

# LỢI ÍCH VÀ THÁCH THỨC CỦA VIỆC ỨNG DỤNG BIG DATA TRONG GIÁO DỤC ĐẠI HỌC

Nguyễn Phan Tinh  
Học viện Ngân hàng

**Tóm tắt:** Trong bối cảnh chuyển đổi số, dữ liệu lớn (Big Data) đã trở thành một nguồn lực quan trọng hỗ trợ đổi mới hoạt động đào tạo và quản trị trong giáo dục đại học. Việc khai thác dữ liệu từ các hệ thống quản lý học tập, hồ sơ học tập và các nền tảng số cho phép các cơ sở giáo dục nâng cao chất lượng ra quyết định, cá nhân hóa trải nghiệm học tập và tối ưu hóa hiệu quả vận hành. Tuy nhiên, cùng với những lợi ích đó là các thách thức liên quan đến quyền riêng tư, bảo mật dữ liệu và tính công bằng của các mô hình phân tích. Bài báo giới thiệu tổng quan về Big Data và phân tích những ứng dụng tiêu biểu trong học tập và quản trị, đồng thời đánh giá các vấn đề đạo đức phát sinh trong quá trình triển khai. Trên cơ sở đó, đề xuất một số khuyến nghị nhằm hỗ trợ các cơ sở giáo dục khai thác dữ liệu một cách hiệu quả và có trách nhiệm.

**Từ khóa:** Dữ liệu lớn, Big Data, giáo dục đại học, phân tích học tập, đạo đức dữ liệu.

## BENEFITS AND CHALLENGES OF APPLYING BIG DATA IN HIGHER EDUCATION

**Abstract:** In the context of digital transformation, Big Data has emerged as a critical resource for enhancing teaching, learning, and institutional management in higher education. By leveraging data generated from learning management systems, academic records, and digital platforms, higher education institutions can improve evidence-based decision-making, personalize learning experiences, and optimize operational performance. This paper reviews the role of Big Data in higher education and examines its major applications in learning analytics and educational management. In addition, it discusses key challenges associated with data privacy, security, and algorithmic fairness. Based on the findings, several recommendations are proposed to support the effective, ethical, and sustainable use of Big Data in higher education.

**Keywords:** Big data, higher education, Learning analytics, Data Ethics.

Nhận bài: 20/04/2026

Phản biện: 23/05/2026

Duyệt đăng: 28/05/2026

### I. ĐẶT VẤN ĐỀ

Sự phát triển mạnh mẽ của công nghệ số đã làm gia tăng đáng kể khối lượng dữ liệu được tạo ra trong các hoạt động giáo dục. Các hệ thống quản lý học tập (Learning Management System - LMS), hệ thống thông tin sinh viên (Student Information System - SIS), nền tảng học trực tuyến và các thiết bị kết nối liên tục tạo ra những dòng dữ liệu có quy mô lớn và tốc độ cập nhật cao. Trong bối cảnh đó, Big Data được xem là nền tảng quan trọng hỗ trợ các cơ sở giáo dục đại học nâng cao chất lượng đào tạo, cải thiện hiệu quả quản trị và tăng cường khả năng thích ứng với những thay đổi của môi trường giáo dục hiện đại.

Việc ứng dụng Big Data không chỉ giúp các trường đại học thu thập và lưu trữ dữ liệu mà còn cho phép chuyển đổi dữ liệu thành tri thức phục vụ quá trình ra quyết định. Thông qua các công cụ phân tích dữ liệu và mô hình dự báo, nhà quản lý có thể nhận diện sớm các vấn đề phát sinh, trong khi giảng viên có thể xây dựng các chiến lược giảng dạy phù hợp với nhu cầu của người học. Tuy nhiên, việc sử dụng dữ liệu quy mô lớn cũng làm gia tăng các quan ngại về quyền riêng tư, tính minh bạch và công bằng trong giáo dục.

Nghiên cứu này nhằm tổng hợp và phân tích các ứng dụng của Big Data trong giáo dục đại học, đồng thời đánh giá những lợi ích, thách thức và yêu cầu quản trị đặt ra trong quá trình triển khai.

### II. NỘI DUNG NGHIÊN CỨU

#### 2.1. Khái niệm và đặc trưng của Big Data trong giáo dục

Big Data được hiểu là các tập dữ liệu có quy mô lớn, tốc độ phát sinh nhanh và cấu trúc đa dạng, vượt quá khả năng xử lý của các hệ thống cơ sở dữ liệu truyền thống.

Trong lĩnh vực giáo dục, dữ liệu được tạo ra từ nhiều nguồn khác nhau như hoạt động học tập trực tuyến, kết quả đánh giá, tương tác trên nền tảng số và các quy trình quản trị.

