

CHUYỂN ĐỔI SỐ TRONG PHÁT TRIỂN TÀI NGUYÊN THÔNG TIN KHOA HỌC VÀ CÔNG NGHỆ SỐ

ThS Cao Minh Kiểm

Tổng thư ký Hội Thông tin khoa học và công nghệ Việt Nam

● **Tóm tắt:** Giới thiệu khái niệm chuyển đổi số và một số khía cạnh chuyển đổi số trong phát triển, chia sẻ tài nguyên thông tin số khoa học và công nghệ.

● **Từ khóa:** Chuyển đổi số; tài nguyên thông tin; thông tin khoa học và công nghệ.

DIGITAL TRANSFORMATION IN THE DEVELOPMENT OF DIGITAL INFORMATION RESOURCES ON SCIENCE AND TECHNOLOGY

● **Abstract:** To introduce the concept of digital transformation and some aspects of digital transformation in developing and sharing digital information resources on science and technology.

● **Keywords:** Digital transformation; information resources; science and technology information.

Mở đầu

Cuộc Cách mạng công nghiệp lần thứ 4, thường được gọi là Cách mạng công nghiệp 4.0 (CMCN 4.0), đang diễn ra trên thế giới và đang làm thay đổi căn bản nền sản xuất của thế giới. Đây là cuộc cách mạng công nghiệp dựa trên công nghệ số mà nền tảng là tích hợp cao độ hệ thống kết nối số hóa - vật lý - sinh học với sự phát triển của internet kết nối vạn vật (IoT), dữ liệu lớn (Big data), trí tuệ nhân tạo (AI). CMCN 4.0 có tác động mạnh mẽ đến sự chuyển đổi của toàn bộ hệ thống, các quốc gia, các doanh nghiệp, các ngành công nghiệp và toàn xã hội. Để CMCN 4.0 được thực hiện thành công, các nước trên thế giới phải thực hiện “Chuyển đổi số” (Digital transformation). Việc chuyển đổi số sẽ làm thay đổi công nghệ, biến đổi hoạt động sản xuất - kinh doanh, thay đổi các hoạt động xã hội. Gần đây, Chính phủ đã yêu cầu Bộ Thông tin và Truyền thông xây dựng và trình Chính phủ Đề án chuyển đổi số quốc gia (1) để đẩy mạnh chuyển đổi số ở Việt Nam. Một trong những nội dung của dự thảo đề án là phát triển tài nguyên dữ liệu quốc gia. Việc phát triển và chia sẻ tài nguyên thông tin KH&CN là một khía cạnh

quan trọng của hoạt động thông tin KH&CN và cần được quan tâm phát triển trong bối cảnh chuyển đổi số quốc gia và chuyển đổi số trong lĩnh vực thông tin KH&CN.

Bài viết này giới thiệu khái niệm chuyển đổi số và một số khía cạnh chuyển đổi số trong phát triển, chia sẻ tài nguyên thông tin số KH&CN.

1. KHÁI QUÁT VỀ CHUYỂN ĐỔI SỐ

1.1. Khái niệm chuyển đổi số

Hiện nay, chưa có định nghĩa thống nhất về chuyển đổi số (Digital transformation). Định nghĩa trên Wikipedia cho rằng, chuyển đổi số là việc sử dụng những công nghệ số mới và thay đổi nhanh chóng để giải quyết các vấn đề, thường bằng điện toán đám mây, giảm sự phụ thuộc vào phần cứng do khách hàng sở hữu nhưng nâng cao việc sử dụng các dịch vụ dựa trên điện toán đám mây [Wikipedia]. Công ty tư vấn hàng đầu thế giới Garner cho rằng, chuyển đổi số có nội hàm khác nhau tùy thuộc vào ngữ cảnh. Nó có thể đề cập đến mọi việc từ hiện đại hóa công nghệ thông tin (như điện toán đám mây) đến tối ưu hóa số (digital optimization), hay sáng tạo ra mô hình

1. Số liệu tra cứu từ CSDL của Cục Thông tin KH&CN quốc gia http://sti.vista.gov.vn/Pages/ket-qua-nghien-cuu-moi-nhat.aspx?Type_CSDL=KETQUANHIEMVU

kinh doanh số mới. Trong lĩnh vực các tổ chức công, chuyển đổi số cũng có thể liên quan đến những sáng kiến khiêm tốn như thực hiện các dịch vụ công trực tuyến hoặc hiện đại hóa di sản (legacy modernization) [Gartner IT Glossary]. Theo Garner, thuật ngữ “chuyển đổi kinh doanh số” (digital business transformation) có thể được hiểu là việc sử dụng các công nghệ số để thay đổi mô hình kinh doanh và tạo ra cơ hội doanh thu và giá trị mới [trích theo Cục Tin học hóa, 2018; Gartner IT Glossary].

Như vậy, có thể thấy rằng, chuyển đổi số là một sự chuyển đổi mạnh mẽ các hoạt động, quá trình, mô hình và hiểu biết về kinh doanh và tổ chức để tận dụng triệt để những thay đổi và cơ hội của sự kết hợp các công nghệ kỹ thuật số và tác động tăng tốc của chúng trên toàn xã hội theo cách chiến lược và ưu tiên, với những thay đổi hiện tại và tương lai [i-SCOOP].

Hiện nay ở Việt Nam, chúng ta có thể thống nhất với định nghĩa của Bộ Thông tin và Truyền thông rằng “Chuyển đổi số được hiểu là việc sử dụng dữ liệu và công nghệ số để thay đổi một cách tổng thể và toàn diện tất cả các khía cạnh của đời sống kinh tế - xã hội, tái định hình cách chúng ta sống, làm việc và liên hệ với nhau” [Bộ TT&TT, 2019]. Như vậy, chuyển đổi số có sự liên hệ chặt chẽ với thông tin/dữ liệu số, các công nghệ số và việc ứng dụng chúng một cách mạnh mẽ trong mọi hoạt động của cơ quan, tổ chức và cả quy mô quốc gia.

1.2. Sự khác biệt giữa “Số hóa”, “Số hóa hoạt động” và “Chuyển đổi số”

Chuyển đổi số không đơn thuần chỉ là việc số hóa, tạo ra dữ liệu số hoặc việc ứng dụng công nghệ thông tin (công nghệ số) vào các hoạt động. Thực tế, chuyển đổi số có sự khác biệt với số hóa (digitization) và số hóa hoạt động (digitalization) [i-SCOOP; Igniger Anna, 2017].

Số hóa (Digitization) là quá trình chuyển đổi từ dạng tương tự sang dạng số (dạng số hóa). Thực chất số hóa là quá trình chuyển đổi những thông tin trên những đối tượng thực sang dạng điện tử hay còn gọi là dạng số. Những đối tượng thực chứa thông tin có thể là các tài liệu dạng văn bản, hình ảnh, bản đồ, băng ghi âm, ghi hình,... trên các vật mang tin vật lý (trên giấy, trên phim, giấy ảnh, vi hình, băng ghi âm băng ghi hình,...). Kết quả của số hóa là những đối tượng thực được chuyển sang đối tượng số dưới hình thức tệp tin. Mục đích của số hoá có thể là: tăng cường sự truy cập đến tài nguyên thông tin; cải thiện chất lượng dịch vụ cho những người dùng tin thông qua khả năng truy cập được cải thiện; giảm việc tiếp xúc trực tiếp đến những tài nguyên quý, hiếm, cổ hoặc được sử dụng nhiều; tạo ra bản sao lưu trữ; cho phép cơ quan, đơn vị phát triển hạ tầng kỹ thuật và kỹ năng của nhân viên; phát triển khả năng chia sẻ tài nguyên thông tin [IFLA, 2002]. Nói một cách đơn giản, số hóa mới chỉ là việc chuyển từ dạng giấy sang định dạng số (digital format). Ngày nay, trong hoạt động thông tin thư viện, với việc ứng dụng công nghệ thông tin, nhiều thông tin đã được tạo lập ở dạng số ngay từ đầu mà không cần phải trải qua công đoạn số hóa. Vì thế, số hóa cũng có thể được hiểu rộng hơn: đó là quá trình tạo lập thông tin/dữ liệu số, có thể là thông qua việc chuyển đổi từ dạng vật lý (tương tự) sang dạng số hoặc đã là ở dạng số ngay từ khi được tạo ra (digital born).

