

XÂY DỰNG KHO DỮ LIỆU TÍCH HỢP TRƯỜNG ĐẠI HỌC THỦ DẦU MỘT

Trịnh Trọng Thành ⁽¹⁾, Bùi Thanh Khiết ⁽¹⁾

(1) Trường Đại học Thủ Dầu Một

Ngày nhận bài: 15/01/2026; Chấp nhận đăng: 28/02/2026

Email tác giả: khietbt@tdmu.edu.vn

Tóm tắt

Trong bối cảnh tự chủ đại học và chuyển đổi số, việc ra quyết định dựa trên dữ liệu (Data-driven decision making) trở thành yêu cầu cấp thiết. Tuy nhiên, thách thức lớn nhất tại các trường đại học hiện nay là tình trạng phân mảnh dữ liệu, thiếu tính liên thông giữa các phòng ban chức năng. Bài báo này trình bày quy trình xây dựng và triển khai hệ thống Kho dữ liệu (Data Warehouse - DWH) tích hợp tại Trường Đại học Thủ Dầu Một. Nghiên cứu đề xuất kiến trúc hệ thống hiện đại dựa trên mô hình Lakehouse, kết hợp sức mạnh lưu trữ của Data Lake và khả năng quản lý cấu trúc của DWH, sử dụng công nghệ Apache Doris cho xử lý dữ liệu lớn và Apache Superset cho trực quan hóa. Kết quả thực nghiệm cho thấy hệ thống đã tích hợp thành công dữ liệu từ các phân hệ đào tạo, tuyển sinh và nghiên cứu khoa học, cung cấp các dashboard phân tích đa chiều thời gian thực. Hệ thống giúp lãnh đạo nhà trường có cái nhìn toàn cảnh, từ đó nâng cao hiệu quả quản trị và hoạch định chiến lược.

Từ khóa: *Apache Doris, Data Warehouse, kho dữ liệu, quản trị đại học.*

Abstract

DEVELOPMENT OF AN INTEGRATED DATA WAREHOUSE AT THU DAU MOT UNIVERSITY

In the context of university autonomy and digital transformation, data-driven decision-making has become an urgent requirement. However, the most significant challenge facing universities today is data fragmentation and the lack of interoperability between functional departments. This paper presents the process of building and deploying an integrated Data Warehouse (DWH) system at Thu Dau Mot University. The study proposes a modern system architecture based on the Lakehouse model, which combines the flexible storage capabilities of a Data Lake with the structural management of a traditional DWH. The system utilizes Apache Doris technology for high-speed big data processing and Apache Superset for data visualization. Experimental results demonstrate that the system has successfully integrated data from training, admissions, and scientific research modules, providing real-time multi-dimensional analysis dashboards. This system enables university leadership to gain a comprehensive view of operations, thereby enhancing governance efficiency and strategic planning.

1. Giới thiệu

Tự chủ đại học được xem là xu thế phát triển tất yếu, là điều kiện cần để các trường đại học tồn tại và mở ra nhiều cơ hội phát triển. Muốn thực hiện được tự chủ đại

học, đòi hỏi các trường đại học cần có những chiến lược quản trị đại học chuyển đổi sang tổ chức dựa trên dữ liệu (Huisman et al., 2024). Trong bối cảnh hiện nay, ứng dụng công nghệ thông tin không chỉ dừng lại ở mức độ kỹ thuật mà đã trở thành nền tảng cốt lõi trong quản trị đại học. Việc xây dựng hệ thống dữ liệu đồng bộ giúp nâng cao hiệu quả điều hành của bộ máy quản lý, đồng thời cải thiện trải nghiệm của người học thông qua việc minh bạch hóa thông tin. Sinh viên có thể chủ động theo dõi lộ trình đào tạo và các yêu cầu nghiệp vụ từ lúc nhập học cho đến khi hoàn thành chương trình, tạo tiền đề vững chắc cho sự nghiệp sau ra trường.