Các đặc trưng của Big Data thường được mô tả thông qua mô hình 5Vs gồm: Volume (khối lượng dữ liệu lớn), velocity (tốc độ tạo và xử lý dữ liệu cao), variety (sự đa dạng về định dạng dữ liệu), veracity (độ tin cậy của dữ liệu) và value (giá trị thu được từ việc khai thác dữ liệu). Những đặc điểm này tạo điều kiện cho các tổ chức giáo dục triển khai các hoạt động phân tích chuyên sâu nhằm hỗ trợ ra quyết định.

#### 2.2. Các hướng nghiên cứu chủ yếu về Big data trong giáo dục

Dựa trên các tài liệu được thu thập từ các công trình nghiên cứu quốc tế, báo cáo của các cơ sở giáo dục và các nguồn học thuật liên quan, có thể thấy hai hướng nghiên cứu chủ đạo về ứng dụng Big Data trong giáo dục được quan tâm rộng rãi là phân tích học tập (Learning Analytics – LA) và khai phá dữ liệu giáo dục (Educational Data Mining – EDM).

EDM tập trung vào việc áp dụng các thuật toán khai phá dữ liệu và học máy nhằm phát hiện các mô hình, xu hướng hoặc hành vi học tập tiềm ẩn trong dữ liệu giáo dục. Ngược lại, LA nhấn mạnh việc sử dụng các kết quả phân tích để hỗ trợ giảng viên, nhà quản lý và người học đưa ra quyết định phù hợp nhằm nâng cao hiệu quả đào tạo.

Sự kết hợp giữa EDM và LA đã tạo nên tảng cho nhiều ứng dụng như dự báo kết quả học tập, phát hiện nguy cơ bỏ học và xây dựng các hệ thống học tập thích ứng.

### 2.3. Các ứng dụng của Big data trong giáo dục đại học

#### 2.3.1. Cá nhân hóa Trải nghiệm Học tập và Công nghệ Thích ứng

Big Data cho phép xóa bỏ cách tiếp cận “một kích cỡ cho tất cả” (one-size-fits-all). Bằng cách phân tích tương tác và hiệu suất của từng cá nhân, giảng viên có thể điều chỉnh nội dung dạy học phù hợp với nhu cầu cụ thể. Các công nghệ học tập thích ứng như ALEKS hoặc Knewton sử dụng thuật toán máy học để thay đổi việc phân phối nội dung trong thời gian thực dựa trên phản hồi của sinh viên, đóng vai trò như các “gia sư kỹ thuật số”.

Tại Đại học Bang Arizona (ASU), việc áp dụng nền tảng thích ứng ALEKS trong môn Đại số Đại học đã nâng tỷ lệ đạt (điểm C trở lên) từ 57% lên 85%. Tương tự, tại môn kinh tế vi mô, tỷ lệ sinh viên đạt điểm A tăng từ 15% lên 47% sau khi triển khai hệ thống này.

#### 2.3.2. Hệ thống Cảnh báo Sớm và Giữ chân Sinh viên

Một trong những ứng dụng nổi bật của Big Data trong giáo dục đại học là các hệ thống cảnh báo sớm nhằm nhận diện sinh viên có nguy cơ bỏ học. Tại Đại học Purdue, hệ thống Course Signals sử dụng dữ liệu từ LMS và hệ thống thông tin sinh viên để dự báo mức độ rủi ro học tập theo mô hình đèn giao thông (đỏ, vàng, xanh). Kết quả cho thấy tỷ lệ duy trì học tập của sinh viên tham gia ít nhất một học phần có sử dụng Course Signals đạt 87,42%, cao hơn đáng kể so với 69,40% của nhóm đối chứng. Tương tự, Đại học Bang Georgia triển khai hệ thống phân tích hơn 800 chỉ số rủi ro để theo dõi khoảng 40.000 sinh viên mỗi ngày, qua đó góp phần tăng tỷ lệ tốt nghiệp thêm 7 điểm phần trăm và thu hẹp khoảng cách thành tích giữa các nhóm sinh viên. Bên cạnh đó, chương trình Panther Retention Grants sử dụng dữ liệu để xác định sinh viên gặp khó khăn tài chính ngắn hạn và cung cấp hỗ trợ kịp thời, giúp 86% sinh viên thụ hưởng hoàn thành chương trình học.

2.3.3. Tối ưu hóa Quản lý và Vận hành Tổ chức  
Big Data hỗ trợ quản lý tuyển sinh chiến lược (SEM) thông qua mô hình hóa dự báo để tối ưu hóa ngân sách tuyển dụng và giảm tình trạng “summer melt” (sinh viên đã trúng tuyển nhưng không nhập học). Các hệ thống Hoạch định Nguồn lực Doanh nghiệp (ERP) hiện đại kết nối các hệ thống rời rạc như nhân sự, tài chính, và registrar vào một kho lưu trữ thống nhất, giúp tự động hóa theo dõi hàng tồn kho và tối ưu hóa lịch trình khóa học dựa trên nhu cầu lịch sử.

