

BẢO VỆ DỮ LIỆU NGƯỜI HỌC TRONG ĐÀO TẠO TRỰC TUYẾN: LIÊN HỆ TRƯỜNG ĐẠI HỌC MỞ HÀ NỘI

Nguyễn Thị Phan Mai^{1*}, Đồng Thị Toàn¹, Nguyễn Tiến Dũng¹

*Tác giả liên hệ, Email: maintp@hou.edu.vn. ORCID: 0009-0001-5060-1678

Ngày tòa soạn nhận được bài báo: 18/05/2026

Ngày phản biện đánh giá: 02/07/2026

Ngày bài báo được duyệt đăng: 20/07/2026

DOI: 10.59266/houjs.2026.1361

Tóm tắt: Đào tạo trực tuyến làm gia tăng mạnh hoạt động thu thập, lưu trữ, phân tích và chia sẻ dữ liệu người học, từ hồ sơ hành chính, kết quả học tập đến nhật ký LMS, dữ liệu tương tác, bản ghi âm - ghi hình, dữ liệu sinh trắc học và dữ liệu suy luận từ trí tuệ nhân tạo (AI).. Bài viết phân tích cơ sở lý thuyết và pháp lý của bảo vệ dữ liệu người học; làm rõ quan hệ giữa dữ liệu cá nhân, dữ liệu nhạy cảm và dữ liệu giáo dục; đồng thời liên hệ với thực tiễn Trường Đại học Mở Hà Nội.

Từ khóa: bảo vệ dữ liệu cá nhân, dữ liệu người học, đào tạo trực tuyến, giáo dục đại học, trí tuệ nhân tạo

I. Đặt vấn đề

Chuyển đổi số trong giáo dục đại học không chỉ làm thay đổi phương thức dạy và học, mà còn làm thay đổi cấu trúc quản trị của nhà trường. Khi hoạt động tuyển sinh, tổ chức lớp học, học liệu số, tương tác sư phạm, kiểm tra đánh giá, khảo sát phản hồi và hỗ trợ học vụ được thực hiện trên nền tảng số, dữ liệu của người học trở thành một điều kiện vận hành thiết yếu của đào tạo trực tuyến.

Trong đào tạo trực tuyến, người học để lại dấu vết dữ liệu trên hầu hết các hoạt động học tập. Ngoài thông tin định danh, hồ sơ sinh viên và điểm số, hệ thống còn có thể ghi nhận thời điểm đăng nhập, thời lượng học tập, số lần xem bài giảng,

thao tác trên LMS, nội dung trao đổi, tệp bài tập, kết quả kiểm tra, hình ảnh, giọng nói, dữ liệu thiết bị, địa chỉ IP, dữ liệu xác thực danh tính và dữ liệu giám sát thi trực tuyến. Đây là bước phát triển có giá trị quản trị, nhưng đồng thời cũng làm gia tăng nguy cơ xâm phạm quyền riêng tư, quyền được đánh giá công bằng và quyền tự chủ của người học.

Bối cảnh pháp lý của Việt Nam đang chuyển từ cách tiếp cận rải rác về an toàn thông tin và bảo mật thông tin sang mô hình quản trị dữ liệu cá nhân có hệ thống. Bên cạnh các quy định về quyền riêng tư trong Hiến pháp năm 2013, an toàn thông tin mạng, an ninh mạng, giáo dục đại học và đào tạo trực tuyến, pháp luật hiện đã

¹ Trường Đại học Mở Hà Nội, Hà Nội, Việt Nam

đặt ra yêu cầu chặt chẽ hơn về quyền của chủ thể dữ liệu, căn cứ xử lý, đánh giá tác động, chuyển dữ liệu, xử lý dữ liệu nhạy cảm và trách nhiệm của bên kiểm soát, bên xử lý dữ liệu (Quốc hội, 2013, 2015, 2018, 2024, 2025a; Chính phủ, 2025). Đối với các ứng dụng AI trong giáo dục, yêu cầu minh bạch, giám sát của con người, quản trị rủi ro và bảo vệ quyền, lợi ích của cá nhân càng trở nên cấp thiết (Quốc hội, 2025b; UNESCO, 2023).

