảng biểu hay văn bản.

- Có các công cụ phân tích thống kê.

- Có khả năng giải quyết nhiều bài toán trên những công việc khác nhau.

- Được sử dụng trực tiếp bởi người quản lý điều hành không cần qua trung gian.

EIS trở nên quen thuộc trong những năm gần đây và ngày càng được sử dụng rộng rãi trong các hoạt động quản lý trung hạn.

Vai trò con người trong hệ thống thông tin quản lý

Một hệ thống thông tin quản lý không thể thiếu yếu tố con người. Đó là những cán bộ chuyên môn trên lĩnh vực tin học cùng với nhà quản lý có trách nhiệm phân tích, thiết kế và vận hành hệ thống thông tin quản lý. Trong đó người phân tích thiết kế hệ thống và người quản lý có vai trò quan trọng nhất.

Người phân tích hệ thống là những người được đào tạo cơ bản, có kiến thức về hệ thống, về kinh tế xã hội và quản lý. Nhiệm vụ của họ là nghiên cứu các ứng dụng và thiết kế xây dựng hệ thống thông tin. Họ là cầu nối giữa tin học và quản lý.

Người quản lý có trách nhiệm phối hợp và hỗ trợ cho người phân tích hệ thống trong việc thiết kế hệ thống thông tin, sao cho phù hợp với cơ chế quản lý của tổ chức. Người quản lý cũng cần được đào tạo. Ngoài kiến thức về chuyên môn, về quản lý, họ còn cần phải nắm được công cụ tin học để

có khả năng sử dụng và khai thác hệ thống thông tin trong công việc thường ngày.

4.3. Các hệ chuyên gia

Khi một tổ chức hay cá nhân cần giải quyết một vấn đề phức tạp, họ thường hay tìm đến các chuyên gia. Các chuyên gia này có tri thức chuyên sâu và có kinh nghiệm về lĩnh vực hẹp. Đối với nhà quản lý, có thể xây dựng một hệ thống thông tin có thể cung cấp thông tin giúp đỡ họ giải quyết vấn đề giống như tham khảo ý kiến các chuyên gia.

Một hệ chuyên gia (ES - Expert System) là hệ thống thông tin dựa trên máy tính, sử dụng các tri thức về một lĩnh vực xác định để trợ giúp cho người sử dụng như kiểu trợ giúp của một chuyên gia.

Sở dĩ gọi là hệ chuyên gia vì cơ sở dữ liệu của nó lưu trữ mô tả các kỹ năng ra quyết định của người chuyên gia trên một số lĩnh vực chuyên môn hẹp như chẩn đoán bệnh qua kết quả chiếu điện, cách tính thuế, sửa chữa các thiết bị vận hành kém hiệu quả, thực hiện quá trình lên men bia,...

Các thành phần của một hệ chuyên gia bao gồm:

Cơ sở tri thức: bao gồm các sự kiện về lĩnh vực xác định và các phép suy lý của các chuyên gia về các sự kiện đó, gọi là luật suy diễn. Luật suy diễn được xây dựng dựa trên cơ sở các phép toán của logic mệnh đề, bao gồm: phép hội, phép tuyển, phép phủ định, phép kéo theo... Ví dụ: Nếu nước ở thể rắn và trong suốt và nếu áp suất là 760 mm thủy ngân thì nhiệt độ bằng hay nhỏ hơn 0°. Cơ sở tri thức do các chuyên gia xây dựng và được lưu trữ trong bộ nhớ của máy tính.

Các phần mềm: bao gồm một chương trình xử lý tri thức (sự kiện và luật) liên quan đến một vấn đề xác định, gọi là mô tơ suy diễn và chương trình giao diện dùng để liên lạc với người sử dụng, bao gồm cả việc giải thích cách thức suy diễn nếu họ yêu cầu. Mô tơ suy diễn chính là một hệ thống tình chế, lựa chọn, vận hành, kiểm tra các luật suy diễn lấy ra từ cơ sở tri thức để xây dựng các suy luận logic phù hợp, nó không phụ thuộc vào lĩnh vực chuyên môn mà phụ thuộc vào cơ sở tri thức.