### 2.4. Lợi ích của việc ứng dụng Big data trong giáo dục đại học

Việc ứng dụng Big Data mang lại nhiều lợi ích quan trọng, góp phần nâng cao năng lực cạnh tranh và hiệu quả hoạt động của các cơ sở giáo dục đại học. Thông qua việc thu thập, lưu trữ và phân tích dữ liệu từ các hoạt động học tập, giảng dạy và quản lý, các nhà quản lý có thể đưa ra quyết định nhanh chóng và chính xác hơn dựa trên bằng chứng thực tiễn thay vì kinh nghiệm hoặc nhận định chủ quan. Cách tiếp cận này không chỉ nâng cao chất lượng quản trị mà còn giúp các cơ sở giáo dục chủ động nhận diện xu hướng, dự báo nhu cầu và xây dựng chiến lược phát triển phù hợp.

Bên cạnh đó, Big Data góp phần tối ưu hóa các quy trình vận hành thông qua tự động hóa hoạt động quản lý, phân bổ nguồn lực hợp lý và giảm thiểu chi phí không cần thiết. Việc khai thác hiệu quả dữ liệu giúp nâng cao hiệu suất sử dụng cơ sở vật chất, tối ưu hóa kế hoạch đào tạo và cải thiện các hoạt động hỗ trợ sinh viên. Nhờ đó, các nguồn lực về tài chính, nhân sự và hạ tầng được sử dụng hiệu quả hơn, góp phần giảm lãng phí và nâng cao hiệu quả hoạt động của tổ chức.

### 2.5. Thách thức

Tuy mang lại nhiều lợi ích nhưng việc sử dụng Big Data trong giáo dục đại học cũng đặt ra những thách thức to lớn:

#### 2.5.1. Định kiến thuật toán và công bằng

Mặc dù các thuật toán học máy hỗ trợ hiệu quả trong phân tích và dự báo dữ liệu giáo dục, chúng cũng có nguy cơ tái tạo và duy trì những bất bình đẳng vốn tồn tại trong dữ liệu lịch sử. Do được huấn luyện trên các tập dữ liệu phản ánh sự khác biệt về điều kiện kinh tế, xã hội và học tập, các mô hình dự báo có thể đánh giá sai lệch mức độ rủi ro của sinh viên thuộc các nhóm thiểu số hoặc có hoàn cảnh khó khăn. Bên cạnh đó, các hệ thống tư vấn dựa trên dữ liệu có thể dẫn đến hiện tượng “định hướng giáo dục” (educational steering), khi sinh viên thuộc nhóm yếu thế được khuyến nghị lựa chọn các ngành học ít cạnh tranh hơn thay vì các lĩnh vực có yêu cầu cao như STEM, qua đó vô tình duy trì các bất bình đẳng mang tính cấu trúc.

### 2.5.2. Quyền riêng tư và giám sát dữ liệu

Việc thu thập và phân tích dữ liệu quy mô lớn đòi hỏi theo dõi nhiều khía cạnh trong hoạt động học tập của sinh viên, bao gồm mức độ tương tác trên các nền tảng số, việc sử dụng thiết bị và các hành vi trực tuyến. Mặc dù những dữ liệu này hỗ trợ nâng cao chất lượng đào tạo và quản lý, chúng cũng làm gia tăng các quan ngại về quyền riêng tư và quyền tự chủ cá nhân. Vì vậy, các cơ sở giáo dục cần xây dựng cơ chế quản trị dữ liệu minh bạch và tuân thủ các quy định bảo vệ dữ liệu cá nhân như FERPA hoặc GDPR nhằm bảo đảm quyền lợi của người học.

### 2.5.3. Yếu tố con người và sự chấp nhận công nghệ

Big Data là công cụ hỗ trợ, không thay thế vai trò của con người trong quá trình ra quyết định giáo dục. Tuy nhiên, việc triển khai các hệ thống phân tích dữ liệu thường gặp trở ngại do lo ngại về tính minh bạch của các thuật toán “hộp đen” và khả năng ảnh hưởng đến quyền tự chủ học thuật. Đồng thời, việc truyền đạt các kết quả dự báo rủi ro cần được thực hiện thận trọng nhằm tránh tạo áp lực tâm lý hoặc tác động tiêu cực đến động lực học tập của sinh viên. Do đó, sự kết hợp giữa phân tích dữ liệu và đánh giá của con người vẫn là điều kiện cần thiết để bảo đảm tính hiệu quả, công bằng và nhân văn trong giáo dục đại học.