Trường Đại học Mở Hà Nội là cơ sở giáo dục đại học có đặc thù gắn với đào tạo mở, đào tạo từ xa và đào tạo trực tuyến. Đặc thù này làm cho vấn đề bảo vệ dữ liệu cá nhân người học không chỉ là yêu cầu tuân thủ pháp luật, mà còn là điều kiện để duy trì niềm tin số, nâng cao trải nghiệm người học và bảo đảm chất lượng đào tạo. Trên cơ sở đó, bài viết phân tích ở bình diện chung của cơ sở giáo dục đại học, đồng thời liên hệ với thực trạng và nhu cầu quản trị dữ liệu tại Trường Đại học Mở Hà Nội.

Mục tiêu của bài viết là: làm rõ cơ sở lý thuyết về bảo vệ dữ liệu người học trong đào tạo trực tuyến; phân loại dữ liệu cá nhân, dữ liệu nhạy cảm và dữ liệu giáo dục; đánh giá yêu cầu và thách thức đối với cơ sở giáo dục đại học; và đề xuất định hướng hoàn thiện cơ chế bảo vệ dữ liệu cá nhân người học phù hợp với bối cảnh Trường Đại học Mở Hà Nội.

II. Cơ sở lý thuyết

Quyền riêng tư là nền tảng lý luận quan trọng của bảo vệ dữ liệu cá nhân. Trong pháp luật quốc tế, quyền không bị can thiệp tùy tiện vào đời sống riêng tư được ghi nhận trong Tuyên ngôn thế giới về quyền con người và Công ước quốc tế về các quyền dân sự và chính trị (United

Nations, 1948, 1966). Trong xã hội số, quyền riêng tư không chỉ dừng lại ở việc bảo mật thông tin cá nhân, mà mở rộng thành quyền kiểm soát hợp lý đối với việc thu thập, sử dụng, chia sẻ, lưu giữ và suy luận thông tin về cá nhân. Đối với người học, quyền này gắn trực tiếp với nhân phẩm, cơ hội học tập, quyền được đánh giá công bằng và sự an toàn trong môi trường giáo dục số.

Tiếp cận quản trị dữ liệu nhấn mạnh rằng bảo vệ dữ liệu cá nhân không thể chỉ dựa vào bảo mật kỹ thuật. Cơ sở giáo dục đại học cần xác lập chủ thể chịu trách nhiệm, quy trình xử lý dữ liệu, phân quyền truy cập, thời hạn lưu trữ, cơ chế chia sẻ, quy trình xử lý sự cố và kênh để người học thực hiện quyền của mình. Quản trị dữ liệu vì vậy là sự kết hợp giữa pháp luật, tổ chức, công nghệ và đạo đức học thuật.

Trong giáo dục trực tuyến, phân tích học tập và AI làm cho bảo vệ dữ liệu trở nên phức tạp hơn. Các nghiên cứu về learning analytics chỉ ra rằng dữ liệu có thể hỗ trợ phát hiện khó khăn học tập và cá nhân hóa hỗ trợ, nhưng cũng có thể gây ra rủi ro giám sát, thiên lệch, gán nhãn và xâm phạm quyền tự chủ của người học nếu thiếu minh bạch và trách nhiệm giải trình (Pardo và Siemens, 2014; Slade và Prinsloo, 2013). Drachsler và Greller (2016) đề xuất cách tiếp cận thận trọng, trong đó sự tin cậy của phân tích dữ liệu học tập (learning analytics) phụ thuộc vào minh bạch, đồng thuận, bảo mật, quyền kiểm soát và trách nhiệm của tổ chức. Những yêu cầu này có ý nghĩa trực tiếp đối với trường đại học khi triển khai LMS, công cụ khảo thí trực tuyến, phần mềm chống đạo văn, chatbot, hệ thống dự báo nguy cơ học vụ hoặc công cụ AI hỗ trợ đánh giá.