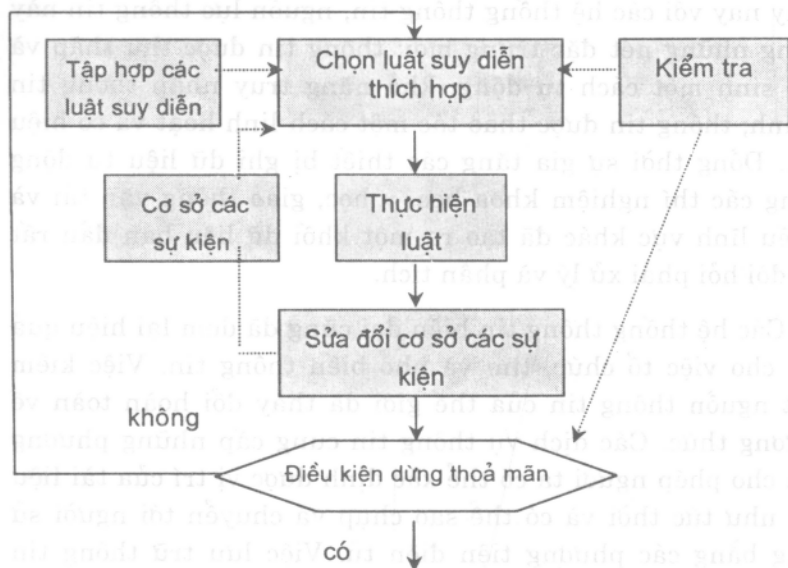
Phần cứng: bao gồm các trạm làm việc bằng máy tính, các thiết bị đầu cuối nối với máy chủ trong một mạng viễn thông.

Hệ chuyên gia tình chế tri thức và có cơ chế suy diễn, cho phép suy luận trên các tri thức và cung cấp câu trả lời cho người sử dụng. Nó cũng cho người ta biết cách thức mà nó đã suy luận.

Nguyên tắc làm việc của hệ chuyên gia là: Xuất phát từ một vấn đề được mô tả bởi một tập hợp các sự kiện được ghi trong cơ sở tri thức, nhờ sự hỗ trợ của mô tả suy diễn, hệ thống sẽ xây dựng một suy luận logic trên các sự kiện để rút ra câu trả lời cho người sử dụng, bằng cách sử dụng các luật suy diễn được cung cấp bởi các chuyên gia và được ghi trong cơ sở tri thức.

Hệ chuyên gia là một lớp chương trình máy tính đa dạng đã được thương mại hoá. Tri thức trong hệ chuyên gia là không có cấu trúc (tức là quá trình giải quyết vấn đề đặt ra không được rõ ràng) và nó được phát biểu trong các mối quan hệ hay suy luận từ một chuỗi các mệnh đề. Vì thế mỗi điều

kiện có thể xảy ra phải được mô tả bằng các luật và hệ chuyên gia không thể xử lý được các sự kiện không dự kiến trước mà chỉ giới hạn giải quyết trong một phạm vi hẹp. Sự vận hành của hệ chuyên gia được mô tả bằng sơ đồ dưới đây:



Hình 36. Sự vận hành của một hệ chuyên gia

Hệ chuyên gia là một ngành ứng dụng của trí tuệ nhân tạo (AI - Artificial Intelligence). Nó được ứng dụng trong nhiều lĩnh vực khác nhau như y học, kỹ thuật, vật lý và hoạt động kinh doanh. Hệ chuyên gia có thể giúp cho việc chẩn đoán bệnh, khai thác mỏ, phân tích các hợp chất,... Ngày nay hệ chuyên gia còn vươn tới các bài toán phức tạp như lập kế hoạch tài chính, quản lý tín dụng, tư vấn thuế, đánh giá điều khiển nội bộ, quản lý sản xuất, phân tích đúng sai.

