

quan trọng trong phát triển kinh tế-xã hội và kinh tế tri thức. Dịch vụ công về thông tin KHCN sẽ đóng vai trò khơi dậy, nuôi dưỡng và thúc đẩy các hoạt động sáng tạo, đổi mới của mọi tổ chức, cá nhân trong xã hội.

Tới năm 2015, Trung tâm Thông tin KHCN Quốc gia phấn đấu trở thành Tập đoàn dịch vụ công về thông tin KHCN, trong đó có:

- Thư viện điện tử quốc gia về KHCN đạt trình độ tiên tiến trong khu vực;
- Ngân hàng dữ liệu quốc gia cung cấp thông tin

KHCN cần thiết cho các doanh nghiệp, các tổ chức KHCN;

- Trung tâm xúc tiến thị trường công nghệ, bao hàm Sàn giao dịch điện tử về công nghệ, Techmart Việt Nam, Techmart khu vực,...;

- Mạng thông tin KHCN Việt Nam - mạng nòng cốt của Hệ thống thông tin quốc gia về KHCN;

- Trung tâm nghiên cứu và đào tạo nguồn nhân lực về thông tin KHCN;

- Bảo tàng quốc gia về KHCN góp phần giáo dục truyền thống và nâng cao

nhận thức của xã hội về KHCN.

### Tài liệu tham khảo

1. Đảng Cộng sản Việt Nam. Dự thảo Các văn kiện trình Đại hội X của Đảng. Hà Nội, tháng 9 năm 2005.

2. Nghị định số 159/2004/NĐ-CP ngày 31 tháng 8 năm 2004 của Chính phủ về Hoạt động thông tin Khoa học và công nghệ.

3. Điều lệ Tổ chức và hoạt động của Trung tâm Thông tin Khoa học và Công nghệ Quốc gia (Ban hành kèm theo Quyết định số 11/QĐ-BKH&CN ngày 13/05/2004 của Bộ trưởng Bộ Khoa học và Công nghệ).

## VẤN ĐỀ PHÁT TRIỂN VÀ CHIA SẺ NGUỒN LỰC THÔNG TIN SỐ HOÁ TẠI VIỆT NAM

PGS.TS. Nguyễn Hữu Hùng  
Trung tâm Thông tin KHCN Quốc gia

*Trình bày khái niệm và luận chứng vai trò trung tâm của tài nguyên thông tin số trong hệ thống thông tin quốc gia. Giới thiệu 3 kịch bản trong tạo lập tài nguyên số: số hoá toàn phần, số hoá hồi cố và song song tồn tại tài nguyên số và tư liệu. Đưa ra các điều kiện và yếu tố cần thiết để thực hiện việc chia sẻ tài nguyên số trên qui mô hệ thống.*

### Đặt vấn đề

Thế kỷ XXI được đánh dấu bằng sự ra đời của nền kinh tế mới, được gắn bằng nhiều thuật ngữ: kinh tế tri thức, kinh tế số, kinh tế mạng,... Trong nền kinh tế này, xuất hiện một loại

nguồn lực quan trọng là tài nguyên thông tin số. Trong những năm qua, trong khuôn khổ của Hệ thống thông tin quốc gia, tại không ít các cơ quan thông tin, tư liệu, thư viện ở các cấp (trung ương, địa

phương), các khu vực (kinh tế, thương mại, KHCN, giáo dục, đào tạo) đã có những nỗ lực đáng kể để hình thành bước đầu phần tài nguyên này. Cũng như bất kỳ loại nguồn lực nào trong xã hội, vấn đề cần thiết phải

xác lập được phương thức quản lý và chiến lược phát triển thích hợp.