Theo Gartner, Digitalization (tạm dịch là số hóa hoạt động) (2) là việc sử dụng các công nghệ số để thay đổi mô hình kinh doanh và cung cấp những cơ hội tạo ra giá trị và doanh thu mới; đó là quá trình chuyển sang kinh doanh số [Gartner Glossary]. Anna Igniger (2017) cho rằng, “digitalization” là quá trình làm cho thông tin số có giá trị cho con người và coi đó là quá trình xem xét sao cho áp dụng tốt nhất thông tin số/số hóa

2. Cũng có tài liệu gọi là “số hóa”

để đơn giản hóa mọi hoạt động. Trong kinh doanh, ứng dụng số thường đề cập đến việc cho phép, cải thiện và/hoặc chuyển đổi các hoạt động kinh doanh và/hoặc chức năng kinh doanh và/hoặc mô hình/quy trình và/hoặc hoạt động kinh doanh bằng cách tận dụng các công nghệ kỹ thuật số, sử dụng rộng rãi hơn dữ liệu số, được chuyển thành hành động, kiến thức cho mục đích cụ thể [i-SCOOP]. Nói tóm lại, “digitalization” được hiểu là việc sử dụng công nghệ số (digital technologies) và dữ liệu (được số hóa hoặc là dạng số ngay từ đầu (digital-born hay natively digital)) để tạo ra doanh thu, cải thiện kinh doanh, thay thế/chuyển đổi quá trình kinh doanh (không chỉ đơn giản là số hóa nó) và tạo ra môi trường kinh doanh số, trong đó thông tin/dữ liệu số là then chốt. “Digitalization” sẽ dẫn đến kinh doanh/hoạt động số, còn chuyển đổi số đòi hỏi kinh doanh số và số hóa.

Chuyển đổi số là sự chuyển đổi sâu sắc và tăng tốc các hoạt động, quy trình, năng lực và mô hình kinh doanh để tận dụng triệt để những thay đổi và cơ hội của công nghệ số và tác động của chúng trên toàn xã hội theo cách chiến lược và ưu tiên [i-SCOOP]. Đó là quá trình tạo lập, phát triển ra các ứng dụng kinh doanh mới tích hợp tất cả dữ liệu số và các ứng dụng số này [Irniger Anna, 2017]. Giáo sư Hồ Tú Bảo (Viện nghiên cứu cao cấp

về toán) đã tổng kết một công thức thể hiện định nghĩa chuyển đổi số [Hồ Tú Bảo, 2018]:

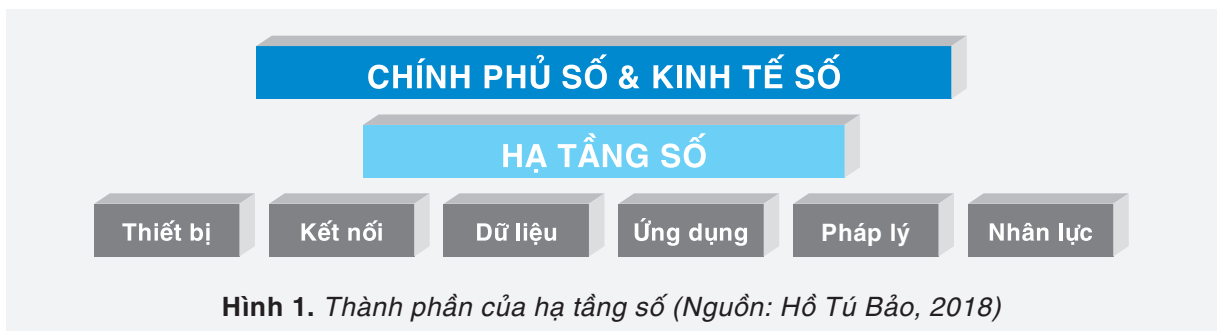
Chuyển đổi số = Số hóa + Công nghệ số

Như vậy, chuyển đổi số không chỉ đơn thuần là việc số hóa, tạo lập tài nguyên dữ liệu số mà là bao gồm việc chuyển đổi sâu sắc quá trình tạo lập thông tin số bằng việc ứng dụng các công nghệ số, hình thành và phát triển những dịch vụ mới, tạo ra giá trị gia tăng từ tài nguyên số bằng việc sử dụng những công nghệ số tiên tiến.

1.3. Hạ tầng dữ liệu có vai trò quan trọng trong chuyển đổi số

Để thực hiện chuyển đổi số, cần đảm bảo các yếu tố nền tảng (hạ tầng), bao gồm [Bộ Thông tin và Truyền thông, 2019]:

- Hạ tầng số (phát triển mạng di động thế hệ mới, kết nối cáp quang đến các gia đình, doanh nghiệp, cung cấp wifi miễn phí tại khu vực công cộng, phát triển điện toán đám mây- cloud computing, hạ tầng IoT, BigData,...);
- Nhân lực có kỹ năng số (digital skills);
- Nghiên cứu và phát triển công nghệ số mới;
- Xây dựng môi trường pháp lý để bảo đảm môi trường an toàn, tin cậy, thúc đẩy chuyển đổi số.



Hình 1. Thành phần của hạ tầng số (Nguồn: Hồ Tú Bảo, 2018)

Trong hạ tầng số, hạ tầng dữ liệu đóng vai trò then chốt, quyết định đến chuyển đổi số [Cục Tin học hóa, 2018; Hồ Tú Bảo, 2018]. Những thành phần cơ bản của hạ tầng số bao gồm: hạ tầng thiết bị, hạ tầng kết nối (các mạng truyền thông), hạ tầng dữ

liệu; những ứng dụng; các quy định pháp lý và nhân lực (Hình 1).

Có thể cho rằng, dữ liệu trong hoạt động thông tin KH&CN chính là tài nguyên thông tin KH&CN số. Như vậy, phát triển và chia

sở tài nguyên thông tin KH&CN số đóng vai trò cốt lõi trong chuyển đổi số trong hoạt động thông tin KH&CN.

2. KHÁI QUÁT VỀ TÀI NGUYÊN THÔNG TIN KHOA HỌC VÀ CÔNG NGHỆ SỐ Ở VIỆT NAM

2.1. Khái niệm tài nguyên thông tin KH&CN số

Thông tin KH&CN có thể được định nghĩa là “thông tin/dữ liệu về tài liệu và dữ kiện thu nhận được trong quá trình hoạt động khoa học, khoa học kỹ thuật, đổi mới sáng tạo và xã hội” [МОДЕЛЬНЫЙ ЗАКОН О НАУЧНО-ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ]. Nghị định 11/2014/NĐ-CP ngày 18/02/2014 của Chính phủ về hoạt động thông tin KH&CN xác định, thông tin KH&CN không chỉ là thông tin/dữ liệu về tài liệu và dữ kiện, mà còn là dữ liệu, dữ kiện, số liệu, tin tức được tạo ra trong các hoạt động KH&CN, đổi mới sáng tạo. Tri thức được coi là thông tin khi nó được trình bày, được hiển thị trên vật mang tin hoặc ở dạng đọc, xem được (không phải là tri thức ẩn trong não của người có tri thức (3)).

Luật Thư viện 2019 định nghĩa: tài nguyên thông tin là “tập hợp các loại hình tài liệu, bao gồm: tài liệu in, bản chép tay, tài liệu nghe nhìn, tài liệu vi dạng (vi phim, vi phiếu), tài liệu số và tài liệu đặc biệt cho người khuyết tật” [Luật Thư viện 2019]. Theo Nghị định 11/2014/NĐ-CP của Chính phủ về hoạt động thông tin KH&CN, tài nguyên thông tin được hiểu “là các thông tin KH&CN được thể hiện dưới dạng sách, báo, tạp chí khoa học; kỷ yếu hội nghị, hội thảo khoa học; thuyết minh nhiệm vụ, báo cáo kết quả thực hiện, ứng dụng kết quả nhiệm vụ KH&CN; tài liệu sở hữu trí tuệ, tiêu chuẩn, quy chuẩn kỹ thuật; catalô công nghiệp; luận án khoa học; tài liệu thiết kế, kỹ thuật; CSDL; trang

thông tin điện tử; tài liệu thống kê KH&CN; tài liệu đa phương tiện và tài liệu trên các vật mang tin khác” (4). Như vậy, tài nguyên tin KH&CN là toàn bộ thông tin KH&CN được ghi lại, trình bày hoặc thể hiện ở nhiều dạng thức và trên nhiều dạng vật mang tin (ở dạng truyền thống và dạng điện tử) mà tổ chức, quốc gia có được. Những thông tin này có thể là nội sinh/trong nước (do tổ chức hoặc quốc gia đó tạo ra, lưu giữ, phổ biến), thu thập được qua các kênh thông tin khác nhau (mua, trao đổi, biếu tặng,...) hoặc truy cập được một cách ổn định, lâu dài (thông qua mua quyền truy cập hoặc có thể truy cập được do bên có thông tin tạo điều kiện truy cập lâu dài).