5. VAI TRÒ CỦA HỆ THỐNG THÔNG TIN TRONG XÃ HỘI HIỆN ĐẠI

Sự quan tâm đối với thông tin và tri thức như một nguồn lực của mỗi cá nhân, tổ chức và xã hội ngày nay tăng nhanh gấp bội so với bất cứ thời kỳ nào trong lịch sử. Tuy nhiên ngày nay với các hệ thống thông tin, nguồn lực thông tin này mang những nét đặc trưng mới: thông tin được thu thập và sản sinh một cách tự động; khả năng truy nhập thông tin nhanh; thông tin được thao tác một cách linh hoạt và có hiệu quả. Đồng thời sự gia tăng các thiết bị ghi dữ liệu tự động trong các thí nghiệm khoa học, y học, giao thông vận tải và nhiều lĩnh vực khác đã tạo ra một khối dữ liệu ban đầu rất lớn đòi hỏi phải xử lý và phân tích.

Các hệ thống thông tin hiện đại cũng đã đem lại hiệu quả mới cho việc tổ chức, tìm và phổ biến thông tin. Việc kiểm soát nguồn thông tin của thế giới đã thay đổi hoàn toàn về phương thức. Các dịch vụ thông tin cung cấp những phương tiện cho phép người ta có thể xác định được vị trí của tài liệu gần như tức thời và có thể sao chụp và chuyển tới người sử dụng bằng các phương tiện điện tử. Việc lưu trữ thông tin bằng kỹ thuật số cho phép cá nhân có thể sở hữu những bộ sưu tập tài liệu tương đương với cả một thư viện. Việc truy nhập các nguồn thông tin trên mạng cho phép các đơn vị và cá nhân tích lũy và chia sẻ nguồn lực thông tin.

Dưới tác động của sự phát triển công nghệ, một số khái niệm cơ bản của lĩnh vực thông tin cũng thay đổi. Chẳng hạn, công nghệ thông tin đã mở rộng khái niệm truyền thống về tài liệu. Ngày nay tài liệu không chỉ là các đối tượng vật chất ghi cố định các thông tin mà bao gồm cả các vật mang tin đa

phương tiện. Nhờ được ghi dưới dạng số hoá, các đối tượng này rất dễ thao tác; chúng có thể chia nhỏ, kết hợp với các phần khác, thay đổi dạng thức, hiển thị dưới nhiều hình thức trên nhiều thiết bị khác nhau.

Hiệu quả của hệ thống thông tin hiện đại là ở khả năng xử lý thông tin tự động của nó. Khả năng này có được là do thông tin được biểu diễn dưới dạng số hoá. Các hệ thống thông tin dựa trên máy tính điện tử có thể thực hiện các phép tính toán, phân tích, phân loại, thiết lập các mối tương quan với mức độ phức tạp và hiệu quả vượt xa khả năng của con người. Chúng có thể mô phỏng việc thực hiện các mô hình logic và toán học của các quá trình vật lý và trạng thái dưới những điều kiện khác nhau. Các hệ thống này cũng bắt đầu bắt chước được quá trình nhận thức của con người: rút ra kết luận bằng suy diễn trong hệ chuyên gia, phân tích ngữ cảnh trong xử lý ngôn ngữ tự nhiên, suy luận tương tự và trực giác trong tìm kiếm thông tin. Khả năng to lớn của công nghệ biến đổi thông tin ngày nay - từ dữ liệu văn bản sang đồ thị, từ tiếng nói sang chữ viết, từ ngôn ngữ này sang ngôn ngữ khác - đã mở rộng giá trị sử dụng của thông tin và nâng cao khả năng giải quyết các bài toán thực tiễn của con người. Trong nhiều trường hợp việc sử dụng các mạng lưới thông tin máy tính đa phương tiện có thể thay thế cho báo chí, cạnh tranh với truyền hình và cung cấp một chiều mới trong tương lai cho giáo dục và đào tạo ở mức độ toàn xã hội.