### 2.6. Khuyến nghị

Để khai thác Big Data hiệu quả và có trách nhiệm, các cơ sở giáo dục đại học cần xây dựng khung quản trị dữ liệu và trí tuệ nhân tạo toàn diện. Trước hết, việc thành lập Hội đồng Đạo đức AI là cần thiết nhằm giám sát quá trình triển khai các hệ thống phân tích, đánh giá tính công bằng và

thực hiện kiểm toán định kỳ để giảm thiểu nguy cơ thiên lệch thuật toán. Đồng thời, các cơ sở giáo dục cần ưu tiên tính minh bạch của thuật toán, hạn chế sử dụng các hệ thống “hộp đen” và yêu cầu nhà cung cấp công bố đặc điểm dữ liệu huấn luyện cũng như hiệu suất mô hình trên các nhóm đối tượng khác nhau.

Bên cạnh đó, hiệu quả của Big Data chỉ được phát huy khi kết hợp với các nguồn lực hỗ trợ từ con người như giảng viên, cố vấn học tập và các chương trình hỗ trợ sinh viên. Việc nâng cao năng lực dữ liệu cho đội ngũ quản lý và giảng viên cũng đóng vai trò quan trọng trong việc khai thác, diễn giải và sử dụng kết quả phân tích một cách phù hợp. Trong tương lai, các cơ sở giáo dục cần tiếp tục hoàn thiện cơ chế quản trị dữ liệu và thúc đẩy nghiên cứu về đạo đức AI nhằm bảo đảm việc ứng dụng Big Data phục vụ hiệu quả, công bằng và bền vững cho giáo dục đại học.

### III. KẾT LUẬN

Big Data đang mở ra những cơ hội quan trọng cho giáo dục đại học thông qua việc nâng cao chất lượng đào tạo, tối ưu hóa quản trị và hỗ trợ ra quyết định dựa trên bằng chứng. Những ứng dụng như học tập cá nhân hóa, hệ thống cảnh báo sớm và phân tích dự báo đã chứng minh tiềm năng cải thiện kết quả học tập cũng như hiệu quả vận hành của các cơ sở giáo dục. Tuy nhiên, việc triển khai Big Data cần đi kèm với các cơ chế quản trị phù hợp nhằm bảo đảm quyền riêng tư, tính công bằng và trách nhiệm đạo đức. Chỉ khi kết hợp hài hòa giữa công nghệ, quản trị và yếu tố con người, Big Data mới có thể phát huy đầy đủ giá trị trong quá trình phát triển giáo dục đại học bền vững.

### TÀI LIỆU THAM KHẢO

- Bird, K. (2023). Predictive analytics in higher education: The promises and challenges of using machine learning to improve student success. *The AIR Professional File*, (161). <https://doi.org/10.34315/apf1612023>
- Cambridge Spark. (2024). 5 ways to use big data to enhance higher education. <https://www.cambridgespark.com/blog/5-ways-to-use-big-data-to-enhance-higher-education>
- EAB. (n.d.). Data management best practices for higher education. <https://eab.com/higher-education-glossary/data-management-for-higher-education/>
- Evans, M. (2026). The power of prediction: How early warning systems are changing student retention. *Insight Into Academia*. <https://insightintoacademia.com/the-power-of-prediction/>
- I'MNOVATION #Hub. (n.d.). How big data is shaping the future of education. <https://www.imnovation-hub.com/digital-transformation/big-data-applications-education>
- Learning Guild. (2025). The impact of learner data on adaptive AI-driven learning systems. <https://www.learningguild.com/articles/the-impact-of-learner-data-on-adaptive-ai-driven-learning-systems>
- Lesjak, D., Natek, S., & Kohun, F. (2021). Big data analytics in higher education. *Issues in Information Systems*, 22(4), 320-333. [https://doi.org/10.48009/4\\_iis\\_2021\\_346-359](https://doi.org/10.48009/4_iis_2021_346-359)
- Mapademics. (2025). Transforming student success: How predictive retention analytics is revolutionizing higher education in 2025. <https://www.mapademics.com/resources/whitepapers/predicting-student-success-with-smart-analytics>
- Markdown Source. (n.d.). Big Data Analytics in Higher Education: Systemic Applications, Technological Architectures, and Ethical Governance Frameworks.
- Merrimack College. (2025). Big data in education: How it's transforming learning in 2025. <https://online.merrimack.edu/big-data-in-education/>
- Zhou, J., & Li, X. (2024). Application of big data in college student education management based on data warehouse technology and integrated learning. *International Journal of e-Collaboration*, 20(1). <https://doi.org/10.4018/IJeC.346368>