Tài nguyên thông tin KH&CN có thể được thể hiện dưới hai dạng tài nguyên thông tin chính là: tài liệu KH&CN và dữ liệu nghiên cứu. Tài liệu KH&CN gồm các loại như công bố KH&CN được đăng tải trên các tạp chí KH&CN (mà chúng ta thường gọi là bài báo khoa học hay bài báo nghiên cứu), tài liệu sáng chế, chuyên khảo, sách, các báo cáo kết quả nhiệm vụ KH&CN, các bài trình bày tại các hội nghị, hội thảo KH&CN, một số loại hình tài liệu KH&CN khác,...). Công bố KH&CN là dạng tài liệu KH&CN trình bày những kết quả và tri thức thu nhận được từ hoạt động nghiên cứu KH&CN hoặc được rút ra, được nhận thức từ các hoạt động KH&CN và các hoạt động khác, được đăng tải trên các tạp chí KH&CN. Công bố KH&CN có thể bao gồm những bài báo nghiên cứu, những thông báo (communication) ngắn, những bài tổng quan, tổng luận, được đăng trên các tạp chí KH&CN. Dữ liệu nghiên cứu (research data) có thể được định nghĩa là “những sự kiện, con số, ký tự, và các ký hiệu mô tả một đối tượng, ý tưởng, điều kiện, hoàn cảnh, hoặc các yếu tố khác” hoặc “sự trình bày có thể

3. Trước đây, Nghị định 159/2004/NĐ-CP về hoạt động thông tin KH&CN định nghĩa thông tin KH&CN là “là các dữ liệu, số liệu, dữ kiện, tin tức, tri thức KH&CN (bao gồm KH tự nhiên, KH công nghệ, KH xã hội và nhân văn) được tạo lập, quản lý và sử dụng nhằm mục đích cung cấp dịch vụ công, phục vụ quản lý nhà nước hoặc đáp ứng nhu cầu của tổ chức, cá nhân trong xã hội”.

4. Nghị định 14/2014/NĐ-CP sử dụng thuật ngữ “nguồn tin KH&CN”.

diễn giải lại của thông tin theo cách thức hình thức hóa phù hợp để truyền thông, giải thích, hoặc xử lý [National Academy of Sciences, 2009]. Tổng cục Nghiên cứu và Đổi mới sáng tạo của Ủy ban châu Âu EC định nghĩa dữ liệu nghiên cứu (Research data) là thông tin, dữ kiện, số liệu được thu thập để nghiên cứu và xem xét làm cơ sở cho suy luận, thảo luận, tính toán [EC. Directorate - General for Research & Innovation, 2016].

Tài nguyên thông tin KH&CN số (còn gọi là tài nguyên điện tử) là những tài nguyên ở dạng số/điện tử mà khi truy cập và khai thác cần sự trợ giúp của các thiết bị công nghệ thông tin và truyền thông (máy tính, mạng thông tin, thiết bị di động/cầm tay...) và

thường được ghi và lưu giữ trên các vật mang tin số như: ổ cứng máy tính (cố định hoặc lưu động), đĩa quang, trên bộ nhớ của hệ thống máy chủ, hoặc lưu giữ trên mạng internet.

2.2. Hiện trạng tài nguyên thông tin khoa học và công nghệ số ở Việt Nam

2.2.1. Sách KH&CN

Theo thống kê, trong những năm gần đây, mỗi năm Việt Nam xuất bản khoảng từ 28.000 đến 30.000 đầu tên sách với số lượng bản in là khoảng hơn 310 triệu bản. Trong đó có khoảng 1.200 đến 1.500 đầu sách khoa học kỹ thuật, 7.500 đến 7.900 đầu sách khoa học xã hội (Bảng 1) [Cao Minh Kiểm, 2019].

Bảng 1. Số lượng sách xuất bản tại Việt Nam giai đoạn 2015-2017

Sách xuất bản	Năm		
	2015	2016	2017
Số đầu tên sách xuất bản	29.014	30.069	28.717
Số bản sách xuất bản (đơn vị: triệu bản)	362,8	334,3	313,9
Sách khoa học kỹ thuật			
- Đầu tên sách	1.532	1.599	1.293
- Số bản (đơn vị: triệu bản)	7,6	4,1	3,3
Sách khoa học xã hội			
- Đầu tên sách	7.593	7.950	7.878
- Số bản (đơn vị: triệu bản)	24,8	15,0	20,4

(Nguồn: Tổng Cục Thống kê, <http://www.gso.gov.vn/default.aspx?tabid=723>. Truy cập ngày 17/8/2019)

Với công nghệ xuất bản hiện nay, có thể nói tất cả các sách KH&CN được xuất bản đều qua khâu chế bản điện tử, như vậy đều có bản số/điện tử của mỗi đầu sách. Tuy nhiên, không có số liệu về việc lưu giữ hoặc tạo lập, phát hành bản điện tử của các sách KH&CN.

2.2.2. Tạp chí KH&CN

Ước tính, trong số trên 500 tên tạp chí được xuất bản trong nước, có 334 tạp chí khoa học được Hội đồng chức danh giáo sư nhà nước đưa vào danh sách xem xét, tính

điểm cho các bài báo được công bố trong đó. Tất cả các tạp chí KH&CN đều được chuẩn bị bằng công nghệ chế bản điện tử, nghĩa là có bản điện tử của các số tạp chí được xuất bản. Tuy nhiên, số tạp chí KH&CN được xuất bản ở dạng số không nhiều. Khảo sát của tác giả Dương Thị Phương và cộng sự [2019] ở Cục Thông tin KH&CN quốc gia ở 155 tạp chí KH&CN cho thấy chỉ có một (01) tạp chí xuất bản điện tử và 52 (chiếm 33%) tạp chí KH&CN có xuất bản ở cả hai dạng giấy và điện tử (Bảng 2).

Bảng 2. Tổng hợp về dạng xuất bản của các tạp chí KH&CN

Định dạng tài liệu	TC thuộc các trường ĐH, học viện		TC thuộc Bộ/ngành, viện NC&PT		TC thuộc các hội, hiệp hội		Tổng số	
	SL	%	SL	%	SL	%	SL	%
Chỉ ở dạng giấy	48	55	39	75	15	94	102	66
Chỉ ở dạng điện tử	1	1	0	0	0	0	1	1
Ở cả hai định dạng	38	44	13	25	1	6	52	33
Tổng số	87	100	52	100	16	100	155	100

(Nguồn: Dương Thị Phương và cộng sự, 2019)

Một nguồn công bố KH&CN trên các tạp chí KH&CN định dạng số là hệ thống Tạp chí khoa học Việt Nam trực tuyến (VJOL - Vietnam Journals Online) do Cục Thông tin KH&CN quốc gia duy trì. Theo thông tin trên website này, hiện tại hệ thống có 98 tạp chí KH&CN tham gia, cung cấp 46.245 bài viết được cung cấp toàn văn dưới định dạng PDF [VJOL, 2019].

2.2.3. Tài liệu xám

Nguồn tài nguyên thông tin KH&CN không công bố (tài liệu xám - Grey literature) có 2 loại hình quan trọng là: Báo cáo kết quả thực hiện nhiệm vụ KH&CN và luận án tiến sĩ.

- Báo cáo kết quả thực hiện nhiệm vụ KH&CN: Theo quy định tại Thông tư 14/2014/TT-BKHCN ngày 11/6/2019, các báo cáo kết quả thực hiện nhiệm vụ KH&CN sử dụng ngân sách nhà nước phải được thu thập, đăng ký và lưu giữ tại cơ quan có thẩm quyền. Các cá nhân/tổ chức thực hiện nhiệm vụ KH&CN phải nộp bản điện tử báo cáo tổng hợp kết quả thực hiện nhiệm vụ (file hoàn chỉnh, không tách riêng các file chương, mục,...); bản điện tử Báo cáo tóm

tắt kết quả thực hiện nhiệm vụ; bản điện tử phụ lục tổng hợp số liệu điều tra, khảo sát, bản đồ, bản vẽ, ảnh, sách chuyên khảo, bài tạp chí... tài liệu đa phương tiện, phần mềm (nếu có). Bản điện tử phải sử dụng định dạng Portable Document (.pdf) và phải sử dụng phông chữ tiếng Việt - Unicode (Time New Roman) theo tiêu chuẩn Việt Nam (TCVN 6909:2001). Đây là nguồn tài nguyên thông tin KH&CN dạng số rất có giá trị. Cục Thông tin KH&CN quốc gia là cơ quan nhà nước có thẩm quyền về đăng ký, lưu giữ và phổ biến các kết quả nghiên cứu khoa học và phát triển công nghệ. Tính đến nay, Cục Thông tin KH&CN quốc gia đang lưu giữ trên 27.800 báo cáo kết quả thực hiện nhiệm vụ KH&CN các cấp. Ngoài Cục Thông tin KH&CN quốc gia, các cơ quan thông tin trực thuộc các bộ, ngành cũng lưu giữ các kết quả thực hiện nhiệm vụ trong phạm vi mình quản lý, các cơ quan thông tin - thư viện thuộc trường đại học, cao đẳng, cơ quan thông tin thuộc các viện nghiên cứu cũng đều lưu giữ các báo cáo kết quả thực hiện nhiệm vụ KH&CN có sử dụng ngân sách nhà nước.

