

Dữ liệu số và vận dụng dữ liệu số trong hoạt động giáo dục và đào tạo hiện nay

Nguyễn Thị Thuý*

* *Thiếu tá, ThS. Bộ môn Tin học - Khoa Khoa học Cơ bản - Trường Sĩ quan Lục quân 2*

Received: 27/01/2023; Accepted: 30/01/2023; Published: 06/02/2023

Abstract: Digital data is a core element in the Industrial Revolution 4.0. Currently, digital data has been applied in all aspects of social life. In education and training, digital data has fundamentally changed from management and advocacy to teaching and research activities. The article deals with digital data and its application in education and training, on which the topic proposes some solutions to effectively apply digital data in education and training currently not.

Keywords: Digital data, applied, education and training.

1. Đặt vấn đề

Dữ liệu số (DLS) là dữ liệu dưới dạng ký hiệu, chữ viết, chữ số, hình ảnh, âm thanh hoặc dạng tương tự được biểu diễn bằng tín hiệu số [1]. Thuật ngữ này chỉ một tập hợp dữ liệu rất lớn và phức tạp mà những công cụ, ứng dụng xử lý dữ liệu truyền thống không thể đảm nhiệm được. Nó mang thông tin số và được chia sẻ dưới dạng thông điệp dữ liệu. Bên cạnh đó nó còn là yếu tố cốt lõi để sử dụng, phát triển internet vạn vật (IoT) và trí tuệ nhân tạo (AI). DLS chứa nhiều thông tin phục vụ cho việc kinh doanh, nghiên cứu khoa học, giáo dục, dự đoán dịch bệnh sắp phát sinh và thậm chí là xác định tình trạng giao thông theo thời gian thực.

2. Nội dung nghiên cứu

2.1. Vai trò của việc vận dụng DLS trong hoạt động giáo dục đào tạo

DLS đóng vai trò quan trọng trong mọi mặt đời sống xã hội nói chung. Đem đến những lợi thế vượt bậc trong giáo dục như chương trình giảng dạy tùy chỉnh, cải thiện hệ thống đánh giá, định hướng nghề nghiệp cho người học và đề xuất phương pháp học tập phù hợp. Sự phát triển của nó giúp các cơ sở đào tạo kiểm soát tốt hơn thông tin người học; đổi mới phương thức học và giảng dạy theo hướng tích cực, hiệu quả hơn.

Hiện nay, việc vận dụng DLS trong hoạt động giáo dục và đào tạo thường tồn tại dưới dạng dữ liệu hành chính và dữ liệu quá trình học tập. Sự xuất hiện của DLS trong giáo dục bắt nguồn từ ít nhất hai xu hướng chính trong kỹ nguyên kỹ thuật số. Thứ nhất, việc ghi chép và lưu trữ dữ liệu trong môi trường truyền thống ngày càng được số hóa, dẫn đến lượng lớn thông tin được chuẩn hóa. Cụ thể, Hệ thống thông tin người học (SIS) đã được áp dụng rộng rãi để lưu trữ và sắp xếp thông tin hồ sơ người học như: nhân khẩu học,

nền tảng học vấn,... và hồ sơ học tập như: ghi danh khóa học, kết quả học tập,... SIS giúp quản lý và phân tích các dữ liệu của người học tại một cơ sở giáo dục trong nhiều thập kỷ trên quy mô lớn. Thứ hai, các hoạt động học tập vốn khó ghi chú trong các lớp học truyền thống (face-to-face) giờ đây có thể được ghi lại một phần thông qua Hệ thống quản lý học tập (LMS). Trong hầu hết các trường hợp, LMS được người dạy sử dụng để phân phối tài liệu giảng dạy, quản lý bài tập của người học và giao tiếp với họ. Từ việc truy cập vào các môđun khóa học đến sửa chữa bài tiểu luận, những thông tin này có thể dễ dàng được lưu trữ tới hàng nghìn mục dữ liệu cho mỗi người. Ngoài SIS và LMS, những đổi mới trong môi trường học tập kỹ thuật số làm phong phú thêm các hình thức sự phạm, đồng thời thu thập “DLS” của người học. Sự đa dạng này đã làm cho kho dữ liệu được đa phương thức với khối lượng lớn.

2.2. Cấp độ DLS và kỹ thuật khai thác DLS trong hoạt động giáo dục và đào tạo hiện nay

- Cấp độ DLS

DLS trong lĩnh vực giáo dục hiện nay được chia làm ba cấp độ. Đó là: vi mô, trung bình và vĩ mô.