Trong những năm gần đây, CNTT đã được ứng dụng mạnh mẽ vào nhiều lĩnh vực như quản lý tuyển sinh, đào tạo, nghiên cứu khoa học và hợp tác doanh nghiệp. Tuy nhiên, thực tế cho thấy đa số các hệ thống phần mềm hỗ trợ quản trị hiện nay còn thiếu tính liên thông. Dữ liệu thường bị phân tán ở nhiều định dạng khác nhau và không được đồng bộ giữa các đơn vị. Các giải pháp tích hợp dữ liệu truyền thống thường gặp khó khăn trong việc xử lý khối lượng dữ liệu lớn và đa dạng.

Để giải quyết vấn đề này, việc xây dựng một Kho dữ liệu (Data Warehouse) tích hợp là giải pháp cấp thiết. Nghiên cứu này tập trung xây dựng hệ thống DWH tại Trường Đại học Thủ Dầu Một nhằm tích hợp dữ liệu từ nhiều nguồn, chuẩn hóa và cung cấp công cụ phân tích trực quan, hỗ trợ đắc lực cho quá trình ra quyết định của lãnh đạo nhà trường.

2. Kiến trúc xây dựng kho dữ liệu

2.1. Cơ sở lý luận về kho dữ liệu

Kho dữ liệu (Data Warehouse – DWH) là hệ thống lưu trữ dữ liệu có tổ chức, được thiết kế chuyên biệt để hỗ trợ truy vấn, phân tích và ra quyết định. Theo Inmon, DWH là tập hợp dữ liệu có tính hướng chủ đề, tích hợp, không biến động và có yếu tố thời gian – đáp ứng nhu cầu phân tích quản lý trong tổ chức (Inmon, 2005). So với cơ sở dữ liệu giao dịch truyền thống, kho dữ liệu tập trung vào dữ liệu lịch sử và khả năng phân tích xu hướng dài hạn.

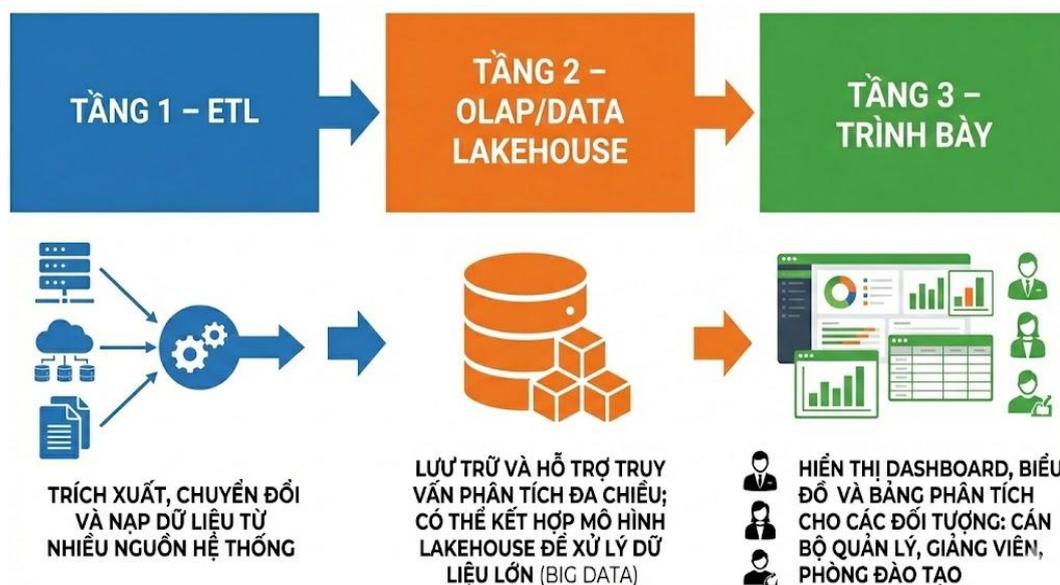
Trong bối cảnh giáo dục đại học, DWH giúp tích hợp dữ liệu từ nhiều hệ thống nghiệp vụ như SIS (quản lý sinh viên), LMS (quản lý học tập), hệ thống thư viện, tài chính, giúp đồng bộ hóa và chuẩn hóa dữ liệu phục vụ phân tích và quản trị tổng thể (Ismaili & Besimi, 2024).