Một điểm cần làm rõ là quan hệ giữa dữ liệu cá nhân, dữ liệu cá nhân nhạy cảm và dữ liệu giáo dục. Dữ liệu giáo dục là dữ liệu phát sinh trong hoạt động tuyển sinh, đào tạo, khảo thí, hỗ trợ người học và bảo đảm chất lượng. Khi dữ liệu giáo dục gắn với một cá nhân xác định hoặc có thể xác định được, nó là dữ liệu cá nhân. Một số dữ liệu giáo dục có thể trở thành dữ liệu nhạy cảm hoặc dữ liệu có rủi ro cao khi phản ánh sức khỏe, khuyết tật, sinh trắc học, hoàn cảnh cá nhân, kỷ luật, nghi vấn gian lận, hồ sơ hành vi hoặc khi được dùng để phân loại, dự báo, giám sát và ra quyết định có ảnh hưởng đáng kể đến người học. Do đó, không nên đồng nhất mọi dữ liệu học tập với dữ liệu nhạy cảm, nhưng cũng không nên coi dữ liệu học tập chỉ là dữ liệu hành chính thông thường.

III. Phương pháp nghiên cứu

Bài viết sử dụng phương pháp nghiên cứu định tính, kết hợp phân tích luật học, tổng hợp tài liệu và nghiên cứu trường hợp. Phương pháp phân tích luật học được sử dụng để nhận diện các yêu cầu pháp lý đặt ra bởi khung pháp luật về bảo vệ dữ liệu cá nhân, dữ liệu, an toàn thông tin, an ninh mạng, giáo dục đại học, đào tạo trực tuyến và quản trị trí tuệ nhân tạo.

Phương pháp tổng hợp tài liệu được sử dụng để tiếp cận các nghiên cứu quốc tế về quyền riêng tư, quản trị dữ liệu giáo dục, learning analytics và AI trong giáo dục. Các tài liệu của UNESCO, OECD, Liên minh châu Âu và một số nghiên cứu học thuật nền tảng được sử dụng nhằm đặt vấn đề của Việt Nam trong bối cảnh rộng hơn của quản trị giáo dục số (European Parliament và Council of the European Union, 2016; OECD, 2021; UNESCO, 2023).

Phương pháp nghiên cứu trường hợp được sử dụng thông qua liên hệ với Trường Đại học Mở Hà Nội. Do bài viết không thực hiện kiểm toán hệ thống kỹ thuật hay khảo sát định lượng nội bộ, phần liên hệ thực tiễn được trình bày ở mức nhận diện yêu cầu quản trị chung xuất phát từ đặc thù đào tạo mở, đào tạo từ xa và đào tạo trực tuyến của nhà trường.

Bài viết cũng xác định giới hạn nghiên cứu: chưa thực hiện khảo sát định lượng, chưa kiểm toán kỹ thuật các hệ thống hiện hữu và chưa tiếp cận hồ sơ vận hành nội bộ. Vì vậy, phần liên hệ Trường Đại học Mở Hà Nội được trình bày như một phân tích định hướng quản trị, kèm các chỉ báo và quy trình khả thi để nhà trường có thể tiếp tục kiểm chứng trong giai đoạn triển khai.

IV. Kết quả và thảo luận

4.1. Khung pháp lý mới và yêu cầu chuyển đổi mô hình quản trị dữ liệu người học

Khung pháp lý hiện nay đặt cơ sở giáo dục đại học trước yêu cầu quản trị dữ liệu toàn diện hơn giai đoạn chỉ nhấn mạnh bảo mật thông tin. Quyền về đời sống riêng tư, bí mật cá nhân và an toàn thông tin cá nhân được ghi nhận ở cấp độ hiến định; các luật về an toàn thông tin mạng, an ninh mạng, dữ liệu, bảo vệ dữ liệu cá nhân, giáo dục đại học và quản trị trí tuệ nhân tạo tạo thành nền tảng điều chỉnh việc thu thập, lưu trữ, sử dụng, chia sẻ và xử lý tự động dữ liệu trong giáo dục số (Quốc hội, 2013, 2015, 2018, 2024, 2025a, 2025b; Chính phủ, 2025, 2026).

