

DỮ LIỆU MỞ TRONG BỐI CẢNH CÁCH MẠNG CÔNG NGHIỆP 4.0

Nguyễn Tuấn Anh

Bộ Thông tin và Truyền thông

Dữ liệu mở (DLM) có vai trò quan trọng trong nền kinh tế số và bối cảnh cách mạng công nghiệp 4.0. Ở Việt Nam hiện nay đã hình thành một số DLM như: Cổng DLM quốc gia, truy cập mở thông tin về khoa học và công nghệ (KH&CN), tài nguyên giáo dục mở... Tuy nhiên, việc cung cấp DLM còn rất nhiều rào cản, khó khăn, nhất là việc từ chối quản lý và giám sát dữ liệu, quy định về sự riêng tư không thể công khai thông tin. Chính vì vậy, cần có sự quan tâm chỉ đạo mạnh mẽ từ Chính phủ tới các bộ, ngành, địa phương để triển khai thực hiện đưa dữ liệu vào khai thác sử dụng, chia sẻ cho xã hội, đem lại hiệu quả kinh tế. Bên cạnh đó, cần xây dựng khung pháp lý, hướng dẫn và chính sách khuyến khích thúc đẩy thực thi DLM.

DLM và vai trò trong bối cảnh cách mạng công nghiệp 4.0

Khái niệm

DLM là các dữ liệu điện tử hoặc dữ liệu số. Bất kỳ dữ liệu nào được tạo ra, lưu trữ, thu thập, phát hành bởi Nhà nước, các cơ quan chuyên trách, chính quyền địa phương trong các hoạt động hành chính hoặc liên quan đến các quyết định đưa ra của mình đều có giá trị sử dụng không những cho cơ quan, tổ chức tạo lập ra nó mà còn có thể mở ra, chia sẻ cho các cơ quan đơn vị khác hoặc cộng đồng sử dụng. Dữ liệu có thể bao gồm những thông tin địa lý, môi trường, sức khỏe cộng đồng, giáo dục, những thống kê định kỳ, thư mục, các danh sách, các số liệu về ngân sách đầu tư, chi tiêu... Chúng được lưu trữ dưới rất nhiều dạng khác nhau và được giữ trong các kho dữ liệu riêng biệt của mỗi cơ quan, tổ chức. Gần như không tồn tại các liên kết giữa những kho dữ liệu này.

Như vậy, DLM là dữ liệu có thể được bất kỳ ai tự do sử dụng, sử dụng lại, phân phối lại, chỉ yêu cầu phải ghi nhận nguồn và chia sẻ tương tự. Tính năng quan trọng của

DLM là khả năng tiếp cận và sẵn sàng sử dụng như là các yêu cầu về dữ liệu mà máy có thể đọc được. Các tính năng chính của sự mở là: tính sẵn sàng và sự truy cập; tái sử dụng và phân phối lại; sự tham gia toàn cầu.

Về mặt kỹ thuật, công bố dữ liệu trên một cổng DLM tương đối giống với việc công bố các tài liệu trên một trang web. Tuy nhiên, về bản chất lại hoàn toàn khác nhau. Mục tiêu của việc công bố thông tin trên trang web của Chính phủ là nhằm thông báo thông tin trực tiếp cho người dân và các trang web này được thiết kế cho đối tượng sử dụng là người dân. Cổng DLM lại được thiết kế để lưu trữ các tập DLM với hai đặc điểm cốt lõi là: i) Định dạng: tập hợp DLM được xuất bản ở định dạng mở về kỹ thuật, tức là được sử dụng ở định dạng không độc quyền và có thể được dùng bởi các chương trình máy tính; ii) Giấy phép: tập DLM được đăng cùng giấy phép (điều khoản sử dụng), cho phép sử dụng lại thông tin cho mục đích thương mại và phi thương mại. Hai đặc điểm cốt lõi này tạo ra sự khác biệt với các văn bản thông thường đăng trên các trang web

của Chính phủ, cơ quan nhà nước, địa phương hiện nay.

Lợi ích của DLM

- DLM cho phép Chính phủ điều hành, tạo ra sự sáng tạo trong kinh doanh, dịch vụ cung cấp giá trị xã hội và thương mại: là chìa khóa để nâng cao chất lượng dịch vụ và cuộc sống; đẩy nhanh tiến độ lan truyền các dịch vụ kỹ thuật số trên web và nền tảng di động; trao quyền và thu hút sự tham gia của người dân; là đầu vào cho nghiên cứu và giáo dục.