Việc ứng dụng các hệ thống thông tin được thúc đẩy bởi sự đòi hỏi gia tăng các chức năng xử lý thông tin mang tính trí tuệ của con người. Hiệu quả của nó thể hiện nổi bật trên ba lĩnh vực sau: kinh tế, quản lý xã hội và môi trường tồn tại của mỗi cá nhân.

Hệ thống thông tin là công cụ chủ yếu để nâng cao hiệu quả của đầu tư xã hội. Trong lĩnh vực kinh tế, chúng có thể thúc đẩy nâng cao sức sản xuất, nhất là trong khu vực công nghiệp và dịch vụ, thông qua việc tự động hoá các quá trình chế tạo, thông qua sử dụng hệ thống thông tin để hỗ trợ việc ra quyết định, giải quyết các bài toán thực tiễn và hỗ trợ các công việc văn phòng.

Ngày nay hệ thống thông tin trở thành bộ phận hữu cơ của bất cứ hệ thống tổ chức xã hội nào. Nó đem lại hiệu quả cho hoạt động quản lý xã hội, đồng thời cũng tác động tới cấu trúc và phong cách quản lý của các hệ thống tổ chức xã hội.

Các hệ thống thông tin tác động rõ ràng tới trình độ sống và phong cách sống của mỗi cá nhân. Các hệ thống thông tin mở rộng khả năng tiếp thu tri thức và hưởng thụ văn hoá, nâng cao chất lượng dịch vụ xã hội đối với mọi người. Sự gia tăng mạnh mẽ khối lượng tri thức mới cũng như tốc độ lỗi thời của tri thức đã sử dụng buộc con người phải thường xuyên tiếp xúc với các nguồn thông tin và các hệ thống thông tin. Việc sử dụng có hiệu quả các nguồn thông tin và các hệ thống thông tin phải trở thành thói quen, tập quán của con người trong xã hội hiện đại.

TÀI LIỆU THAM KHẢO

1. Đoàn Phan Tân. *Cơ sở Thông tin học*. Trường Đại học Văn hoá Hà Nội, 1990.
2. Claire Guinchat et Michel Menou. *Introduction aux sciences et techniques de l'information et de la documantation*. Les Presses de l' Unesco, 1981.
3. *Manuel de bibliothécaire-documentalistes travaillant dans les pays en voie de développement*. Paris, Presses Universitaire de France, 1977.
4. Peter Zorkoczy. *Information technology - An introduction*. Pitman, London, 1985.
5. Jacques Chaumier. *Systèmes d'information - Marché et Technologies*. Entreprise moderne d'édition, Paris, 1986.
6. Guy Pujolle. *Télécommunications et réseaux*. Eyrolles, Paris, 1992.
7. Des banques de données. Ministère de la Recherche et de l' Enseignement Supérieur, 1987.
8. John N. Olsgaard, editor. *Principles and applicatios os Information Science for library professionals*. American Library Association. Chicago and London, 1989.

9. James A.O' Brien. *Management Information Systems*. IRWIN, Burr Ridge, Illinois, Boston, Massachusetts, Sydney, Australia, 1993.

10. Jean-Charles Pomerol. *Les systèmes experts - Hermès*, Paris, 1988.

11. Peter Hodson and Mike Watkins. *Information Technology* · DP Publication Ltd, Aldine Place, London, 1995.

12. Jan C.A. van der Lubbe. *Information Theory*. Cambridge University Press, 1997.

13. Understanding MARC Bibliographic: Machine Readable Cataloguing/ published by the Cataloguing Distribution Service, Library of Congress, in collaboration with The Follett Software Co. in Betty Furrie, 2000.- 51p.