Điểm thay đổi cốt lõi là dữ liệu người học không còn chỉ là thông tin quản lý đào tạo, mà là đối tượng pháp lý gắn

với quyền của chủ thể dữ liệu và trách nhiệm giải trình của nhà trường. Vì vậy, một chính sách bảo mật chung trên LMS là chưa đủ; cơ sở giáo dục cần có quy chế, quy trình, phân quyền, hồ sơ đánh giá tác động, cơ chế xử lý yêu cầu của người học và phương án ứng phó sự cố. Khi AI được dùng để chấm điểm, phát hiện đạo văn, giám sát thi, dự báo nguy cơ học vụ hoặc khuyến nghị hỗ trợ, yêu cầu bảo vệ dữ liệu phải gắn với minh bạch thuật toán, giám sát của con người và quyền yêu cầu xem xét lại của người học.

4.2. Nhận diện và phân loại dữ liệu người học theo mức độ rủi ro

Bước đầu tiên của quản trị dữ liệu là nhận diện đúng dữ liệu nào đang được xử lý. Trong đào tạo trực tuyến, dữ liệu người học phân tán trên nhiều hệ thống: tuyển sinh, quản lý đào tạo, LMS, học liệu số, hợp trực tuyến, khảo thí, chống đạo văn, thanh toán, thư điện tử, lưu trữ đám mây, khảo sát và công cụ AI. Nếu chưa có bản đồ dữ liệu, nhà trường khó xác định dữ liệu nào cần thông báo riêng, dữ liệu nào phải hạn chế truy cập, dữ liệu nào cần đánh giá tác động và dữ liệu nào phải xóa hoặc ẩn danh sau khi hết mục đích sử dụng.

Có thể phân loại dữ liệu người học thành bốn nhóm. Nhóm thứ nhất là dữ liệu định danh và hành chính như họ tên, mã sinh viên, thông tin liên hệ, tài khoản học tập, lớp, ngành và tình trạng học vụ. Nhóm thứ hai là dữ liệu học tập và dữ liệu giáo dục như điểm số, bài làm, kết quả kiểm tra, nhận xét của giảng viên, phản hồi học phần và hồ sơ hỗ trợ học tập. Nhóm thứ ba là dữ liệu hành vi số và dữ liệu suy luận như nhật ký truy cập, thời lượng học, thao tác trên nền tảng, mô hình tương tác, dự báo nguy cơ bỏ học hoặc phân nhóm mức độ tham gia. Nhóm thứ tư là dữ liệu nhạy cảm hoặc rủi ro cao như sinh trắc học,

hình ảnh, giọng nói, bản ghi thi trực tuyến, dữ liệu sức khỏe, khuyết tật, kỷ luật, nghi vấn gian lận và kết quả AI có ảnh hưởng đáng kể đến cơ hội học tập.

4.3. Thực trạng quản trị dữ liệu người học tại Trường Đại học Mở Hà Nội và giới hạn chứng cứ thực nghiệm

Trường Đại học Mở Hà Nội có đặc thù gắn với đào tạo mở, đào tạo từ xa và đào tạo trực tuyến. Dữ liệu người học được tạo lập liên tục trong toàn bộ vòng đời đào tạo, từ tuyển sinh, nhập học, xác thực người học, tổ chức học trực tuyến, cung cấp học liệu, kiểm tra đánh giá, khảo sát phản hồi, hỗ trợ học vụ đến xét tốt nghiệp và quản lý văn bằng. Dữ liệu giúp nâng cao trải nghiệm người học và bảo đảm chất lượng, nhưng đồng thời tạo ra yêu cầu quản trị liên thông giữa đào tạo, khảo thí, công nghệ thông tin, pháp chế, bảo đảm chất lượng, công tác sinh viên, các khoa và đơn vị đào tạo trực tuyến.