- Luận án tiến sĩ: Theo quy định, tất cả các nghiên cứu sinh bảo vệ luận án tiến sĩ trên lãnh thổ Việt Nam (người Việt Nam và người nước ngoài) và nghiên cứu sinh người Việt Nam bảo vệ luận án ở nước ngoài đều phải nộp một bản luận án và tóm tắt luận án tại Thư viện Quốc gia Việt Nam. Hiện nay, Thư viện Quốc gia Việt Nam lưu giữ khoảng 29.200 bộ luận án tiến sĩ của người Việt Nam bảo vệ trong nước và nước ngoài, và của người nước ngoài bảo vệ tại Việt Nam, với tổng số trang tài liệu hơn 4,5 triệu trang.

2.2.4. Cơ sở dữ liệu

Cơ sở dữ liệu (CSDL) KH&CN nội sinh cũng là tài nguyên thông tin KH&CN số quan trọng. Khảo sát năm 2011 của Phan Huy Quế và cộng sự [2011] cho thấy, các cơ quan thông tin - thư viện bộ, ngành địa phương xây dựng hơn 140 CSDL nội sinh các loại. Khảo sát gần đây của Cục Thông tin KH&CN quốc gia cho thấy, các đơn vị thuộc bộ, ngành đã xây dựng khoảng 2.620 CSDL; các trường đại học xây dựng khoảng 15.322 CSDL. Tuy nhiên, số lượng biểu ghi là rất hạn chế. Một số CSDL KH&CN nội sinh trong nước có quy mô tương đối lớn là:

- CSDL công bố KH&CN Việt Nam (còn gọi là CSDL tài liệu KH&CN Việt Nam) lưu giữ thông tin về các bài báo KH&CN từ hơn 236 tạp chí KH&CN có tính điểm công trình. Đến tháng 12/2019, CSDL có trên 264.500 biểu ghi thư mục, trong đó khoảng 70% được đính kèm file toàn văn [Cục Thông tin KH&CN quốc gia, 2019].

- CSDL nhiệm vụ KH&CN là CSDL thư mục lớn nhất Việt Nam về các nhiệm vụ KH&CN, lưu giữ thông tin về kết quả thực hiện nhiệm vụ (Báo cáo tổng hợp), nhiệm vụ KH&CN đang tiến hành và kết quả ứng dụng nhiệm vụ KH&CN. Đến tháng

12/2019, CSDL có hơn 34.650 biểu ghi, trong đó có hơn 29.230 biểu ghi về kết quả thực hiện nhiệm vụ KH&CN (Báo cáo tổng hợp), khoảng 4.300 biểu ghi nhiệm vụ KH&CN đang tiến hành (5).

- CSDL luận án tiến sĩ ở Thư viện Quốc gia Việt Nam. Đến tháng 12/2019, Thư viện Quốc gia Việt Nam tạo lập được CSDL luận án tiến sĩ với 20.477 luận án, tổng số 314.239 trang [Thư viện Quốc gia Việt Nam].

2.2.5. Dữ liệu nghiên cứu

Dữ liệu nghiên cứu số là một nguồn tài nguyên thông tin KH&CN số quan trọng. Tuy nhiên, chưa có nghiên cứu đầy đủ về nguồn tài nguyên dữ liệu nghiên cứu này. Về cơ bản, dữ liệu nghiên cứu vẫn được tự lưu giữ tại các cơ quan, đơn vị hoặc bộ phận tạo ra chúng. Theo Nguyễn Thị Kim Lân [2017], nói chung, các trung tâm TT-TV và viện nghiên cứu chưa có hình thức dịch vụ lưu trữ dữ liệu nghiên cứu.

Nghiên cứu về dữ liệu mở ở Việt Nam cho thấy, dữ liệu ở Việt Nam còn hạn chế [Trần Minh và cộng sự, 2017, 2018]. Hiện tại, ở Việt Nam cũng đã có một số hệ thống cung cấp dữ liệu mở như: Cổng dữ liệu mở itrithuc (<http://dulieu.itrithuc.vn>) được xây dựng theo Đề án phát triển Hệ tri thức Việt số hóa. Tính đến hết tháng 10/2018 trên cổng dữ liệu mở của itrithuc có 10.420 tập dữ liệu [Trần Minh và cộng sự, 2018].

2.3. Một số hạn chế của tài nguyên thông tin KH&CN số ở Việt Nam

Tìm hiểu hiện trạng tài nguyên thông tin KH&CN của Việt Nam cho thấy việc xây dựng và phát triển tài nguyên thông tin KH&CN số còn nhiều hạn chế:

- Hầu hết CSDL được xây dựng và phát

5. Số liệu tra cứu từ CSDL của Cục Thông tin KH&CN quốc gia http://sti.vista.gov.vn/Pages/ket-qua-nghien-cuu-moi-nhat.aspx?Type_CSDL=KETQUANHIEMVU

triển như những CSDL độc lập, riêng lẻ, không liên kết tạo thành hệ thống;

- Hầu hết CSDL được xây dựng mới chỉ là CSDL thư mục. Dữ liệu toàn văn của các đối tượng thông tin (bài báo, báo cáo kết quả nghiên cứu, luận án,...) nếu có mới chỉ là ở dạng tệp đính kèm, chưa được đưa vào quản lý, xử lý bên trong hệ thống CSDL. Những CSDL như vậy mới chỉ hỗ trợ việc tra cứu tìm kiếm thông tin (thư mục hoặc bài toàn văn (nếu đính kèm)). Trong kỷ nguyên kỹ thuật số, những dữ liệu toàn văn nếu được đưa vào quản lý, xử lý bên trong CSDL có thể được xử lý tự động; hỗ trợ phân tích thông tin dựa trên các công nghệ số như trí tuệ nhân tạo, dữ liệu lớn,....

- Cấu trúc dữ liệu của các CSDL khác nhau; không có chuẩn cấu trúc dữ liệu được chấp nhận chung (trừ trường hợp các CSDL về sách tuân thủ theo khổ mẫu MARC 21); điều này hạn chế sự liên thông, chia sẻ, kết nối dữ liệu của các CSDL;

- Có sự trùng lặp giữa các CSDL tương tự của các tổ chức (như CSDL về nhiệm vụ KH&CN có thể được nhiều bộ, ngành hoặc địa phương (Sở KH&CN) cùng thực hiện;

- Chưa có nhiều quy định về chuẩn xử lý dữ liệu (như các bộ từ điển từ chuẩn; các bảng phân loại; các mã định danh được thống nhất).

3. MỘT SỐ Ý KIẾN VỀ CHUYỂN ĐỔI SỐ TRONG PHÁT TRIỂN TÀI NGUYÊN THÔNG TIN KHOA HỌC VÀ CÔNG NGHỆ SỐ Ở VIỆT NAM

Phát triển và chia sẻ tài nguyên thông tin KH&CN, nhất là nguồn tài nguyên thông tin số, phục vụ hoạt động thông tin KH&CN trong nước đã được nhiều tác giả đề cập [Nguyễn Hữu Hùng, 2005; Nguyễn Tiến Đức, 2006; Phạm Văn Hùng, 2015; Cao Minh Kiểm và Nguyễn Tuấn Hải, 2013; Cao Minh Kiểm và cộng sự, 2015; Cao Minh Kiểm và Lê Thị Hoa, 2017; Đào Mạnh

Thắng và Trần Thị Hải Yến 2016, 2017;]. Sự quan tâm của Chính phủ và Bộ KH&CN đến phát triển và chia sẻ tài nguyên thông tin KH&CN được cụ thể hóa bằng việc Thủ tướng chính phủ ban hành Quyết định 1258/QĐ-TTg ngày 01/10/2018 phê duyệt đề án “Phát triển nguồn tin KH&CN phục vụ NC&PT đến năm 2025 và định hướng đến năm 2030” và Bộ KH&CN ban hành kế hoạch triển khai thực hiện Đề án [Quyết định số 3999/QĐ-BKHCN]. Trước xu thế của Cách mạng công nghiệp 4.0 và chuyển đổi số, hoạt động thông tin KH&CN cũng cần triển khai thực hiện chuyển đổi số trong phát triển tài nguyên thông tin KH&CN, tác giả đề xuất một số nội dung chuyển đổi số trong phát triển tài nguyên thông tin KH&CN.