### 1. Khái niệm

Trong lịch sử phát triển nhân loại, trình độ lưu giữ và quản trị thông tin trong xã hội bao giờ cũng là một chỉ tiêu quan trọng thể hiện trình độ phát triển văn hoá, khoa học kỹ thuật và công nghệ. Trong thang cuốn lịch sử thông tin, tiếng nói là kỹ thuật truyền tin cổ xưa nhất, nó là vật mang tin không thể lưu giữ được theo thời gian và khả năng của nó hết sức hạn chế trong không gian. Chữ viết là nấc kỹ thuật cách mạng, nhờ nó mà người ta có thể ghi chép được các sự kiện và giúp con người thực hiện chức năng chuyển giao thông tin giữa các thế hệ. Thế kỷ XV, máy in ra đời, đó là một thành tựu kỹ thuật rất lớn tạo điều kiện để con người giao lưu được rộng khắp và thông tin mang rõ tính chất quốc tế của nó. Sau chiến tranh thế giới lần thứ II, một hệ thống các công cụ kỹ thuật thông tin mới dựa trên các thành tựu của kỹ thuật vi điện tử đã làm chấn động trong thông tin. Đầu tiên là máy điện toán, tiếp là máy vi tính, mạng thông tin và nay là các công cụ số hoá.

Tất cả những thành tựu kỹ thuật đó đã tích cực tạo dựng và góp phần làm hiện thực hoá một "xã hội thông tin".

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

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

việc:

- Kiểm soát tài nguyên thông tin;

- Bảo vệ an toàn và lâu dài các tài liệu gốc (điều này đặc biệt có ý nghĩa khi số hoá các tài liệu có giá trị quý, hiếm như các chứng cứ của lịch sử, là di sản văn hoá...);

- Nâng cao năng lực khai thác thông tin của người dùng tin;

- Dễ dàng tạo lập các loại sản phẩm và dịch vụ thông tin mới;

- Thúc đẩy mở rộng việc chia sẻ thông tin trong Hệ thống thông tin quốc gia.

Nói về các ưu điểm trên đây đối với tài nguyên số không có nghĩa là tài nguyên số không có các hạn chế của nó. Cũng cần lưu ý rằng, đến nay, nhiều chuyên gia trong lĩnh vực này cũng chưa thực dám khẳng định, công nghệ số hoá là phương thức và là giải pháp bảo quản thông tin tốt nhất.

Để có được một nguồn tài nguyên số từ việc số hoá các tài liệu trong Hệ thống thông tin quốc gia có thể thực hiện thông qua các đề án số hoá bằng:

- Tạo lập các loại CSDL (thư mục, tóm tắt, số liệu, dữ

kiện, toàn văn) bằng các phần mềm ứng dụng;

- Kỹ thuật mã hoá và chụp ảnh số (digital camera);

- Dùng máy quét (scanner).

## 2. Tài nguyên thông tin số - hợp phần trung tâm của Hệ thống thông tin quốc gia

Hệ thống thông tin quốc gia bao gồm bốn hợp phần cơ bản:

- Kỹ thuật: lõi là kết cấu hạ tầng thông tin;

- Thông tin: nguồn lực thông tin, trong đó phần được số hoá là hợp phần tích cực;

- Con người: cán bộ thông tin và cộng đồng người dùng tin;

- Tổ chức: mức độ tập trung - phân tán, qui mô, hình thái...

Ở bất cứ hệ thống thông tin nào, giá trị của hệ thống thể hiện ở chỗ chúng quản trị được những loại thông tin gì và tổ chức khai thác các thông tin đó như thế nào. Trong Hệ thống thông tin quốc gia, phần tài nguyên số giữ vai trò trung tâm.

Hiện nay, trên thế giới có nhiều chiến lược khác nhau để số hoá tài nguyên thông

tin. Trên quy mô quốc gia, trong điều kiện hiện tại, việc phát triển tài nguyên thông tin số có thể nhìn nhận theo 3 mức như 3 kịch bản:

- Số hoá toàn phần (fully digital resources);

- Song song tồn tại tài liệu và nguồn lực số (parallel resources);

- Số hoá hồi cố (Retrospective digitization).

Cách đây gần 40 năm (năm 1966), Licklider đã đưa ra đề xuất giải pháp điện tử cho các tài liệu trong hệ thống liên lạc xã hội với kỳ vọng, giải pháp điện tử sẽ chấm dứt những vấn đề thông tin nhức nhối của xã hội bấy giờ, như: kim hãm bùng nổ thông tin, đẩy nhanh lưu thông, giảm chi phí, thay thế vãn nạn giấy tờ trong từng cơ quan và trong toàn xã hội...