Do chưa có khảo sát định lượng hoặc kiểm toán kỹ thuật nội bộ, bài viết không đưa ra kết luận về mức độ tuân thủ cụ thể của từng hệ thống. Tuy nhiên, để tăng tính thực nghiệm cho các nghiên cứu tiếp theo, nhà trường có thể thu thập một số chỉ báo tối thiểu: số hệ thống đang xử lý dữ liệu người học; số nhóm dữ liệu cá nhân và dữ liệu nhạy cảm; số nhà cung cấp công nghệ có tiếp cận dữ liệu; số hoạt động dùng AI hoặc phân tích học tập; số yêu cầu thực hiện quyền dữ liệu; số sự cố hoặc cảnh báo an toàn dữ liệu; số cán bộ, giảng viên được tập huấn về bảo vệ dữ liệu.

Từ góc độ quản trị, có thể nhận diện bốn vấn đề chính: dữ liệu phân tán theo đơn vị và hệ thống; cơ chế thông báo và đồng ý của người học cần được chuẩn hóa hơn; hợp đồng với bên thứ ba cần bổ sung điều khoản xử lý dữ liệu, chuyển dữ liệu, xóa dữ liệu, thông báo sự cố và cấm dùng dữ liệu

người học để huấn luyện AI thương mại nếu chưa có căn cứ hợp pháp; và việc sử dụng AI cần được kiểm soát không chỉ ở dữ liệu đầu vào mà cả ở kết quả suy luận, nhất là các cảnh báo về gian lận, nguy cơ học vụ hoặc phân loại năng lực.

Dữ liệu phân tán có thể làm cho cơ chế đồng ý khó minh bạch; đồng ý thiếu rõ ràng làm cho việc chia sẻ cho bên thứ ba khó kiểm soát; hợp đồng công nghệ thiếu điều khoản dữ liệu làm tăng nguy cơ sử dụng dữ liệu ngoài mục đích; còn AI có thể biến các dữ liệu tưởng như thông thường thành hồ sơ suy luận có tác động đáng kể đến người học. Vì vậy, giải pháp không nên được thiết kế rời rạc theo từng phòng, ban, mà cần đặt trong một mô hình quản trị thống nhất, trong đó mỗi hoạt động xử lý dữ liệu đều có chủ thể chịu trách nhiệm, căn cứ xử lý, biện pháp bảo vệ và cơ chế kiểm tra.

4.4. Các thách thức trọng tâm trong bảo vệ dữ liệu người học

Thách thức thứ nhất là cân bằng giữa bảo đảm chất lượng đào tạo và quyền riêng tư. Đào tạo trực tuyến cần dữ liệu để xác nhận sự tham gia, kiểm soát tiến độ, hỗ trợ người học và phòng chống gian lận; tuy nhiên, nếu ghi hình quá rộng, yêu cầu bật camera thường xuyên, theo dõi thao tác màn hình liên tục hoặc lưu giữ dữ liệu quá lâu, nhà trường có thể tạo ra cảm giác giám sát thường trực.

Thách thức thứ hai là rủi ro xử lý tự động và thiên lệch thuật toán. AI có thể hỗ trợ phát hiện khó khăn học tập, cá nhân hóa hỗ trợ và nâng cao hiệu quả đánh giá, nhưng cũng có thể bất lợi cho người học có điều kiện thiết bị hạn chế, đường truyền không ổn định, học không đồng bộ hoặc có phong cách học tập khác biệt. Các kết quả tự động không nên là căn cứ duy nhất cho quyết định ảnh hưởng đến điểm số,

cảnh báo học vụ, kỷ luật, học bổng, điều kiện dự thi hoặc tốt nghiệp.

Thách thức thứ ba là kiểm soát dữ liệu sinh trắc học và thi trực tuyến. Xác thực danh tính là cần thiết nhưng nhà trường cần phân tầng rủi ro: không phải mọi hình thức kiểm tra đều cần giám sát sinh trắc học; với kiểm tra thường xuyên có thể ưu tiên ngân hàng câu hỏi, đề mở, bài tập ứng dụng hoặc vấn đáp ngắn; chỉ kỳ thi rủi ro cao mới cần nhắc biện pháp giám sát mạnh hơn.