3.1. Triển khai thực hiện chuyển đổi số trong xây dựng và phát triển CSDL quốc gia về KH&CN

Tài nguyên thông tin KH&CN số là hạ tầng dữ liệu cốt lõi của hạ tầng thông tin KH&CN, trong đó CSDL quốc gia về KH&CN có vai trò quan trọng. CSDL quốc gia được định nghĩa là tập hợp thông tin của một hoặc một số lĩnh vực kinh tế - xã hội được xây dựng, cập nhật và duy trì đáp ứng yêu cầu truy nhập và sử dụng thông tin của các ngành kinh tế và phục vụ lợi ích công cộng [Luật Công nghệ thông tin 2006]. Tiêu chí xác định CSDL quốc gia gồm [Nguyễn Trọng Khánh, 2016]:

- Lưu trữ thông tin quốc gia, có quy mô lớn, tầm quan trọng ảnh hưởng đến chính trị, sự phát triển kinh tế, xã hội của đất nước;

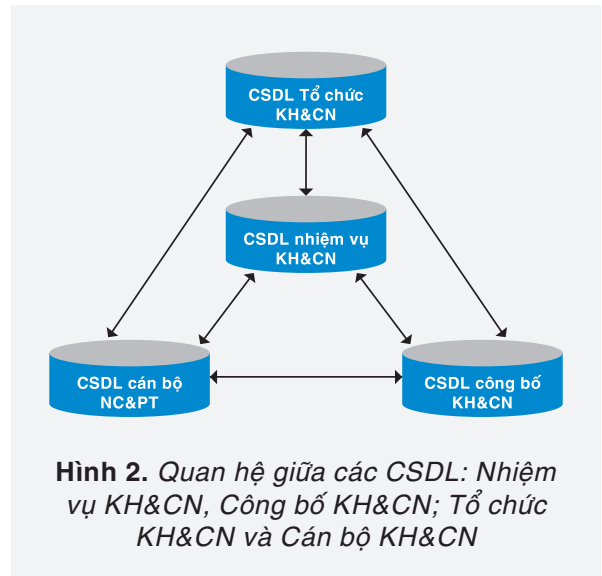
- Dữ liệu được dùng chung, chia sẻ giữa các bộ, ngành, địa phương;

- Có phạm vi, đối tượng, thuộc tính dữ liệu phủ rộng toàn quốc;

- Làm hạ tầng thông tin, tạo nền tảng cho các hệ thống thông tin khác hoạt động và phát triển.

CSDL quốc gia về KH&CN được hiểu là tập hợp thông tin về tiềm lực và kết quả hoạt động KH&CN quốc gia được xây dựng, duy trì và phát triển nhằm cung cấp dữ liệu và thông tin đầy đủ, chính xác, kịp thời về KH&CN, phục vụ hoạch định chính sách, lãnh đạo, quản lý, nghiên cứu khoa học và phát triển công nghệ, giáo dục và đào tạo, sản xuất, kinh doanh [Cao Minh Kiểm 2015]. Nghị định 11/2014/NĐ-CP của Chính phủ về thông tin KH&CN đã quy định CSDL quốc gia về KH&CN là một bộ phận của hạ tầng thông tin KH&CN và bao gồm những CSDL thành phần chính: CSDL về tổ chức KH&CN; CSDL về cán bộ NC&PT; CSDL về nhiệm vụ KH&CN (đang tiến hành, đã kết thúc và kết quả ứng dụng); CSDL về tài liệu sở hữu trí tuệ, các công bố KH&CN và chỉ số trích dẫn; CSDL về công nghệ, công nghệ cao, chuyển giao công nghệ; CSDL về thông tin, thống kê KH&CN; CSDL về doanh nghiệp KH&CN; CSDL thông tin về KH&CN trong khu vực và trên thế giới; Các thông tin khác có liên quan).

Những tìm hiểu về hiện trạng hạ tầng thông tin, thống kê KH&CN cho thấy, hiện tại chúng ta đã có những cơ sở ban đầu cho hệ thống CSDL quốc gia về KH&CN, tuy vậy, những CSDL thành phần của CSDL quốc gia về KH&CN còn phân tán, không được quản lý tập trung. Đề án “Phát triển nguồn tin KH&CN phục vụ NC&PT đến năm 2025 và định hướng đến năm 2030” đã yêu cầu tăng cường xây dựng và hoàn thiện CSDL nhiệm vụ KH&CN và CSDL công bố KH&CN. Hai CSDL (CSDL nhiệm vụ KH&CN và CSDL công bố KH&CN), cùng với CSDL tổ chức KH&CN và CSDL cán bộ NC&PT là những CSDL quan trọng của CSDL quốc gia về KH&CN được đề xuất xây dựng, phát triển và duy trì tập trung tại Cục Thông tin KH&CN. Đây là những CSDL thành phần có mối liên quan chặt chẽ với nhau, trong đó CSDL nhiệm vụ KH&CN có vị trí quan trọng (Hình 2).



Hình 2. Quan hệ giữa các CSDL: Nhiệm vụ KH&CN, Công bố KH&CN; Tổ chức KH&CN và Cán bộ KH&CN

CSDL nhiệm vụ KH&CN Việt Nam là CSDL thư mục lớn nhất Việt Nam về các nhiệm vụ KH&CN nghệ, bao gồm: CSDL nhiệm vụ KH&CN nghệ đang tiến hành, CSDL kết quả thực hiện nhiệm vụ KH&CN và CSDL ứng dụng kết quả thực hiện nhiệm vụ KH&CN. CSDL Nhiệm vụ KH&CN hiện đã được xây dựng, duy trì và quản lý tập trung tại Cục Thông tin KH&CN quốc gia.

CSDL Công bố KH&CN Việt Nam đã được Cục Thông tin KH&CN quốc gia xây dựng từ năm 1987 và hiện là CSDL lớn nhất của Việt Nam về công bố KH&CN đăng tải trên các tạp chí KH&CN trong nước, trong đó có nhiều tạp chí KH&CN được tính điểm công trình. Chúng ta có thể sử dụng kinh phí để mua các CSDL thư mục của quốc tế nhưng không thể mua được CSDL về công bố KH&CN của các nhà nghiên cứu của Việt Nam đăng tải trên các tạp chí KH&CN của Việt Nam.

Dữ liệu cơ bản về tổ chức KH&CN đã có (thông qua các cuộc điều tra thống kê NC&PT, điều tra tiềm lực KH&CN của tổ chức KH&CN do Cục Thông tin KH&CN quốc gia thực hiện, tuy nhiên những dữ liệu này vẫn chưa được tích hợp để tạo thành CSDL hoàn chỉnh về tổ chức KH&CN.

Hiện nay, những CSDL này tuy được đưa chung vào Hệ thống thông tin KH&CN do Cục Thông tin KH&CN quốc gia thực hiện [http://sti.vista.gov.vn/pages/

trang-chu.aspx], và dường như chúng vẫn là các CSDL riêng rẽ, không có sự liên thông, liên kết hữu cơ như một CSDL quốc gia (Hình 3)



Hình 3. Trang chủ Hệ thống thông tin KH&CN do Cục Thông tin KH&CN quốc gia xây dựng

Vì thế, thực tế chúng ta vẫn chưa có CSDL quốc gia về KH&CN đúng nghĩa. Để thực hiện chuyển đổi số trong xây dựng và phát triển CSDL quốc gia về KH&CN cần thiết:

- Hoàn thiện mô hình kiến trúc CSDL quốc gia về KH&CN. Theo Thông tư 14/2014/TT-BKHCN, CSDL quốc gia về KH&CN bao gồm 10 CSDL thành phần, trong đó có 7 CSDL được xây dựng và quản lý tập trung tại Cục Thông tin KH&CN quốc gia và 3 CSDL thành phần khác được xây dựng và quản lý ở tổ chức ngoài Cục Thông tin KH&CN quốc gia (CSDL thông tin sở hữu trí tuệ; CSDL thông tin Tiêu chuẩn - đo lường - Chất lượng và CSDL doanh nghiệp KH&CN). Tuy nhiên, những CSDL này được xây dựng độc lập, trên các nền tảng khác nhau; áp dụng các quy trình và phương tiện

thu nhập dữ liệu khác nhau, không thống nhất; không có một từ điển dữ liệu đặc tả (metadata) dùng chung cho toàn hệ thống. Những công nghệ được sử dụng để xây dựng CSDL thành phần hiện nay chỉ có thể đáp ứng được những yêu cầu đơn giản (tra cứu, tìm kiếm,...) mà chưa có khả năng phân tích dữ liệu; có có khả năng thu thập dữ liệu lớn không có cấu trúc (big data). Do đó, việc xây dựng được mô hình CSDL về KH&CN phù hợp sẽ đảm bảo cho phát triển hệ thống phần mềm phù hợp;