- Giúp Chính phủ cải thiện được tính minh bạch và công khai vì người dân luôn có được các thông tin cần thiết để có thể so sánh và đối chiếu; cho phép tiếp cận thông tin, dữ liệu chính là tạo ra điều kiện thuận lợi nhất cho người dân thực hiện quyền dân chủ và tham gia quản lý.

- Giúp giảm chi phí và tăng hiệu quả hoạt động cho các cơ quan nhà nước, các tổ chức, doanh nghiệp và công dân.

- Các dữ liệu được công khai sẽ được xử lý, phân tích bởi các trường đại học, nhà khoa học, công ty, nhà

phát triển ứng dụng, nhà báo... để tạo ra các ứng dụng, dịch vụ mới có giá trị cao, ví dụ như các ứng dụng về du lịch, giao thông, tìm kiếm dịch vụ, từ đó tạo động lực cho sự phát triển kinh tế và khoa học.

- Các cơ quan, tổ chức, doanh nghiệp và công dân có thể sử dụng dữ liệu của Chính phủ, cải thiện kết quả nghiên cứu, cải tiến công việc.

Trên thực tế, DLM mang lại rất nhiều lợi ích, tạo động lực cho sự phát triển kinh tế và khoa học thông qua sự thuận tiện hơn khi tiếp cận các thông tin hợp nhất từ nhiều kho dữ liệu khác nhau trong nhiều lĩnh vực khác nhau, từ giao thông, xây dựng, y tế, giáo dục đến các thông tin liên quan đến hoạt động của cộng đồng doanh nghiệp. Việc hệ thống hóa, chuẩn hóa và chia sẻ dữ liệu công và bán công sẽ giúp nâng cao hiệu quả của việc xây dựng chính sách và quản lý hành chính công, cũng như sự hợp tác giữa các tổ chức nhà nước. Thêm vào đó, việc công bố các dữ liệu công theo lộ trình nhất định sẽ làm tăng niềm tin của người dân vào bộ máy hành chính, qua đó nâng cao việc huy động nguồn lực xã hội vào các hoạt động kinh tế, văn hóa và xã hội.

DLM được xem như “tài sản” của nền kinh tế thể hiện qua các trường hợp ứng dụng dữ liệu mạnh mẽ trong thời gian gần đây như Uber, Grab hay một số loại hình dịch vụ đặt phòng khách sạn toàn cầu. Đây là những mô hình kinh doanh mới, khá hiệu quả nhờ khai thác nguồn DLM. Cung cấp cơ sở DLM đang là xu hướng tại các nước phát triển và là một trong các chỉ tiêu đánh giá mức độ phát triển chính phủ điện tử của Liên hợp quốc đối với các nước.

Trong nền kinh tế cạnh tranh công bằng, cộng đồng khởi nghiệp rất cần những cơ sở dữ liệu thông

tin mở, chính xác liên quan đến nhiều lĩnh vực sẵn có để đưa ra quyết định kinh doanh cho riêng mình mà không mất công tìm kiếm những thông tin cơ bản ban đầu. Hiện nay, ở Việt Nam vẫn còn thiếu cơ sở dữ liệu hỗ trợ cho cộng đồng khởi nghiệp nên nhiều doanh nghiệp khởi nghiệp vẫn còn loay hoay tìm kiếm dữ liệu, thông tin liên quan đến lĩnh vực mình kinh doanh.

Trong bối cảnh cách mạng công nghiệp 4.0 và hướng đến mục tiêu phát triển các “thành phố thông minh” thì nhu cầu giới thiệu hoạt động và khai thác DLM của tổ chức hành chính, người dân và doanh nghiệp càng trở nên bức thiết nhằm phát triển chính quyền điện tử, phát triển kinh tế, tạo hệ sinh thái khởi nghiệp. Lâu nay, các cơ sở dữ liệu hiện hữu vẫn còn nằm rải rác ở các quận/huyện, sở/ngành, chưa được tập trung. Do vậy, việc tích hợp các nguồn dữ liệu về một đầu mối, làm cơ sở dữ liệu dùng chung hoàn chỉnh là rất quan trọng để các sở, ban, ngành, doanh nghiệp, người dân tham khảo một cách thống nhất.