Thách thức thứ tư là bảo đảm quyền dữ liệu trong thực tế và xây dựng văn hóa dữ liệu an toàn. Người học chỉ thực sự là chủ thể dữ liệu khi biết gửi yêu cầu đến đâu, thời hạn phản hồi là bao lâu, ai xử lý và trường hợp nào có thể bị từ chối. Đồng thời, nhiều rủi ro không đến từ tấn công mạng phức tạp mà từ thói quen gửi danh sách sinh viên qua email cá nhân, công khai link chứa dữ liệu, chia sẻ bảng điểm không đúng phạm vi hoặc đưa bài làm của sinh viên lên công cụ AI công cộng.

Bên cạnh đó, thách thức về tổ chức nhân sự cũng rất đáng kể. Mô hình phù hợp là thiết lập một đầu mối điều phối bảo vệ dữ liệu, nhưng phân bổ trách nhiệm cụ thể cho từng đơn vị: đơn vị đào tạo chịu trách nhiệm về mục đích và hồ sơ học vụ; đơn vị khảo thí chịu trách nhiệm về dữ liệu thi; đơn vị công nghệ chịu trách nhiệm về hạ tầng và an toàn; đơn vị pháp chế chịu trách nhiệm rà soát căn cứ xử lý và hợp đồng; còn các khoa và giảng viên chịu trách nhiệm trong phạm vi sử dụng dữ liệu lớp học.

4.5. Quy trình rút gọn xây dựng bản đồ dữ liệu người học

Để khắc phục tình trạng giải pháp chỉ dừng ở định hướng, nhà trường có thể xây dựng bản đồ dữ liệu theo sáu bước. Bước 1: thành lập tổ công tác, gồm phòng

quản lý đào tạo, trung tâm công nghệ kỹ thuật số, phòng khảo thí và đảm bảo chất lượng, phòng kiểm tra pháp chế, phòng công tác sinh viên và đơn vị đào tạo trực tuyến. Bước 2: kiểm kê các hệ thống và điểm phát sinh dữ liệu, từ tuyển sinh, quản lý đào tạo, LMS, khảo thí đến lưu trữ đám mây và công cụ AI. Bước 3: xác định nhóm dữ liệu, chủ thể dữ liệu và mục đích xử lý của từng hệ thống. Về mặt sản phẩm, bản đồ dữ liệu không cần khởi đầu bằng một hệ thống phần mềm phức tạp. Bước 4: phân loại mức độ rủi ro dựa trên khả năng nhận diện cá nhân, mức độ nhạy cảm, quy mô dữ liệu, khả năng chia sẻ cho bên thứ ba và tác động đến quyền lợi người học. Bước 5: xác định luồng chia sẻ, thời hạn lưu trữ, phương thức xóa, ẩn danh hoặc tiêu hủy. Bước 6: phê duyệt, công bố nội bộ và cập nhật định kỳ khi có hệ thống mới, nhà cung cấp mới, ứng dụng AI mới hoặc thay đổi mục đích xử lý. Sản phẩm cuối cùng nên là một bảng quản trị gồm: tên dữ liệu, hệ thống, mục đích, căn cứ xử lý, đơn vị phụ trách, bên nhận, thời hạn lưu trữ, mức rủi ro và biện pháp bảo vệ.

4.6. Khung mẫu đánh giá tác động bảo vệ dữ liệu đối với hoạt động đào tạo trực tuyến

Một khung mẫu đánh giá có thể gồm mười nội dung: (1) mô tả hoạt động xử lý; (2) mục đích và căn cứ xử lý; (3) loại dữ liệu và mức độ nhạy cảm; (4) nhóm người học bị ảnh hưởng; (5) hệ thống và bên thứ ba liên quan; (6) rủi ro đối với quyền riêng tư, cơ hội học tập và danh dự; (7) biện pháp giảm thiểu rủi ro; (8) cơ chế thông báo, đồng ý và thực hiện quyền dữ liệu; (9) trách nhiệm của đơn vị phụ trách; (10) kết luận phê duyệt, điều kiện triển khai hoặc yêu cầu điều chỉnh.