- Xây dựng và hoàn thiện kiến trúc dữ liệu, các quy chuẩn, tiêu chuẩn, phương án kỹ thuật của CSDL quốc gia về KH&CN để đảm bảo sự duy trì và phát triển bền vững, sự minh bạch về cấu trúc thông tin và khả năng cung cấp, chia sẻ thông tin từ CSDL quốc gia về KH&CN. Trước mắt, cần xây

dựng được kiến trúc dữ liệu (từ điển dữ liệu đặc tả) cho một số CSDL có liên quan chặt chẽ và do Cục Thông tin KH&CN quốc gia quản lý (CSDL nhiệm vụ KH&CN; CSDL cán bộ NC&PT, CSDL Tổ chức KH&CN; CSLD công bố KH&CN). Kiến trúc dữ liệu chuẩn sẽ đảm bảo việc thể hiện nội dung, cấu trúc và thuộc tính dữ liệu trong CSDL quốc gia. Việc thống nhất kiến trúc dữ liệu đảm bảo thông tin phải hiểu đúng và thống nhất giữa các bên liên quan khi cập nhật và khai thác CSDL;

- Xây dựng các quy định về mô tả dữ liệu, trình diễn dữ liệu và các quy định khác đảm bảo dữ liệu thống nhất và rõ ràng, đơn nghĩa;

- Xây dựng các chuẩn cho công tác xử lý nội dung trong xây dựng CSDL (các bảng phân loại, các công cụ phục vụ định chỉ mục; các quy tắc mô tả);

- Xây dựng được những hệ thống mã định danh cho các đối tượng thông tin để đảm bảo sự liên kết, kết nối giữa các CSDL (ví dụ định danh cho tổ chức KH&CN; định danh cán bộ nghiên cứu; định danh cho khu vực địa lý (tỉnh/thành phố); định danh cho bài báo,..

- Ứng dụng các phần mềm hiện đại, phục vụ cho thiết kế và xây dựng hệ thống phần mềm CSDL tiên tiến, tính năng cao, tương thích với các chuẩn CSDL quốc tế;

Chúng ta có thể tham khảo và học tập mô hình dữ liệu của Hệ thống Thông tin nghiên cứu châu Âu euroCRIS (European Current Research Information System) [euroCRIS]) để nghiên cứu và xây dựng mô hình kiến trúc dữ liệu cho CSDL quốc gia về KH&CN. euroCRIS đã xây dựng Khổ mẫu Thông tin nghiên cứu chung CERIF (Common European Research Information Format) và khuyến nghị áp dụng cho các hệ thống thông tin nghiên cứu, hỗ trợ chia sẻ và đảm bảo tính liên thông kết nối giữa các hệ thống thông tin nghiên cứu [CERIF].

3.2. Phát triển truy cập mở đến tài nguyên thông tin KH&CN số

Truy cập mở đến kết quả nghiên cứu và phát triển đang là một xu thế và được coi là một phương thức phát triển tài nguyên KH&CN. Ở Việt Nam, vấn đề truy cập mở tài nguyên thông tin thư mục cũng đã bước đầu được đề cập, tuy nhiên mới chủ yếu trong lĩnh vực đào tạo và tài liệu học tập [Đỗ Văn Hùng 2016, 2017]. Cục Thông tin KH&CN Quốc gia đã cho truy cập miễn phí đến thông tin toàn văn của CSDL Công bố KH&CN nhưng mới chỉ áp dụng cho thành viên của Liên hợp thư viện Việt Nam và các thành viên Mạng Nghiên cứu và Đào tạo Việt Nam (VinaREN). Một số lượng tương đối tạp chí KH&CN trong nước có thể được truy cập mở trên hệ thống Tạp chí khoa học Việt Nam trực tuyến VJOL (Vietnam Journal On-line) [Lê Thị Hoa, 2013]. Đại học quốc gia Hà Nội đã xây dựng kho tài liệu số nội sinh (<http://repository.vnu.edu.vn/>). Giảng viên, sinh viên, học viên có thể đọc miễn phí các đầu sách, tài liệu học thuật chất lượng, đa ngành, đa lĩnh vực trong kho tài liệu số này. Thư viện học liệu mở Việt Nam - VOER (<http://voer.edu.vn>) do Chương trình giáo dục mở Việt Nam (Vietnam Open Educational Resources - VOER) phát triển là một kho tài nguyên giáo dục mở cho người Việt sử dụng và truy nhập miễn phí nguồn tài liệu hữu ích phục vụ giảng dạy, học tập và nghiên cứu, cũng như phục vụ cho toàn xã hội.

Truy cập mở đến dữ liệu nghiên cứu liên quan đến vấn đề dữ liệu mở (Open data). Rào cản đối với phát triển truy cập đến dữ liệu nghiên cứu có thể là vấn đề pháp lý, từ chối quản lý và giám sát dữ liệu; quy định về sự riêng tư không cho phép công khai thông tin; sự lo ngại về chất lượng dữ liệu; thiếu sự chuẩn hóa; thiếu chính sách dữ liệu mở; thiếu năng lực công khai dữ liệu hoặc tổn thất về doanh thu hay tính bảo mật [Trần Minh, 2017].

Phát triển và chia sẻ tài nguyên thông tin KH&CN thông qua truy cập mở cần được thúc đẩy. Để tăng cường công tác truy cập mở chúng ta cần:

- Nâng cao nhận thức của các ngành, các cấp, các tổ chức KH&CN về truy cập mở, vai trò và lợi ích của truy cập mở;

- Hoàn thiện khung pháp lý để đảm bảo thực hiện truy cập mở đến kết quả nghiên cứu sử dụng kinh phí từ ngân sách nhà nước; xây dựng và thực hiện chính sách về truy cập mở đến kết quả KH&CN từ các nghiên cứu sử dụng ngân sách nhà nước; Các cơ quan cấp kinh phí cho hoạt động KH&CN nghiên cứu ban hành các quy định về truy cập mở;

- Đảm bảo việc truy cập mở đến tất cả các tạp chí KH&CN do nhà nước cấp kinh phí; Tiếp tục hỗ trợ duy trì và phát triển hệ thống Tạp chí khoa học Việt Nam trực tuyến (VJOL); Nghiên cứu xây dựng kho lưu trữ (repository) về các công bố KH&CN của Việt Nam do các tác giả Việt Nam công bố trên các tạp chí KH&CN quốc tế;

- Tăng cường xây dựng các kho lưu trữ số nội bộ truy cập mở tại các tổ chức KH&CN. Các tổ chức KH&CN (các viện nghiên cứu, trường đại học) cần xây dựng những hệ thống lưu trữ số nội bộ (Institutional repositories) là cơ sở hạ tầng kỹ thuật cần thiết cho lưu trữ và cung cấp truy cập mở đến các kết quả của nghiên cứu sử dụng kinh phí từ ngân sách nhà nước.

- Đẩy mạnh truy cập mở đến dữ liệu nghiên cứu, cần nghiên cứu và ban hành những quy định pháp lý để triển khai thực hiện dữ liệu mở, như: xác định rõ phạm vi dữ liệu mở; quy định về phân loại dữ liệu; tiêu chuẩn dữ liệu mở: như quy tắc đặt tên (các tệp dữ liệu nên chứa thông tin về nhà cung cấp, thời gian cung cấp, tên tệp, kiểu dữ liệu), khuôn dạng dữ liệu (các tệp dữ liệu mở phải sẵn có ở dạng máy có thể đọc được, khuyến nghị dưới dạng định dạng mở như: CSV, JSON, XML, etc.); mô tả (tạo lập siêu dữ liệu);...