2.2. Mô hình kiến trúc kho dữ liệu hiện đại

Hiện nay, kiến trúc DWH trong giáo dục hiện đại thường được xây dựng theo mô hình ba tầng:

Về mặt kỹ thuật, kiến trúc của kho dữ liệu thường được chia thành ba tầng chính:

- Tầng 1 – ETL: Trích xuất, chuyển đổi và nạp dữ liệu từ nhiều nguồn hệ thống.
- Tầng 2 – OLAP/Data Lakehouse: Lưu trữ và hỗ trợ truy vấn phân tích đa chiều; có thể kết hợp mô hình Lakehouse để xử lý dữ liệu lớn (big data).
- Tầng 3 – Trình bày: Hiện thị dashboard, biểu đồ và bảng phân tích cho các đối tượng như cán bộ quản lý, giảng viên, phòng đào tạo (Sharma & Singh, 2023).



Hình 1: Kiến trúc DWH trong giáo dục hiện đại

Nguồn: Tác giả, 2025

Mô hình kiến trúc tổng thể của hệ thống kho dữ liệu. Luồng dữ liệu được vận hành khép kín từ khâu tích hợp và chuẩn hóa đa nguồn (ETL), lưu trữ và xử lý phân tích trọng tâm (OLAP/Lakehouse) đến khâu trực quan hóa dữ liệu (Presentation), đảm bảo cung cấp thông tin chi tiết và chính xác cho các cấp quản lý.

2.3. Đề xuất quy trình triển khai DWH trong cơ sở giáo dục

Dựa trên tổng hợp tài liệu và thực tiễn, bài viết đề xuất quy trình triển khai DWH gồm các bước sau:

- Xác định mục tiêu chiến lược và phạm vi dữ liệu.
- Lập kế hoạch kiến trúc hệ thống và lựa chọn công nghệ phù hợp (Apache Doris, Snowflake, BigQuery...).
- Xây dựng quy trình ETL và chuẩn hóa dữ liệu từ các hệ thống SIS, LMS, tài chính.
- Triển khai tầng phân tích (OLAP) và công cụ trực quan hóa (Power BI, Superset).
- Thiết lập chính sách bảo mật, phân quyền truy cập.
- Đánh giá và cải tiến hệ thống theo phản hồi từ người dùng.

2.4. Xu hướng công nghệ mới trong triển khai kho dữ liệu giáo dục

Bên cạnh kiến trúc kho dữ liệu truyền thống, các mô hình hiện đại như Data Lake và Data Lakehouse đang ngày càng được quan tâm trong lĩnh vực giáo dục, nhất là khi các trường đại học cần xử lý dữ liệu lớn từ nhiều nguồn khác nhau (video học trực tuyến, dữ liệu hành vi sinh viên, log từ hệ thống LMS...).

Data Lake cho phép lưu trữ dữ liệu phi cấu trúc và bán cấu trúc ở định dạng gốc, tạo tiền đề cho phân tích nâng cao bằng AI/ML.

Lakehouse kết hợp ưu điểm của DWH và Data Lake: lưu trữ linh hoạt như Data Lake, nhưng vẫn có schema và hỗ trợ truy vấn nhanh như DWH (Rossi et al., 2025).

Một hướng đi đáng chú ý khác là xây dựng trực dữ liệu (Data Fabric hoặc Data Mesh), giúp các đơn vị trong trường (phòng đào tạo, thư viện, tài chính...) vừa chia sẻ dữ liệu vừa đảm bảo quyền riêng tư và quản trị phân tán hiệu quả.

Data Fabric sử dụng AI để kết nối dữ liệu từ nhiều nguồn, tự động hóa việc chuẩn hóa và chia sẻ dữ liệu toàn trường.

Data Mesh đề xuất tổ chức dữ liệu theo miền chức năng (domain-driven), trao quyền làm chủ dữ liệu cho từng bộ phận – đây là xu hướng phù hợp với các trường đại học lớn đang phân quyền quản lý mạnh mẽ (Rongnaimuang et al., 2024).

Việc kết hợp DWH truyền thống với các mô hình mới như trên không chỉ nâng cao hiệu quả lưu trữ và phân tích, mà còn mở rộng khả năng tích hợp các công nghệ như trí tuệ nhân tạo, học máy và dữ liệu thời gian thực trong giáo dục đại học hiện đại.