Khung đánh giá tác động nên được kích hoạt đối với các hoạt động có một

trong các dấu hiệu: xử lý dữ liệu sinh trắc học; ghi âm, ghi hình hoặc giám sát thi trực tuyến; sử dụng AI để chấm điểm, phát hiện gian lận, cảnh báo học vụ hoặc phân loại người học; chia sẻ dữ liệu cho nhà cung cấp ngoài trường; chuyển dữ liệu ra nước ngoài; hoặc xử lý dữ liệu trên quy mô lớn. Kết quả đánh giá không nhất thiết chỉ có hai lựa chọn “cho phép” hoặc “không cho phép”, mà có thể đưa ra điều kiện triển khai như rút ngắn thời hạn lưu trữ, ẩn danh hóa dữ liệu, giới hạn quyền truy cập, bổ sung thông báo cho người học, yêu cầu giám sát của con người hoặc thử nghiệm ở phạm vi hẹp trước khi triển khai toàn trường.

4.7. Lộ trình thực hiện và vai trò chủ động của người học

Lộ trình thực hiện nên chia thành ba giai đoạn. Trong giai đoạn đầu, nhà trường cần thành lập đầu mối bảo vệ dữ liệu người học, rà soát nhanh các hệ thống trọng yếu, chuẩn hóa chính sách quyền riêng tư và ban hành hướng dẫn tạm thời về sử dụng dữ liệu trong lớp học trực tuyến, thi trực tuyến và công cụ AI. Trong giai đoạn thứ 2, cần hoàn thành bản đồ dữ liệu đối với các hệ thống chính, rà soát hợp đồng công nghệ, thiết lập kênh tiếp nhận yêu cầu của người học và áp dụng thử khung đánh giá tác động đối với hoạt động rủi ro cao. Giai đoạn 3, cần kiểm tra định kỳ, hoàn thiện quy chế nội bộ, công bố báo cáo tuân thủ nội bộ, tập huấn cho cán bộ, giảng viên, người học và cập nhật quy trình theo thay đổi công nghệ.

Để lộ trình có tính quản lý, mỗi giai đoạn cần gắn với sản phẩm đầu ra cụ thể. Giai đoạn đầu nên có quyết định thành lập đầu mối, danh mục hệ thống trọng yếu và mẫu thông báo xử lý dữ liệu cơ bản. Giai đoạn giữa cần có bản đồ dữ liệu đối với các hệ thống ưu tiên, phụ lục bảo vệ dữ

liệu trong hợp đồng công nghệ, kênh tiếp nhận yêu cầu của người học và bộ tiêu chí nhận diện hoạt động rủi ro cao. Giai đoạn sau cần có báo cáo tuân thủ nội bộ, kế hoạch đào tạo định kỳ, danh sách hoạt động đã đánh giá tác động và cơ chế cập nhật khi có nền tảng, nhà cung cấp hoặc công cụ AI mới.

Người học cần được nhìn nhận như chủ thể dữ liệu chủ động, không chỉ là đối tượng được bảo vệ thụ động. Nhà trường cần cung cấp chính sách quyền riêng tư dễ hiểu, hướng dẫn người học cách yêu cầu truy cập, chỉnh sửa, hạn chế xử lý hoặc khiếu nại; đồng thời khuyến khích người học phản hồi khi phát hiện việc ghi âm, ghi hình, chia sẻ dữ liệu hoặc sử dụng AI chưa minh bạch.

Vai trò chủ động của người học cũng cần được cụ thể hóa bằng các công cụ để sử dụng. Trên cổng thông tin hoặc LMS, nhà trường có thể bố trí mục “quyền dữ liệu của người học” gồm chính sách quyền riêng tư tóm tắt, biểu mẫu yêu cầu truy cập hoặc chỉnh sửa dữ liệu, hướng dẫn phản ánh sự cố, thông tin về ghi âm - ghi hình lớp học, thông tin về sử dụng AI và đầu mối liên hệ. Trong các học phần trực tuyến, giảng viên cũng nên nhắc ngắn gọn về nguyên tắc sử dụng dữ liệu lớp học, giới hạn chia sẻ bản ghi và trách nhiệm của sinh viên khi sử dụng công cụ AI.