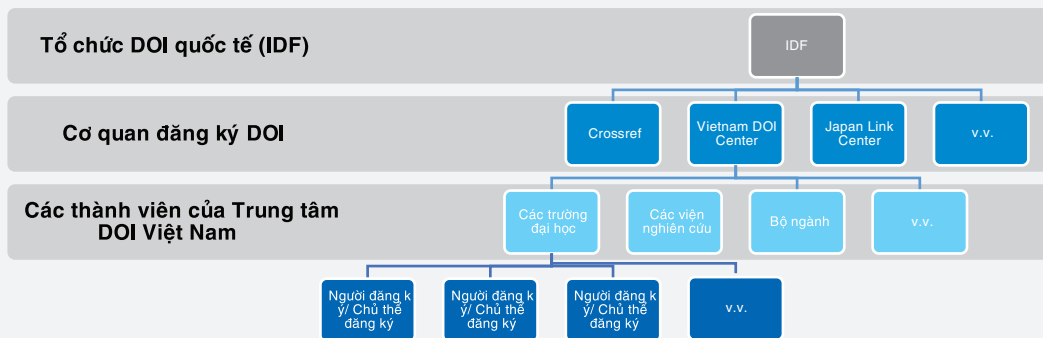
3.3. Triển khai ứng dụng mã định danh số cho tài nguyên thông tin KH&CN số

Trong môi trường số, việc duy trì đảm bảo truy cập đến tài nguyên số là rất quan trọng. Định danh duy nhất bền vững (Persistent Unique Identifier - PUID) cung cấp một phương thức định danh ổn định lâu dài cho các đối tượng số. Mã định danh bền vững (persistent identifier, PI or PID) được hiểu là một tham chiếu tồn tại lâu dài đến một tài liệu, tệp tin, trang web hoặc đối tượng khác [Wiki]. Một định danh bền vững được đảm bảo để quản lý và duy trì không thay đổi trong một thời gian dài [Australian National Data Services]. Những mã định danh này được toàn cầu hóa, tiêu chuẩn hóa và được sử dụng rộng rãi trong môi trường kỹ thuật số và có thể cung cấp thông tin về đối tượng, bất kể đối tượng nằm ở đâu. Việc sử dụng định danh duy nhất bền vững giúp cho người dùng tin chắc chắn rằng mình sẽ lấy được đúng tài nguyên (công trình) mình cần ngay cả khi vị trí thực trên Web của nó đã thay đổi nhờ việc ánh xạ tới một bộ định vị hoặc bộ định vị cập nhật, tạo điều kiện cho việc truy cập vào (các) biểu hiện vật lý của tài nguyên. Hiện nay, tồn tại nhiều loại định danh cho đối tượng số như: Định danh số cho đối tượng (DOI), định danh Khóa Tài nguyên lưu trữ ARK (Archival Resources Key identifier) [ARK], Handles [Sun S. , Lannom L., Boesch B., 2003], mã định danh Nhà nghiên cứu và Cộng sự mở ORCID [Open Researcher and Contributor ID],...

Đối với tài nguyên số thông tin KH&CN, mã định danh DOI được sử dụng rất rộng rãi trên toàn cầu [Dương Thị Phương và cộng sự, 2019]. DOI được sử dụng cho tài liệu số và dữ liệu số. Hiện nay ở Việt Nam, đã có một số tổ chức áp dụng DOI cho tạp chí KH&CN do mình xuất bản. Khảo sát của Dương Thị Phương và cộng sự [2019] cho thấy trong số 155 tạp chí được hỏi, có 42 tạp chí (chiếm tỷ lệ 27%) đã ứng dụng DOI và 113 tạp chí (chiếm 73%) chưa ứng dụng DOI (Hình 14). Theo khảo sát, tất cả

42 tạp chí đã ứng dụng DOI đều áp dụng phương thức đăng ký trực tiếp DOI thông qua tổ chức CrossRef. Việc triển khai áp dụng DOI cho tài liệu số KH&CN và dữ liệu nghiên cứu là cần thiết. Đề xuất của nhóm nghiên cứu của Dương Thị Phương và cộng sự [2019] về mô hình cấp và quản

lý mã định danh DOI cho tài nguyên thông tin KH&CN ở Việt Nam, trong đó hình thành Trung tâm DOI Việt Nam tại Cục Thông tin KH&CN quốc gia là rất đáng quan tâm [Hình 4]. Chúng ta cần nhanh chóng nghiên cứu triển khai vào thực tế mô hình này.



Hình 4. Mô hình cấp và quản lý DOI quy mô quốc gia tại Việt Nam
(Nguồn: Dương Thị Phương và cộng sự, 2019)

4. KẾT LUẬN

Trong bối cảnh của cuộc Cách mạng công nghiệp 4.0, chuyển đổi số đang diễn ra nhanh chóng ở các nước phát triển trên thế giới. Ở Việt Nam, vấn đề chuyển đổi số đã trở thành yêu cầu cấp bách đối với mọi lĩnh vực kinh tế - xã hội. Tầm quan trọng của chuyển đổi số đã được thể hiện ở việc Chính phủ giao nhiệm vụ cho Bộ Thông tin và Truyền thông xây dựng Đề án chuyển đổi số quốc gia.

Lĩnh vực TT-TV KH&CN liên quan chặt chẽ đến tài nguyên thông tin KH&CN, trong đó có tài nguyên thông tin KH&CN số cũng cần nghiên cứu triển khai thực hiện chuyển đổi số. Mặc dù thời gian vừa qua chúng ta đã xây dựng được nguồn tài nguyên thông tin KH&CN số có giá trị phục vụ cho phát triển kinh tế xã hội, tuy nhiên nguồn tài nguyên số này còn nhiều hạn chế cần được chú trọng phát triển trong bối cảnh chuyển đổi số.

Một trong những nội dung thực hiện chuyển đổi số trong hoạt động thông tin

KH&CN là cần tập trung xây dựng CSDL quốc gia về KH&CN. Cần nghiên cứu xây dựng và hoàn thiện mô hình kiến trúc CSDL quốc gia về KH&CN; xây dựng và ban hành kiến trúc dữ liệu chung cho hệ thống CSDL quốc gia về KH&CN, các quy định pháp lý, kỹ thuật, công nghệ cho CSDL quốc gia về KH&CN.

Để đảm bảo hạ tầng dữ liệu cho chuyển đổi số, cần tiếp tục quan tâm phát triển nguồn tài nguyên số KH&CN, triển khai thực hiện truy cập mở đến tài nguyên thông tin KH&CN. Phát triển dữ liệu mở cần được chú trọng trong thời gian tới. Đồng thời chúng ta cần nghiên cứu áp dụng mã định danh số cho đối tượng (DOI), hoàn thiện mô hình cấp và quản lý DOI ở quy mô quốc gia.

TÀI LIỆU THAM KHẢO

- 1. ARK - Archival Resource Key Identifiers.** http://n2t.net/e/ark_ids.html
- 2. Australian National Data Services.** Persistent identifier: awareness level. <https://www.ands.org.au/guides/persistent-identifiers-awareness>