Thực trạng DLM tại Việt Nam

Ở Việt Nam, Chính phủ điện tử cũng đã được giới thiệu và đưa vào chương trình phát triển trọng điểm từ năm 2001. Việt Nam đang trong giai đoạn đầu của số hoá và mạng hoá thông tin trong Chính phủ cũng như ở cấp độ địa phương. Đó là một chương trình dài hơi, cần có sự đầu tư tập trung và giải pháp hợp lý, phù hợp với điều kiện cụ thể của đất nước. Việt Nam cũng có một số thuận lợi nhất định khi người dân đang dần có ý thức và thói quen về việc sử dụng công nghệ nói chung và Internet nói riêng để cập nhật, nắm bắt thông tin và theo dõi các hoạt động của Nhà nước. Hiện tại, ở Việt Nam có thể kể đến một số cơ sở DLM như: thông tin về KH&CN, giáo dục và cổng DLM quốc gia.

Về truy cập mở thông tin KH&CN: ngày 17/5/2017, Cục Thông tin KH&CN quốc gia (Bộ KH&CN) đã khai trương Điểm truy cập mở thông tin KH&CN với mục đích tăng cường tính minh bạch của thông tin KH&CN sử dụng ngân sách nhà nước, đưa kết quả nghiên cứu tới công chúng và nâng cao vai trò của KH&CN đối với sự phát triển kinh tế - xã hội. Điểm truy cập mở thông tin KH&CN quốc gia lưu trữ 380.000 đầu sách khoa học và hơn 17.000 đầu sách chuyên khảo, bao gồm nhiều loại hình từ sách khoa toàn thư, cẩm nang, sổ tay tra cứu, đến các từ điển chuyên ngành, tạp chí tóm tắt. Ngoài ra còn có 7.700 tạp chí thuộc các lĩnh vực KH&CN, chủ yếu là tạp chí quốc tế có hệ số ảnh hưởng cao và 23.000 báo cáo kết quả nghiên cứu cấp quốc gia, cấp bộ, cấp tỉnh. Điểm truy cập mở này còn cung cấp tài liệu điện tử. Mọi người có thể truy cập toàn văn 220.000 bài báo khoa học của Việt Nam, đọc nội dung chính của 22.000 nhiệm vụ KH&CN sử dụng ngân sách nhà nước và toàn văn 40 triệu tài liệu công bố quốc tế từ các cơ sở dữ liệu KH&CN nổi tiếng thế giới như ScienceDirect, Springer Nature, ISI-Web of Knowledge hay Scopus...

Tài nguyên giáo dục mở: trang Thư viện học liệu mở của Việt Nam (<http://voer.edu.vn>) được thực hiện bởi Chương trình giáo dục mở Việt Nam (VOER), hỗ trợ bởi Quỹ Việt Nam (the Vietnam Foundation), với mục tiêu xây dựng kho tài nguyên giáo dục mở cho người Việt sử dụng và truy cập miễn phí nguồn tài liệu hữu ích phục vụ giảng dạy, học tập và nghiên cứu, cũng như phục vụ cho toàn xã hội. Tính đến nay, trang VOER có khoảng 22.341 tài liệu, gần 600 tuyển tập gồm các tài liệu học tập, bài giảng, danh mục tài liệu tham khảo, bài tập về nhà, bài thi, bài thí nghiệm, các tuyển tập, tạp chí... được biên soạn từ 10.903



tác giả. Nguồn tài nguyên học liệu mở gồm các tài liệu tiếng Anh và tiếng Việt của nhiều lĩnh vực.

Cổng DLM quốc gia: đây là cổng dữ liệu do Bộ Thông tin và Truyền thông xây dựng. Hiện cổng dữ liệu này cập nhật đến thời điểm hiện tại là 10.425 bộ dữ liệu, được chia ra thành các lĩnh vực: xã hội: 9.986 bộ dữ liệu, công nghệ: 117, lao động: 108, giáo dục: 97, môi trường: 48, y tế: 17, địa phương: 14, tài chính: 11, năng lượng: 9, cơ sở hạ tầng: 8, kinh tế - thương mại: 8, nông nghiệp: 2. Chia theo cơ quan cung cấp dữ liệu cho Cổng DLM quốc gia thì Đại học Quốc gia Hà Nội có số lượng tài liệu lớn nhất (10.045 bộ dữ liệu), số tài liệu do các bộ/ngành cung cấp còn rất hạn chế.

Có thể khẳng định, mặc dù việc cung cấp DLM Chính phủ đã mang lại lợi ích thiết thực cho người dân, doanh nghiệp, các tổ chức và mang lại lợi thế cạnh tranh cho quốc gia, tuy nhiên việc xây dựng và cung cấp DLM ở Việt Nam còn bị hạn chế bởi những trở ngại, rào cản chính như: từ chối việc quản lý và giám sát dữ liệu, quy định về sự riêng tư không thể công khai thông tin; sự sợ hãi và lo lắng khi công khai tài liệu kém chất lượng; thiếu sự chuẩn hóa và chính sách DLM; thiếu năng lực công khai dữ liệu hoặc tổn thất về doanh thu hay tính bảo mật.