Dữ liệu số cấp vi mô là dữ liệu tương tác “siêu nhỏ” với số giây giữa các hành động có thể thu thập dữ liệu riêng lẻ từ hàng triệu người học tiềm năng. Hầu hết dữ liệu cấp vi mô được thu thập tự động trong quá trình tương tác giữa người học và môi trường học tập tương ứng của họ, bao gồm hệ thống dạy học thông minh, các khóa học trực tuyến lớn (MOOC), các trò chơi mô phỏng,...

Dữ liệu lớn cấp trung bình bao gồm các sản phẩm viết của người học được số hóa và thu thập một cách có hệ thống trong nhiều môi trường học tập khác nhau, từ các bài tập trong khóa học đến tham gia diễn đàn

thảo luận trực tuyến, hệ thống dạy học thông minh và các tương tác trên mạng xã hội. Đáng chú ý, dữ liệu cấp trung bình mang lại cơ hội nắm bắt thông tin sơ cấp về sự tiến bộ của người học trong khả năng nhận thức và xã hội, cũng như trạng thái cảm xúc.

Dữ liệu lớn cấp vĩ mô bao gồm dữ liệu được thu thập ở cấp thể chế như: dữ liệu nhân khẩu học và nhập học của học viên, dữ liệu dịch vụ trong khuôn viên trường, lịch học và dữ liệu đăng ký khóa học, cũng như dữ liệu về yêu cầu chính của nhà trường và dữ liệu về các yêu cầu đối với người học. Mặc dù DLS cấp vĩ mô thường được thu thập trong khoảng thời gian nhiều năm nhưng không được cập nhật thường xuyên, thường chỉ một hoặc hai lần mỗi kỳ như: thông tin về lịch trình khóa học, hồ sơ điểm.

- Kỹ thuật khai thác

Các kỹ thuật khai thác dữ liệu cho DLS trong giáo dục được Baker và Siemens phân loại thành nhiều phương pháp như: suy luận mô hình hóa kiến thức; thuật toán khám phá cấu trúc trong đó tác giả đã chú trọng cấu trúc nội dung, kỹ năng giáo dục và cấu trúc mạng xã hội của người học; khai thác mối quan hệ là quan tâm đến khai thác mẫu tuần tự và khai thác tương quan; hình dung; khám phá mô hình chú ý đến việc sử dụng mô hình trong các phân tích tiếp theo. Với khối lượng, tốc độ và sự đa dạng của DLS cho thấy việc thu thập thông tin cùng với hành vi của người học đã mang lại nhiều giá trị trong lĩnh vực nghiên cứu giáo dục. Các DLS về hành động của người học đã mang lại thông tin cụ thể và chi tiết hơn về quá trình học tập. Bằng cách kết hợp dữ liệu hành vi với khảo sát hoặc thang đo tâm lý, các nhà nghiên cứu có thể lập bản đồ chuỗi hành động với các đặc điểm nhận thức và kiểm tra xem các dấu vết hành vi quan sát được có phù hợp với các giả định lý thuyết hay không và tinh chỉnh lý thuyết ở cấp độ chi tiết. Những thông tin phong phú này giúp họ hiểu rõ cơ chế tác động của chính sách cụ thể và giải quyết các vấn đề liên quan đến chính sách.

2.3. Một số biện pháp vận dụng DLS trong hoạt động giáo dục và đào tạo

2.3.1. Nâng cao nhận thức về ứng dụng công nghệ thông tin, số hóa và ứng dụng dữ liệu lớn trong hoạt động quản lý đào tạo, nghiên cứu, giảng dạy

Công nghệ mang đến những lợi thế vượt bậc trong giáo dục và đào tạo như chương trình giảng dạy tùy chỉnh, cải thiện hệ thống đánh giá, định hướng nghề nghiệp cho người học và đề xuất phương pháp học phù hợp, hỗ trợ quản lý, giám sát hiệu quả, nâng cao chất lượng giảng dạy... Hiện nay, trong giáo dục và đào tạo DLS chủ yếu được vận dụng trong việc lưu trữ, quản lý, phân tích bộ dữ liệu từ hồ sơ người học. Nó có khả

năng thay đổi cách thức tổ chức hoạt động, tương tác với người học cũng như quản lý, vận hành hoạt động của cơ quan. Để DLS thay đổi cách thức tổ chức, quản lý đào tạo, chương trình và phương pháp giảng dạy, giải quyết các vấn đề tồn đọng trong hệ thống, nâng cao hiệu quả chương trình học tập thì nhà trường, các cấp, các ngành cần nâng cao nhận thức về tính cấp thiết của việc vận dụng công nghệ thông tin, số hóa nói chung và ứng dụng DLS nói riêng trong hoạt động quản lý đào tạo, nghiên cứu, giảng dạy mình.