Tuy nhiên, với thời gian, công nghệ vẫn không thể làm được hết và tốt chức năng của hệ thống ấn phẩm "Từ tác giả tới bạn đọc", và hơn nữa cho tới tận ngày nay, người ta vẫn chưa hẳn dám khẳng định về tính lâu bền, tuổi thọ cần thiết của các "công dân điện tử" này. Do vậy, việc tồn tại của hệ tài nguyên thông tin lai (hybrid) vẫn được coi là giải

pháp được lựa chọn ở nhiều nước hiện nay, cho dù, các nhà sản xuất phần mềm cho rằng, họ sẵn sàng cung ứng đủ công cụ hữu hiệu để trao đổi thông tin trên các mạng tốc độ cao, có đủ năng lực kết nối với các kho thông tin khổng lồ được số hoá. Trên thực tế, nhiều người dùng tin vẫn ưa thích các sản phẩm in, thích làm việc với môi trường giấy hơn. Trong bối cảnh như vậy, các nhà xuất bản vẫn tiếp tục cho ra đời các sản phẩm in, các giao dịch trên giấy vẫn còn khá thịnh hành...

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

Việc số hoá các nguồn tin không phải chỉ áp dụng ở các nguồn tài liệu hiện tại và tài liệu sẽ được xuất bản trong tương lai. Tuy nhiên, việc số hoá các tài liệu là công việc tốn kém, đòi hỏi nhiều kinh phí, lao động và trang thiết bị và phí chuyển đổi về tổ chức. Do vậy, xây dựng và phát triển tài nguyên số không thể làm

tràn lan mà phải có chọn lọc, trọng tâm, trọng điểm. Trên thế giới, nhiều cơ quan TT-TV lớn như: các thư viện của các nước Hoa Kỳ, Anh, Pháp, CHLB Đức ... đều có các dự án để số hoá vốn tư liệu hồi cổ. Các dự án như vậy được đặt ra vì:

- Để việc truy cập tới các kho tài liệu dễ và thuận tiện hơn;

- Giảm thiểu không gian để lưu trữ;

- Bảo vệ cho các tài liệu giấy khỏi bị huỷ hoại trong quá trình đưa ra phục vụ;

- Dễ khai thác.

Tuỳ theo các điều kiện cụ thể, các đơn vị thông tin phải có các nghiên cứu cần thiết, như: loại tài liệu nào cần được số hoá, khối lượng, qui mô, phương thức lựa chọn ... để lập kế hoạch sao cho phù hợp.

Tổ chức các tài liệu số hoá có thể nhìn từ hai phương diện:

- Dành cho cán bộ thông tin (người quản trị tài nguyên): Phương thức xây dựng, tổ chức và lưu hành - trao đổi;

- Dành cho người sử dụng (nhà khoa học, doanh nhân, nhà lãnh đạo...): phương thức truy cập, chọn

lọc, hiển thị tài nguyên số.

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

liệu mới, có thể dễ dàng bổ sung vào bằng cách tái xây dựng hoặc tái cấu trúc. Trong thực tiễn, để xây dựng một bộ sưu tập về đề tài xác định thường ta phải đối mặt với cả hai loại tài liệu: tài liệu đã ở dạng điện tử và tài liệu còn ở dạng in cần phải số hoá. Nếu tài liệu ở dạng điện tử thì công việc khá dễ dàng, ngay cả việc sưu tầm, tổ chức tệp tin và chuyển đổi dạng thức. Còn tài liệu còn dạng in cần phải thực hiện các thao tác số hoá bằng cách quét (scanning) tài liệu và chuyển qua dạng PDF - một chuẩn được sử dụng rộng rãi trong số hoá tài liệu bằng phương pháp quét ảnh. Về việc xây dựng các bộ sưu tập số thì những kinh nghiệm gần đây của Thư viện Đại học Khoa học Tự nhiên thuộc Đại học Quốc gia TP. Hồ Chí Minh sử dụng phần mềm mã nguồn mở Greenstone đang được sử dụng ở nhiều cơ quan TT-TV rất đáng được tham khảo.