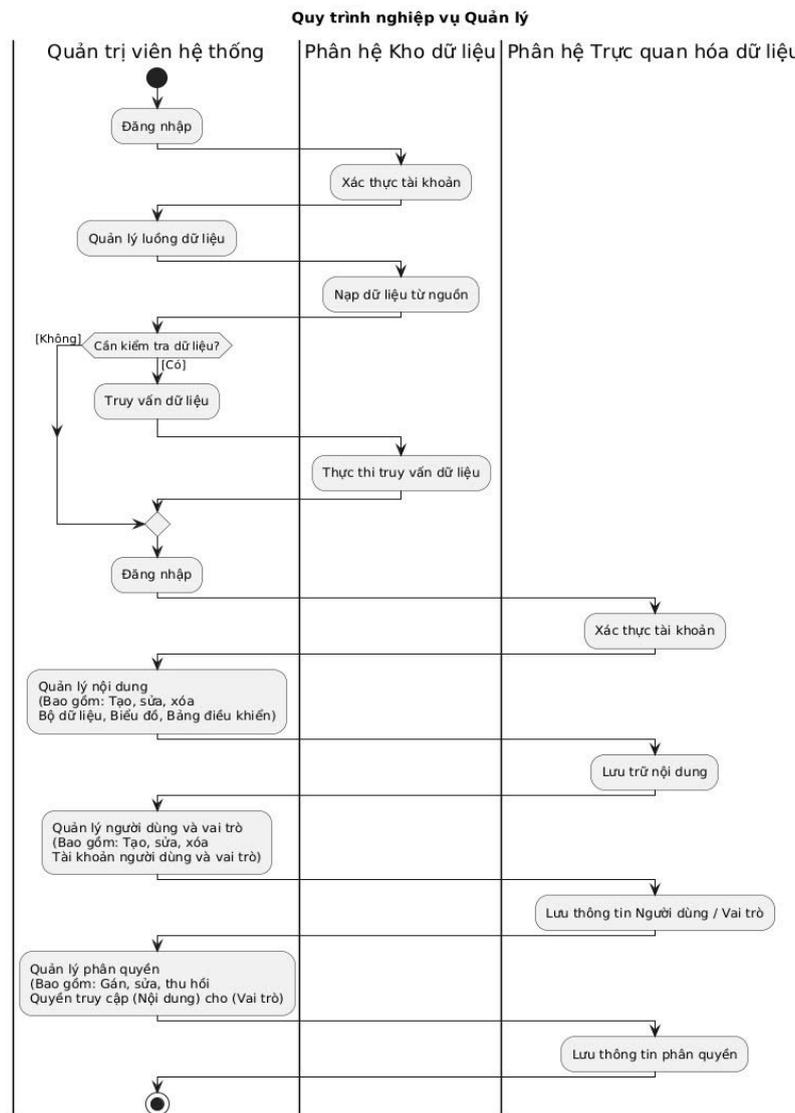
3. Thiết kế hệ thống kho dữ liệu

3.1. Quy trình quản lý của hệ thống

Bước đầu tiên là “Quản lý luồng dữ liệu”, quản trị viên hệ thống sẽ đăng nhập vào phân hệ kho dữ liệu sau đó sử dụng các nguồn dữ liệu đã làm sạch từ các nguồn dữ liệu của trường và “Nạp dữ liệu” vào hệ thống kho dữ liệu đảm bảo tính toàn vẹn và chất lượng dữ liệu tự động (Tamm & Nikiforova, 2024). Ngoài ra, quản trị viên hệ thống có thể dùng thêm các chức năng như “Truy vấn dữ liệu” để kiểm tra, xác thực dữ liệu sau khi nạp và chuẩn bị cho việc chuyển đổi dữ liệu để đưa vào phân hệ trực quan hóa dữ liệu.