2.3.2. Thực hiện một số bước trong quy trình chuẩn hóa dữ liệu nhằm hình thành DLS từ nguồn nội sinh của nhà trường

DLS vừa là cách thức tổ chức vừa là một hình thức lưu trữ dữ liệu trong ngành giáo dục và đào tạo. Trong các nhà trường, nó được hình thành từ nhóm các dữ liệu trong các mảng: trường học, lớp học, học viên, đội ngũ cán bộ, cơ sở vật chất, học liệu, bài thi, điểm thi,... Thực chất, đây là quá trình chuyển đổi số các dữ liệu, tức hoạt động số hóa và công cụ xử lý tương ứng. Do đó, nhà trường cần số hóa, chuẩn hóa lại dữ liệu để có thể thu thập, tập hợp dữ liệu vào cơ sở dữ liệu chủ - dữ liệu lớn dùng chung để khai thác. Dữ liệu của nhà trường có thể thu thập từ tất cả các hoạt động quản lý, phục vụ của các vụ chức năng và hoạt động giảng dạy, nghiên cứu của các viện chuyên ngành cùng cơ sở dữ liệu của thư viện. Tận dụng khai thác các nguồn dữ liệu cấp vĩ mô (dữ liệu về thông tin học viên, chương trình giảng dạy, thời khóa biểu, thông tin về cán bộ, giảng viên,...); dữ liệu cấp trung bình (dữ liệu về các bài luận, tiểu luận, luận văn, luận án, các bài kiểm tra của học viên, đề tài các cấp,...); dữ liệu lớn cấp vi mô (hoạt động của người học trên các ứng dụng học trực tuyến, các tài nguyên trong thư viện số, truy cập mạng internet,...). Việc hình thành dữ liệu nội sinh thông qua chuyển đổi số toàn bộ quy trình hoạt động, dữ liệu các mặt công tác của nhà trường vừa mang tính xây dựng, vừa mang tính ứng dụng. Bởi thực chất, trong quá trình số hóa, các dữ liệu đó được bảo mật ở dạng thức số và quay lại phục vụ cho các hoạt động/lĩnh vực công tác khác của chính nhà trường.

2.3.3. Nghiên cứu lựa chọn mô hình dữ liệu lớn và truy xuất dữ liệu lớn phù hợp với đặc điểm về dữ liệu cũng như chức năng, nhiệm vụ của nhà trường

Cần phân tích cụ thể đặc điểm về cơ cấu tổ chức, chức năng, nhiệm vụ của nhà trường qua các DLS cả cấp vi mô, vĩ mô và trung bình trong quá trình hoạt động của nhà trường, tránh lãng phí nguồn lực và giúp người sử dụng dễ dàng phân tích, truy xuất các thông tin nhằm nâng cao chất lượng giảng dạy, học tập, nghiên cứu, đồng thời giúp nâng cao hiệu quả hoạt

động của các đơn vị chức năng. Thực tế trên thế giới hiện nay, việc ứng dụng DLS đã mang lại nhiều lợi ích cho cả cơ sở đào tạo, người học cũng như các đơn vị liên quan và rộng hơn là quốc gia, quốc tế. Với đặc điểm liên kết mở, môi trường số năng động và tính năng thu thập, lưu trữ, trích xuất ứng dụng. Việc lựa chọn mô hình ứng dụng DLS cần tính tới những yêu cầu về kỹ thuật, công nghệ phù hợp với điều kiện, đặc thù của nhà trường và cần chú trọng vấn đề an toàn, bảo mật thông tin, tránh rò rỉ thông tin mật, thông tin cá nhân; cũng như giải quyết tốt sự xung đột giữa yêu cầu về quyền riêng tư và những yêu cầu về sự cởi mở và minh bạch hơn của môi trường mạng Internet. Chia khóa bảo đảm xây dựng thành công và khai thác hiệu quả cơ sở dữ liệu cho nhà trường là phải xây dựng được cơ chế để vừa đáp ứng an toàn và bảo mật thông tin cá nhân, vừa cho phép phân tích, khai thác dữ liệu để nâng cao chất lượng dạy và học.

2.3.4. Tăng cường quản trị dữ liệu và sử dụng các công cụ để phân tích dữ liệu lớn nhằm mang lại thông tin hữu ích góp phần đổi mới phương pháp đào tạo, bồi dưỡng và nghiên cứu

Quản trị dữ liệu là chìa khóa quan trọng để nâng cao chất lượng đào tạo, bồi dưỡng, nghiên cứu, giảng dạy; nâng cao hiệu suất, hiệu quả học tập của người học. Việc phân tích dữ liệu cũng đang biến đổi cách thức tương tác giữa cơ sở đào tạo với người học và người dạy. Với sự phát triển của khoa học dữ liệu, sự đổi mới trong cách thức thu thập và xử lý thông tin, nhiều đơn vị trong nước và quốc tế đã tận dụng được sức mạnh của Big data trong việc ra các quyết định quan trọng. Đây là con đường tất yếu mà nhà trường cần hướng đến để biến các dữ liệu trở thành thông tin hữu dụng, phục vụ cho chính hoạt động của mình. Bên cạnh đó, nhà trường cần làm chủ hệ thống DLS và sử dụng các công cụ phân tích dữ liệu như là yếu tố hỗ trợ cho việc ra quyết định, nhờ đó tạo ra những tác động tích cực như nâng cao hiệu quả giảng dạy, nghiên cứu của cán bộ giáo viên và hiệu quả học tập của người học.