Thời gian qua, không ít các cơ quan TT-TV ở nước ta đã có những nỗ lực để xây dựng các phần tài nguyên số. Trung tâm Thông tin KHCN Quốc gia đã thực hiện việc số hoá một khối lượng đáng kể các tài liệu

báo cáo các kết quả nghiên cứu được giao nộp và hiện đang được bảo quản tại Trung tâm. Một số các trường đại học lớn như Đại học Quốc gia Hà Nội và Đại học Quốc gia TP. Hồ Chí Minh có các dự án số hoá các tài liệu xây dựng kho tài nguyên học tập phục vụ việc giảng dạy và đào tạo. Thư viện Quốc gia tích cực thực hiện việc xây dựng CSDL toàn văn đối với nguồn tài liệu luận án khoa học thu thập được, Trung tâm TT-TV Văn phòng Quốc hội xây dựng CSDL toàn văn đối với nguồn tài liệu văn bản pháp luật ...

Xây dựng tài nguyên số bằng phương thức trên có nghĩa là tổ chức lại thông tin, biến chúng trở thành nguồn lực, làm cho các thông tin đó trở nên phổ biến hơn đối với đông đảo người dùng tin, giúp cho hệ thống giao lưu khoa học dễ đạt tới chức năng "từ tác giả tới độc giả" mà chỉ với các ấn phẩm truyền thống rất khó, nếu như không muốn nói là không thể thực hiện. Tuy nhiên, vấn đề phổ biến thông tin trong môi trường số hoá có mặt pháp lý, mà những người tham gia cần phải am hiểu các điều khoản của luật xuất bản, sở hữu trí

tuệ, qui định bảo mật, pháp lệnh lưu trữ... để hành động đúng, để không rơi vào lỗi vi phạm pháp luật.

### 3. Chia sẻ nguồn tài nguyên thông tin số

Trên quy mô quốc gia, việc thiết lập nguồn tài nguyên thông tin số là việc làm cần thiết song cũng tốn kém. Để tạo lập được không gian nguồn tài nguyên thông tin số thống nhất trong hệ thống thông tin quốc gia, tránh sự lãng phí do trùng lặp, tổ chức thiếu khoa học cần sớm có dự án hình thành tổ hợp chia sẻ tài nguyên số. Ý tưởng và giá trị của việc chia sẻ thông tin đã được Bernard Shaw từ lâu diễn đạt bằng hình tượng: "Nếu mỗi người có một quả táo mà trao đổi với nhau thì kết cục mỗi người vẫn chỉ có một quả táo. Nhưng nếu mỗi người có một ý tưởng và trao đổi với nhau thì kết quả mỗi người sẽ có hai ý tưởng". Để đảm bảo cho sự chia sẻ được thực hiện thành công, phải chuẩn bị các điều kiện và các yếu tố cần thiết sau giữa các cơ quan TT-TV thành viên:

- Có chương trình hợp tác với sự thống nhất về những nguyên tắc cơ bản như: mục tiêu, quyền lợi, nghĩa vụ,

phương thức hợp tác; phạm vi và khả năng mở rộng hợp tác;

- Đảm bảo phần kiến trúc hạ tầng: tuân thủ tiêu chuẩn chung về tổ chức tài nguyên thông tin số; có phần mềm tích hợp các tiêu chuẩn chung, và có phần nguồn tài nguyên thông tin chia sẻ được;

- Có ban điều hành am hiểu về tổ chức hoạt động TT-TV và hiểu biết về ứng dụng CNTT, có năng lực điều hành và điều chỉnh phù hợp với những biến động;

- Tuân thủ các chuẩn nghiệp vụ trong hoạt động TT-TV như các chuẩn mô tả, trình bày, trao đổi dữ liệu (MARC 21, AACR-2 ...);

- Xây dựng được một chương trình và kế hoạch hành động trên cơ sở nắm vững các đặc điểm, các thuận lợi và khó khăn của các đơn vị thành viên;

- Định hướng và tổ chức huy động được các nguồn lực, bao gồm: nhân lực, vật lực, kỹ lực, tin lực, tài lực, cần thiết đảm bảo hoạt động lâu dài của hợp tác;

- Thực hiện việc đảm bảo các vấn đề pháp lý trong việc số hoá.