Năm 2019, Ngân hàng thế giới cũng đã công bố báo cáo đánh giá về mức độ sẵn sàng DLM. Theo đó thì nhiều tiêu chí còn ở mức độ sẵn sàng trung bình, mặc dù các tiêu chí đó được xác định ở tầm quan

trọng 3 sao (***)*. Đánh giá cụ thể các tiêu chí như sau: (1) Lãnh đạo cấp cao (có sự chủ động và tích cực tham gia, hỗ trợ nhiệt tình của lãnh đạo cấp cao đối với sáng kiến DLM) có mức độ quan trọng *** và mức độ sẵn sàng trung bình; (2) Khung chính sách/pháp lý (tác động của các luật và chính sách hiện hành về phổ biến dữ liệu, bảo vệ thông tin cá nhân và các điều khoản sử dụng hiện tại) có mức độ quan trọng ** và mức độ sẵn sàng trung bình thấp; (3) Cấu trúc, thể chế, trách nhiệm và năng lực của cơ quan quản lý nhà nước (năng lực của các cơ quan để quản lý và phổ biến dữ liệu, điều phối các tiêu chuẩn, quy trình và giải quyết các rào cản thủ tục) có mức độ quan trọng ** và mức độ sẵn sàng trung bình; (4) Chính sách quản lý, quy trình thủ tục và dữ liệu sẵn có của Chính phủ (liệu chính sách hiện tại có tạo điều kiện truy cập dữ liệu hay không và các tập dữ liệu quan trọng đã có sẵn hay có thể có sẵn hay không) có mức độ quan trọng ** và mức độ sẵn sàng trung bình; (5) Nhu cầu đối với DLM (bộ dữ liệu nào đã được yêu cầu hoặc sử dụng và những cộng đồng nào có thể hưởng lợi từ DLM) có mức độ quan trọng *** và mức độ sẵn sàng trung bình; (6) Khả năng tham gia của xã hội và năng lực khai thác DLM (năng lực của các tổ chức ngoài nhà nước và công chúng tham gia với khu vực công như là đối tác và nhà đổi mới) có mức độ quan trọng

1 sao () thể hiện mức độ quan trọng thấp nhất, 3 sao (***) thể hiện mức độ quan trọng cao nhất.

** và mức độ sẵn sàng trung bình cao; (7) Nguồn tài chính triển khai chương trình DLM (nguồn lực sẵn có trực tiếp và gián tiếp để hỗ trợ DLM) có mức độ quan trọng * và mức độ sẵn sàng trung bình cao; (8) Cơ sở hạ tầng công nghệ và năng lực kỹ thuật quốc gia (năng lực và kỹ năng công nghệ thông tin của đội ngũ chuyên gia, công ty công nghệ thông tin và người dân) có mức độ quan trọng ** và mức độ sẵn sàng cao.

*
* *

Từ thực trạng và đánh giá về mức độ sẵn sàng DLM cho thấy, việc cung cấp DLM còn rất nhiều rào cản, khó khăn cần phải vượt qua, trong đó việc từ chối quản lý và giám sát dữ liệu, quy định về sự riêng tư không thể công khai thông tin là trở ngại chính, do vậy cần có sự quan tâm chỉ đạo mạnh mẽ từ Chính phủ tới các bộ, ngành, địa phương để triển khai thực hiện đưa dữ liệu vào khai thác sử dụng, chia sẻ cho xã hội, đem lại hiệu quả kinh tế. Đây là yếu tố quan trọng quyết định tới sự thành công của việc xây dựng và cung cấp DLM. Bên cạnh đó, dẫn hoàn thiện chủ trương, cơ cấu thể chế, khung pháp lý, mô hình, giải pháp quản lý, chính sách thúc đẩy, công nghệ triển khai DLM, trong đó, việc xây dựng chính sách DLM là vấn đề lớn, có tính pháp lý cao cần có sự tham gia của các bộ, ngành có liên quan. Ở Việt Nam, phần lớn dữ liệu là do các bộ, ngành, địa phương quản lý. Để lựa chọn danh sách các dữ liệu chuyển đổi thành DLM đạt các tiêu chuẩn quốc tế cần phải có quy định pháp lý, hướng dẫn và chính sách khuyến khích thúc đẩy thực thi. Các quy định pháp lý này nên nằm trong khung pháp luật