3. **Bộ TT&TT (2019)**. Đề án Chuyển đổi số quốc gia: Dự thảo 1.0, ngày 04/4/2019. Tải xuống ngày http://mic.gov.vn/Upload_Moi/DuThaoVanBan/PL03-DU-THAO-DE-AN-CHUYEN-DOI-SO-QG-VER-1.0.pdf
4. **Cao Minh Kiểm (2015)**. Một số nội dung chủ yếu cho hiện đại hóa hạ tầng thông tin, thống kê KH&CN Việt Nam. TC Thông tin - Tư liệu, số 2/2015, tr. 11-17.
5. **Cao Minh Kiểm (2017)**. Chính sách truy cập mở đến kết quả nghiên cứu sử dụng kinh phí công trên thế giới và đề xuất với Việt Nam. Hội thảo “Phát triển và chia sẻ nguồn tài liệu số nhằm hỗ trợ hoạt động của Quốc hội vì sự phát triển bền vững tại Việt Nam”, 14/4/2017, Hà Nội, Việt Nam.
6. **Cao Minh Kiểm (2018)**. Truy cập mở đến kết quả nghiên cứu sử dụng kinh phí công - tìm hiểu kinh nghiệm trên thế giới và đề xuất đối với Việt Nam. Hội thảo “Số hóa tri thức KH&CN đáp ứng yêu cầu của Cách mạng công nghiệp 4.0”, ngày 19/12/2018, Hà Nội, Việt Nam.
7. **Cao Minh Kiểm và Nguyễn Tuấn Hải (2013)**. Chia sẻ tài nguyên thông tin trên mạng Nghiên cứu và đào tạo Việt Nam (VinaREN). TC Thông tin và Tư liệu: Phụ trương chuyên đề, 2013. Tr.46-52
8. **CERIF**. <https://www.eurocris.org/cerif/main-features-cerif>
9. **Cục Thông tin KH&CN quốc gia**. CSDL Công bố KH&CN. http://sti.vista.gov.vn/Pages/danh-sach-tai-lieu-khcn-moi-nhat.aspx?Type_CSDL=TAILIEUKHCN
10. **Cục Tin học hóa, Bộ Thông tin và Truyền thông (2018)**. Tổng quan chuyển đổi số tại Việt Nam. Bài trình bày tại Hội thảo “Chuyển đổi số: cơ hội và thách thức”, do Hội Truyền thông số Việt Nam tổ chức tại Hà Nội, ngày 19/11/2019.
11. **Đào Mạnh Thắng và Trần Thị Hải Yến (2016)**. Định hướng quốc gia về phát triển nguồn tin khoa học và công nghệ. TC Thông tin và Tư liệu, số 5/2016. Tr.3-12.
12. **Đào Mạnh Thắng và Trần Thị Hải Yến (2017)**. Phát triển nguồn tin khoa học và công nghệ trong thời đại công nghệ số. TC Thư viện Việt Nam, số 1/2017. Tr.12-21
13. **Dương Thị Phương và cộng sự (2019)**. Nghiên cứu cơ sở lý luận và thực tiễn để ứng dụng “Mã định danh tài liệu số (Digital Object Identifier - DOI) đối với tài nguyên số của Việt Nam”. Báo cáo đề tài nghiên cứu cấp bộ. H: Cục Thông tin KH&CN quốc gia, 2019 (Bản bảo vệ cấp cơ sở).
14. **EC. Directorate-General for Research & Innovation, 2016**. Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/h/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf. Truy cập 23/11/2016
15. **euroCRIS**. <https://www.eurocris.org/what-eurocris>
16. **Gartner Glossary**. <https://www.gartner.com/en/information-technology/glossary/digitalization>
17. **Gartner IT Glossary**. <https://www.gartner.com/it-glossary/digital-transformation>
18. **Hồ Tú Bảo (2018)**. Thời chuyển đổi số, trí tuệ nhân tạo và khoa học dữ liệu. [http://www.jaist.ac.jp/~bao/VNAlectures/DSlop1HoBao\(M3\).pdf](http://www.jaist.ac.jp/~bao/VNAlectures/DSlop1HoBao(M3).pdf)
19. **IFLA (2002)**. Guidelines for digitization projects for collections and holdings in the public domain, particularly those held by libraries and archives. March 2002.
20. **Igniger Anna (2017)**. Difference between Digitization, Digitalization and Digital Transformation. <https://www.coresystems.net/blog/difference-between-digitization-digitalization-and-digital-transformation#>
21. **i-SCOOP**. Digital transformation. <https://www.i-scoop.eu/digital-transformation/>
22. **Luật Công nghệ Thông tin - Luật số 67/2006/QH11 ngày 29/6/2006** <https://thuvienphapluat.vn/van-ban/Cong-nghe-thong-tin/Luat-cong-nghe-thong-tin-2006-67-2006-QH11-12987.aspx>
23. **Luật Thư viện 2019 - Luật số 46/2019/QH14**. <https://thuvienphapluat.vn/van-ban/bo-may-hanh-chinh/Luat-Thu-vien-2019-398157.aspx>
24. **National Academy of Sciences (US), National Academy of Engineering (US) and Institute of Medicine (US) Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age (2009)**. Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age. Washington (DC): National Academies Press (US); 2009. <https://www.ncbi.nlm.nih.gov/books/NBK215264/pdf/>

- Bookshelf_ NBK215264.pdf. Truy cập ngày 23/11/2016
- 25. Nguyễn Hữu Hùng (2005).** Vấn đề phát triển và chia sẻ nguồn tin số hóa tại Việt Nam. Thông tin KH&CN Quảng Bình, số 4, 2005.Tr.46-49,51
- 26. Nguyễn Thị Kim Lân (2017).** Quản lý dữ liệu nghiên cứu: Một cách tiếp cận quản trị dữ liệu cho các thư viện Việt Nam. Tạp chí Thư viện Việt Nam. - 2017. - Số 3. - Tr. 19-24,42. <http://nlv.gov.vn/nghep-vu-thu-vien/quan-ly-du-lieu-nghien-cuu-mot-cach-tiep-can-quan-tri-du-lieu-cho-cac-thu-vien-viet-nam.html>
- 27. Nguyễn Tiến Đức.** Bàn về tạo lập và chia sẻ nguồn tin số hóa đối với các cơ quan thông tin KH&CN địa phương. TC Thông tin và Tư liệu. số 1/2006. Tr.11-17.
- 28. Nguyễn Trọng Khánh (2016).** Cơ sở dữ liệu quốc gia: Nền tảng phát triển Chính phủ điện tử. Bài trình bày tại Hội thảo Xây dựng và phát triển CSDL quốc gia về KH&CN, đổi mới sáng tạo. Do Cục Thông tin KH&CN quốc gia tổ chức ngày 6 tháng 10 năm 2016, tại Hà Nội.
- 29. Open Researcher and Contributor ID.** <https://orcid.org/about>
- 30. Phạm Văn Hùng (2015).** Phát triển nguồn tin KH&CN trong giai đoạn mới: Thách thức, cơ hội và giải pháp. TC Thông tin và Tư liệu, số 6/2015, tr.31-37.
- 31. Phan Huy Quế, et al. (2011).** Báo cáo tổng quan về kết quả khảo sát một số nội dung hoạt động của cơ quan thông tin thư viện. Đề tài “Nghiên cứu cơ sở lý luận và thực tiễn phát triển công tác thông tin, thư viện và thống kê KH&CN Việt Nam đến năm 2020”, Chủ nhiệm đề tài: ThS Phan Huy Quế.
- 32. Quyết định số 1285/QĐ-TTg** ngày 01/10/2018 của Thủ tướng Chính phủ phê duyệt Phát triển nguồn tin KH&CN phục vụ NC&PT đến năm 2025 và định hướng đến năm 2030 <https://luatvietnam.vn/khoa-hoc/quyet-dinh-1285-qd-ttg-2018-de-an-phat-trien-nguon-tin-khoa-hoc-va-cong-nghe-167560-d1.html>
- 33. Quyết định số 3999/QĐ-BKHCN** ban hành Kế hoạch của Bộ KH&CN triển khai Quyết định số 1285/QĐ-TTg của Thủ tướng Chính phủ phê duyệt Đề án “Phát triển nguồn tin KH&CN phục vụ nghiên cứu khoa học và phát triển công nghệ đến năm 2025, định hướng 2030”.
- 34. Sun S. , Lannom L., Boesch B., 2003.** Handles system overview. <http://www.rfc-editor.org/rfc/rfc3650.txt>
- 35. The DOI system.** <https://www.doi.org/>
- 36. Thông tư 14/2014/TT-BKHCN** ngày 11/6/2014 Quy định về thu thập, đăng ký, lưu giữ và công bố thông tin về nhiệm vụ KH&CN. <https://thuvienphapluat.vn/van-ban/Cong-nghe-thong-tin/Thong-tu-14-2014-TT-BKHCN-thu-thap-dang-ky-luu-giu-cong-bo-thong-tin-nhiem-vu-khoa-hoc-cong-nghe-238285.aspx>
- 37. Thư viện quốc gia Việt Nam.** CSDL toàn văn luận án tiến sỹ. <http://luanan.nlv.gov.vn/>. truy cập tháng 12/2019.
- 38. Trần Minh, 2017.** Báo cáo tổng quan về dữ liệu mở. Viện Công nghiệp phần mềm và nội dung số Việt Nam. <http://www.nisci.gov.vn/?mdocsfile=806>
- 39. Trần Minh, Đỗ Công Anh, Tô Thị Thiên Ngân, Lương Thị Thu Hà, Chu Thị Thắm, Nguyễn Thị Yến và Đỗ Hoàng Anh Tuấn (2018).** Nghiên cứu cơ sở khoa học và thực tiễn để định hướng chính sách khuyến khích và thúc đẩy dữ liệu mở Việt Nam. Báo cáo đề tài cấp bộ, 2018. 204 tr.
- 40. VJOL (2019).** <http://www.vjol.info/index.php/index/index>. truy cập 12/2019.
- 41. Wiki. Persistent Identifier.** https://en.wikipedia.org/wiki/Persistent_identifier
- 42. Wikipedia. Digital transformation.** https://en.wikipedia.org/wiki/Digital_transformation . Truy cập 22/7/2019.
- 43. МОДЕЛЬНЫЙ ЗАКОН О НАУЧНО-ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ.** Принят на пятнадцатом пленарном заседании Межпарламентской Ассамблеи государств - участников СНГ (постановление № 15-10 от июня 2000 года). “сведения о документах и фактах, получаемых в ходе научной, научно-технической, инновационной и общественной деятельности”
- (Ngày Tòa soạn nhận được bài: 10-11-2019; Ngày phản biện đánh giá: 4-12-2019; Ngày chấp nhận đăng: 15-01-2020).*