V. Kết luận

Bài viết đã làm rõ dữ liệu giáo dục không tách biệt khỏi dữ liệu cá nhân. Khi gắn với một người học xác định hoặc có thể xác định, dữ liệu giáo dục là dữ liệu cá nhân; trong một số trường hợp, dữ liệu này trở thành dữ liệu nhạy cảm hoặc rủi ro cao, đặc biệt khi liên quan đến sinh trắc học, sức khỏe, kỷ luật, giám sát thi cử, phân tích hành vi hoặc xử lý tự động bằng AI.

Đối với Trường Đại học Mở Hà Nội, yêu cầu bảo vệ dữ liệu người học gắn trực tiếp với đặc thù đào tạo mở, đào tạo từ xa và đào tạo trực tuyến. Định hướng khả thi là xây dựng quy chế nội bộ, lập bản đồ dữ liệu, chuẩn hóa thông báo và đồng ý, kiểm soát AI, rà soát hợp đồng công nghệ, kiện toàn đầu mối bảo vệ dữ liệu, áp dụng khung đánh giá tác động và trao quyền cho người học với tư cách là chủ thể dữ liệu. Đây không chỉ là yêu cầu tuân thủ pháp luật, mà còn là điều kiện nâng cao niềm tin, trải nghiệm người học và chất lượng của mô hình giáo dục đại học số.

Tài liệu tham khảo

- Chính phủ. (2025). *Nghị định số 356/2025/NĐ-CP quy định chi tiết Luật Bảo vệ dữ liệu cá nhân*. <https://vanban.chinhphu.vn/>
- Chính phủ. (2026). *Nghị định số 141/2026/NĐ-CP quy định chi tiết và hướng dẫn thi hành Luật Trí tuệ nhân tạo*. <https://vanban.chinhphu.vn/>
- Drachsler, H., và Greller, W. (2016). Privacy and analytics: It's a DELICATE issue: A checklist for trusted learning analytics. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (pp. 89-98). Association for Computing Machinery. <https://doi.org/10.1145/2883851.2883893>
- European Parliament and Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. *Official Journal of the European Union*, L119, 1-88.
- Pardo, A., và Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438-450. <https://doi.org/10.1111/bjet.12152>

- Quốc hội. (2013). *Hiến pháp nước Cộng hòa xã hội chủ nghĩa Việt Nam năm 2013*. <https://vanban.chinhphu.vn/>
- Quốc hội. (2015). *Luật An toàn thông tin mạng số 86/2015/QH13*. <https://vanban.chinhphu.vn/>
- Quốc hội. (2018). *Luật số 34/2018/QH14 sửa đổi, bổ sung một số điều của Luật Giáo dục đại học*. <https://vanban.chinhphu.vn/>
- Quốc hội. (2024). *Luật Dữ liệu số 60/2024/QH15*. <https://vanban.chinhphu.vn/>
- Quốc hội. (2025a). *Luật Bảo vệ dữ liệu cá nhân số 91/2025/QH15*. <https://vanban.chinhphu.vn/>
- Quốc hội. (2025b). *Luật Trí tuệ nhân tạo số 134/2025/QH15*. <https://vanban.chinhphu.vn/>
- Slade, S., và Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510-1529. <https://doi.org/10.1177/0002764213479366>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO.
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- United Nations. (1948). *Universal Declaration of Human Rights*. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>
- United Nations. (1966). *International Covenant on Civil and Political Rights*. <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

PROTECTING LEARNER DATA IN ONLINE LEARNING: A CASE STUDY OF HANOI OPEN UNIVERSITY

Nguyen Thi Phan Mai¹, Dong Thi Toan¹, Nguyen Tien Dung¹

Abstract: *Online learning has significantly increased the collection, storage, analysis, and sharing of learners' data, ranging from administrative records and academic results to LMS logs, interaction data, audio-video recordings, biometric data, and AI-generated inferences. This article examines the theoretical and legal foundations of learner data protection, clarifies the relationship among personal data, sensitive data and educational data, and discusses implications for Hanoi Open University.*

Keywords: *personal data protection; learner data; online learning; higher education; artificial intelligence*

¹ Hanoi Open University, Hanoi, Vietnam