14. Mary Mortimer. *Descriptive cataloguing*. Canberra, DocMatrix Pty Ltd, 1999.

15. Digital Libraries. William Y. Arms. Online Edition, The MIT Press, 2002.

16. Phạm Thị Lệ Hương. Cẩm nang hướng dẫn sử dụng bộ quy tắc biên mục Anh – Mỹ rút gọn, 1988. Ấn phẩm điện tử PDF.- VA: LEAF - Việt Nam, 2004.

17. Đoàn Phan Tân. *Tìm hiểu cấu trúc ngôn ngữ tư liệu*. Tập san Thư viện, số 2/3 - 1992, trang 6 - 10.

18. Đoàn Phan Tân. *Mô hình hoá quá trình lưu trữ và tìm kiếm thông tin*. Tạp chí Thông tin và Tư liệu, số 4 - 1992, trang 9 - 15.

19. Đoàn Phan Tân. *Cơ sở khoa học mới của ngành thông tin - thư viện*. Tạp chí Văn hoá - Nghệ thuật, số 3 - 1999, trang 25 - 27.

20. Đoàn Phan Tân. *Về thị trường thông tin và kinh tế thông tin*. Tạp chí Văn hoá - Nghệ thuật, số 8 - 1999, trang 58 - 60.

21. Đoàn Phan Tân. *Tìm hiểu quá trình phát triển của các kỹ thuật truyền tin*. Tạp chí Văn hoá - Nghệ thuật, số 2 - 2000, trang 54 - 56.

22. Đoàn Phan Tân. *Tìm hiểu quá trình hình thành và phát triển Thông tin học*. Tập san Thư viện, số 3 - 2000, trang 18 - 21.

23. Vũ Văn Sơn. *Giáo trình Biên mục mô tả*. NXB Đại học Quốc gia Hà Nội, 2000.

24. Đoàn Phan Tân. *Về khái niệm thông tin và các thuộc tính làm nên giá trị của thông tin*. Tạp chí Văn hóa - Nghệ thuật, số 3 - 2001, trang 47 - 49.

NHÀ XUẤT BẢN ĐẠI HỌC QUỐC GIA HÀ NỘI

16 Hàng Chuối – Hai Bà Trưng – Hà Nội

Điện thoại: (04) 971 8312; (04) 972 4770. Fax: (04) 971 4899

Email: nxb@vnu.edu.vn

Chịu trách nhiệm xuất bản:

Giám đốc:

PHÙNG QUỐC BẢO

Tổng biên tập:

PHẠM THÀNH HÙNG

Biên tập tái bản:

LAN HƯƠNG

Chế bản:

BÙI VIỆT HỒNG

Trình bày bìa:

MAI THU HÀ

THÔNG TIN HỌC

Mã số: 1L-91 ĐH2006

In 1000 cuốn, khổ 14,5×20,5 cm.

Tại Xưởng in Tin học & Đời sống. ĐT: 7567556.

Số xuất bản 571-2006/CXB/5 – 102/ĐHQGHN, ngày 21/07/2006

Quyết định xuất bản số: 233 LK/XB

In xong và nộp lưu chiểu Quý III năm 2006.



CÙNG MỘT TÁC GIẢ

1. **Cổ sở thông tin học**
Trường Đại học Văn hóa Hà Nội - 1990
2. **Một số phương pháp toán học trong công tác thư viện - thông tin**
Trường Đại học Văn hóa Hà Nội - 1992
3. **Tin học trong hoạt động thông tin - thư viện**
Nhà xuất bản Văn hóa - Thông tin - 1997
4. **Thông tin học**
Nhà xuất bản Đại học Quốc gia Hà Nội - 2001
5. **Tin học trong hoạt động thông tin - thư viện**
Nhà xuất bản Đại học Quốc gia Hà Nội - 2001
(Xuất bản lần thứ 2 có sửa chữa, bổ sung)
6. **Các hệ thống động thông tin quản lý**
Trường Đại học Văn hóa Hà Nội - 2004

thông tin học (34000)



34.000 VNĐ

34.000 VNĐ