2.3.5. Tăng cường đầu tư hạ tầng để phục vụ việc chuyển đổi số cũng như ứng dụng dữ liệu lớn trong quản lý, giảng dạy

Trong bối cảnh chuyển đổi số trong ngành giáo dục đang diễn ra mạnh mẽ, cần tăng cường đầu tư về hạ tầng công nghệ thông tin hiện đại, liên thông thống nhất trong toàn hệ thống nhà trường các cấp; từng bước xây dựng trung tâm tích hợp dữ liệu là đầu mối thu thập, lưu trữ và xử lý DLS. Đặc biệt, cần chú ý đầu tư hạ tầng máy chủ để lưu trữ dữ liệu nhằm bảo đảm tính bảo mật thông tin. Ngoài ra, cần tập trung

thực hiện Đề án Xây dựng thư viện điện tử, đầu tư xây dựng các phòng học đa phương tiện. Đầu tư các trang thiết bị công nghệ thông tin phục vụ giáo dục - đào tạo tại nhà trường cần có trọng tâm, trọng điểm, tránh lãng phí. Ưu tiên các trang thiết bị có nhu cầu sử dụng cao, phục vụ trực tiếp cho nhiệm vụ giáo dục - đào tạo như máy tính, máy chiếu, bảng thông minh, thiết bị mô phỏng...

2.3.6. Nâng cao chất lượng nguồn nhân lực để có thể xây dựng DLS trong nhà trường và tăng cường ứng dụng DLS nhằm nâng cao chất lượng giảng dạy và nghiên cứu

Trong thời gian tới, cần chú trọng xây dựng đội ngũ chuyên gia công nghệ thông tin, đặc biệt là các chuyên gia về khoa học dữ liệu - đóng vai trò trụ cột trong việc thiết kế và xây dựng dữ liệu lớn của nhà trường. Bên cạnh đó, cần tăng cường bồi dưỡng kỹ năng sử dụng công nghệ thông tin cho cán bộ, giáo viên nhằm thích ứng với việc chuyển đổi số. Đa dạng hóa các hình thức đào tạo, bồi dưỡng bằng các giải pháp cụ thể như: đào tạo, bồi dưỡng ngắn hạn, dài hạn; đào tạo, bồi dưỡng tập trung, không tập trung; mở rộng liên kết, hợp tác đào tạo, bồi dưỡng với các cơ quan khác ở địa phương và Trung ương...

3. Kết luận

DLS đóng vai trò quan trọng trong việc phát triển trên mọi lĩnh vực trong đó có giáo dục. Bài viết trình bày DLS và một số biện pháp nhằm vận dụng DLS trong lĩnh vực giáo dục và đào tạo. DLS trong giáo dục không chỉ dừng lại ở việc dạy và học, mà phạm vi vô cùng rộng lớn, ở đó tất cả các hoạt động, quan hệ, thao tác đang thực hiện sẽ được đặt trong nền tảng số để vận hành. DLS cũng không phải để thay cho yếu tố thực tại, trực tiếp, sinh động mà các yếu tố thực tại được đặt vào chuyển đổi số, qua đó vận hành tốt hơn, kết nối hơn, rộng mở hơn, tạo ra nhiều cơ hội học tập hơn, làm cho công cuộc vận hành giáo dục trong thực tại hiệu quả và chất lượng hơn.

Tài liệu tham khảo

1. Baker, R. S., & Siemens, G: *Educational data mining and learning analytics*, In R. K. Sawyer (Ed.): Cambridge handbook of the learning sciences (2nd ed.), Cambridge University Press, 2014, pp.253-274.

2. Chính phủ (2020), *Nghị định số 47/2020/NĐ-CP về quản lý, kết nối và chia sẻ DLS của cơ quan nhà nước*. Hà Nội

3. Thủ tướng Chính phủ (2017), *Quyết định số 117/QĐ-TTg ngày 25/01/2017 phê duyệt Đề án “Tăng cường ứng dụng CNTT trong quản lý và hỗ trợ các hoạt động dạy-học, NCKH góp phần nâng cao chất lượng GDĐT giai đoạn 2016-2020, định hướng đến 2025”*. Hà Nội