Bước thứ hai của quy trình quản lý sẽ là quản lý nội dung báo cáo. Quản trị viên hệ thống sẽ đăng nhập vào phân hệ trực quan hóa dữ liệu sử dụng chức năng “Quản lý nội dung” để tạo các bộ dữ liệu từ nguồn dữ liệu đã được làm sạch, chuyển đổi từ phân hệ kho dữ liệu lên phân hệ trực quan hóa dữ liệu. Sau khi đã tạo các bộ dữ liệu, dùng chức năng tạo biểu đồ để tạo các biểu đồ phù hợp với nội dung chủ đề báo cáo. Cuối cùng là tổng hợp các biểu đồ thuộc cùng một chủ đề thành một bảng báo cáo hoàn chỉnh.

Bước ba, bước cuối của quy trình quản lý là quản lý truy cập. Vẫn trên phân hệ trực quan hóa dữ liệu, sau khi đã có các bảng báo cáo hoàn chỉnh thì quản trị viên hệ thống sẽ bắt đầu sử dụng chức năng “Quản lý người dùng và vai trò” để tạo tài khoản cho các người dùng và gom nhóm họ vào các vai trò (ví dụ như: “Ban giám hiệu”, “Ban Quản lý Đào tạo”). Tiếp theo, quản trị viên hệ thống sử dụng chức năng “Quản lý phân quyền” để gán quyền (ví dụ như: cho phép vai trò “Ban giám hiệu” được xem các báo cáo tuyển sinh).

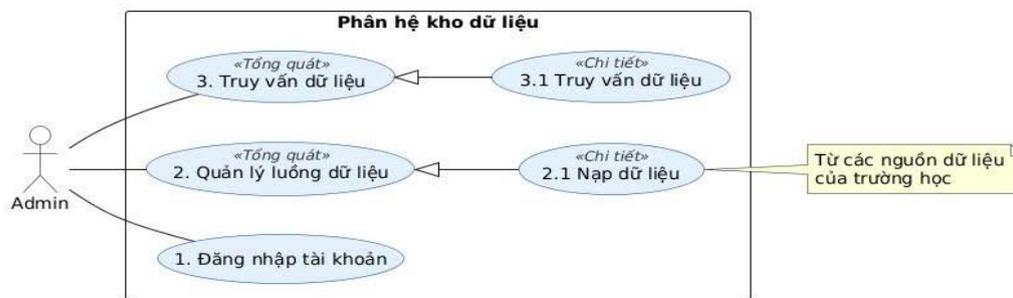


Hình 2: Quy trình nghiệp vụ quản lý

Nguồn: Tác giả, 2025

Đây là quy trình "chuẩn bị" dữ liệu và nội dung, được thực hiện bởi quản trị viên hệ thống. Quy trình này đi từ dữ liệu thô đến báo cáo sẵn sàng cho người dùng cuối.

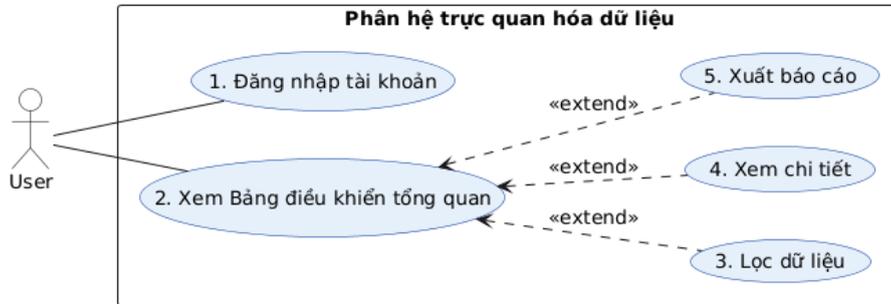
3.2. Biểu đồ usecase



Hình 3: Use case Quy trình Quản lý Phân hệ kho dữ liệu

Nguồn: Tác giả, 2025

Biểu đồ này mô tả các tương tác của Quản trị viên hệ thống với Phân hệ kho dữ liệu. Đây là các chức năng kỹ thuật cốt lõi ở nền tảng dữ liệu, bao gồm Đăng nhập, Quản lý luồng dữ liệu (cụ thể là Nạp dữ liệu từ nguồn) và khả năng Truy vấn dữ liệu để kiểm tra và xác thực.