Trên bình diện quốc gia phải có sự phân hoạch khoa

học để tránh sự trùng lặp, dẫm chân nhau. Thực tế, thời gian qua, tuy chưa có sự tổng kết, song chắc chắn, việc số hoá nhiều nguồn tài liệu, đặc biệt là các tài liệu trong nước, như đề tài nghiên cứu, các dự án, luận án,... xảy ra chông chéo ở nhiều cơ quan. Như vậy, nguyên tắc "xử lý một lần", chia sẻ thông tin vẫn còn bị vi phạm nghiêm trọng.

#### 4. Phát triển trong tương lai

Với sự phát triển nhanh về CNTT và truyền thông, phương thức hoạt động của các cơ quan TT-TV sẽ còn có nhiều thay đổi, trong số đó có trào lưu số hoá. Việc mở rộng phạm vi và qui mô số hoá, hình thành nguồn tài nguyên số và tích cực thực

hiện việc chia sẻ tài nguyên này trong khuôn khổ của hệ thống thông tin quốc gia sẽ thúc đẩy các cơ quan TT-TV thực sự chuyển từ vai trò "mua sách và quản lý sách" sang vai trò "mua thông tin và quản trị thông tin". Điều này cũng có nghĩa, đòi hỏi phải có chiến lược để giải quyết các vấn đề:

- Số hoá;
- Thiết lập và khẳng định các chuẩn nghiệp vụ;
- Thiết lập các mạng;
- Phát triển các máy trạm và các công cụ hỗ trợ;
- Mở rộng phương thức chia sẻ và cung cấp thông tin trong môi trường điện tử.

Trong bối cảnh đó, các thư viện điện tử với các nguồn tài nguyên số hoá sẽ

không ngừng được mở rộng và phát triển vì nhiều lý do, trong đó có :

- Cung cấp các dịch vụ TT-TV trên nền vốn tài nguyên số trở thành một chuẩn mực hoạt động;
- Thay đổi nhu cầu sử dụng thông tin của cộng đồng người dùng tin trẻ tuổi và năng động;
- Áp lực nhu cầu được cung cấp thông tin cho các người dùng tin ở mọi nơi: tại công sở, tại vị trí làm việc, tại tư gia;
- Yêu cầu của xã hội trong việc tiếp tục cắt giảm chi phí cho hoạt động xây dựng, phát triển và duy trì hoạt động của các cơ quan TT-TV.

#### Tài liệu tham khảo

1. *A vision of intelligent island: IT 2000 Report.* National Computer Board, Singapore. 1994
2. Bellardo,t., Bourne, C.P. *The history of the development of online systems and services in the U.S.* New York: Academic Press. 1992
3. Basch, R. *Books online: Visions, Plans and Perspectives for Electronic Texts.*- Online, July 1995, p. 12-25
4. *Chiến lược tăng cường công tác thông tin KH&CN phục vụ CNH và HĐH đất nước.*- H.- Trung tâm Thông tin Tư liệu KH&CN Quốc gia, 1998.
5. *Các nguyên tắc của chính sách nhà nước trong lĩnh vực thông tin.*//Thông tin hoá, 1997. Số 3 (tiếng Nga).
6. Nguyễn Hữu Hùng. *Một số quan điểm về xây dựng chính sách quốc gia về thông tin-tư liệu KH&CN trong giai đoạn CNH và HĐH.*- Kỷ yếu Hội nghị Ngành Thông tin - Tư liệu KH&CN.-H, 1998, tr 57-63.
7. *Information as a raw material for innovation.*- Ministry of Education, Science, Research and Technology.- Born, 1997.
8. Koenig, M. E. D. *Information technology and perestroika. Information Services and Use.* 1990. Số 10. Tr 305-320
9. Kedrovskij O. V. *Nguồn lực thông tin và chính sách thông tin.*//Thông tin KH&CN. Seri 1, 1998. Số 7. Tr.2-4 (tiếng Nga).
10. Martyn, J. *Information UK 2000.* Bowker Saur 1990
11. Williamson, R. *Knowledge Warehouse. Library and Information Research Reporty.* 1997