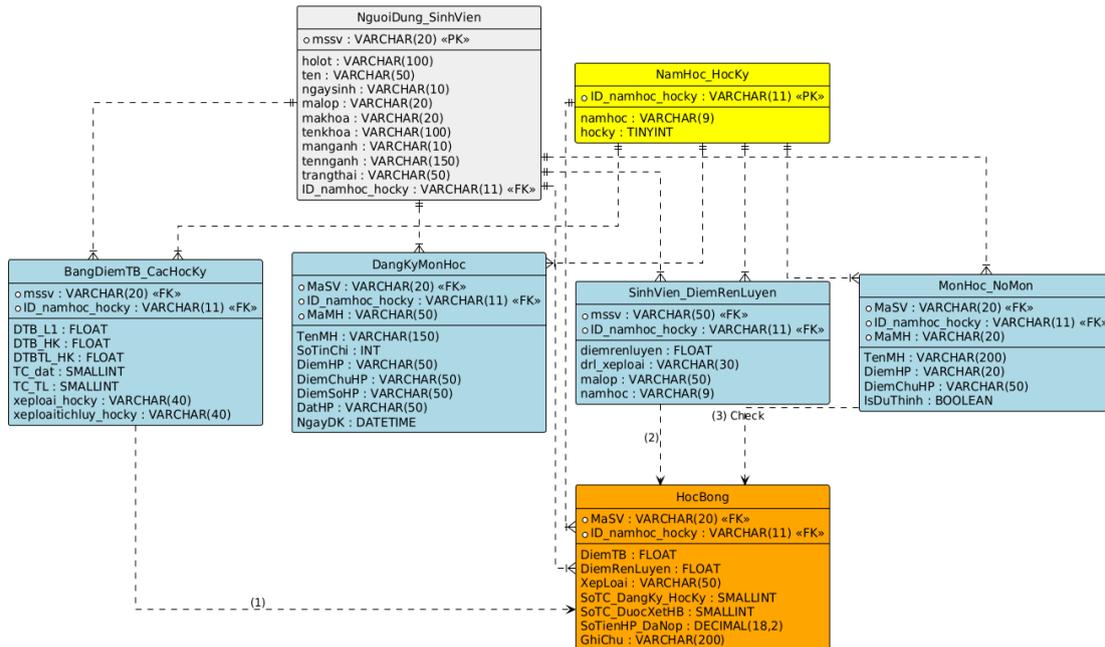


Hình 4: Use case Quy trình người dùng Phân hệ trực quan hóa dữ liệu

Nguồn: Tác giả, 2025

Biểu đồ này mô tả các tương tác của Người dùng cuối với Phân hệ Trực quan hóa dữ liệu. Biểu đồ thể hiện luồng nghiệp vụ khai thác thông tin chính: Đăng nhập và Xem Bảng điều khiển tổng quan. Các chức năng mở rộng cho phép người dùng tương tác sâu hơn với dữ liệu thông qua Lọc dữ liệu, Xem chi tiết và Xuất báo cáo.

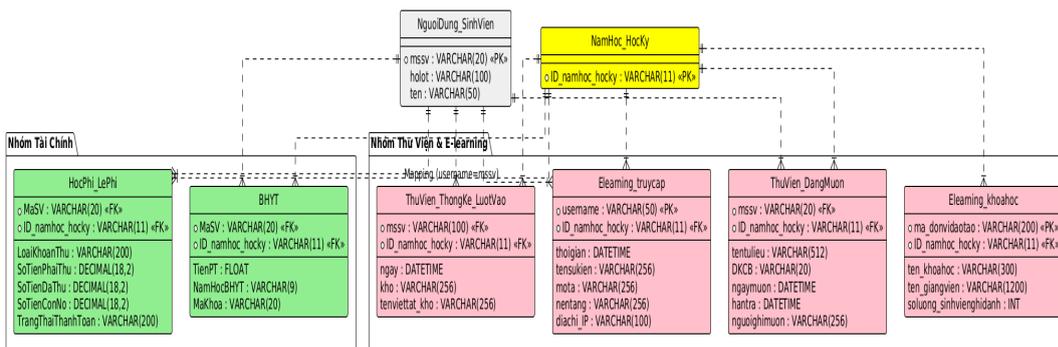
3.3. Mô hình thực thể kết hợp



Hình 5: Mô hình thực thể kết hợp 01

Nguồn: Tác giả, 2025

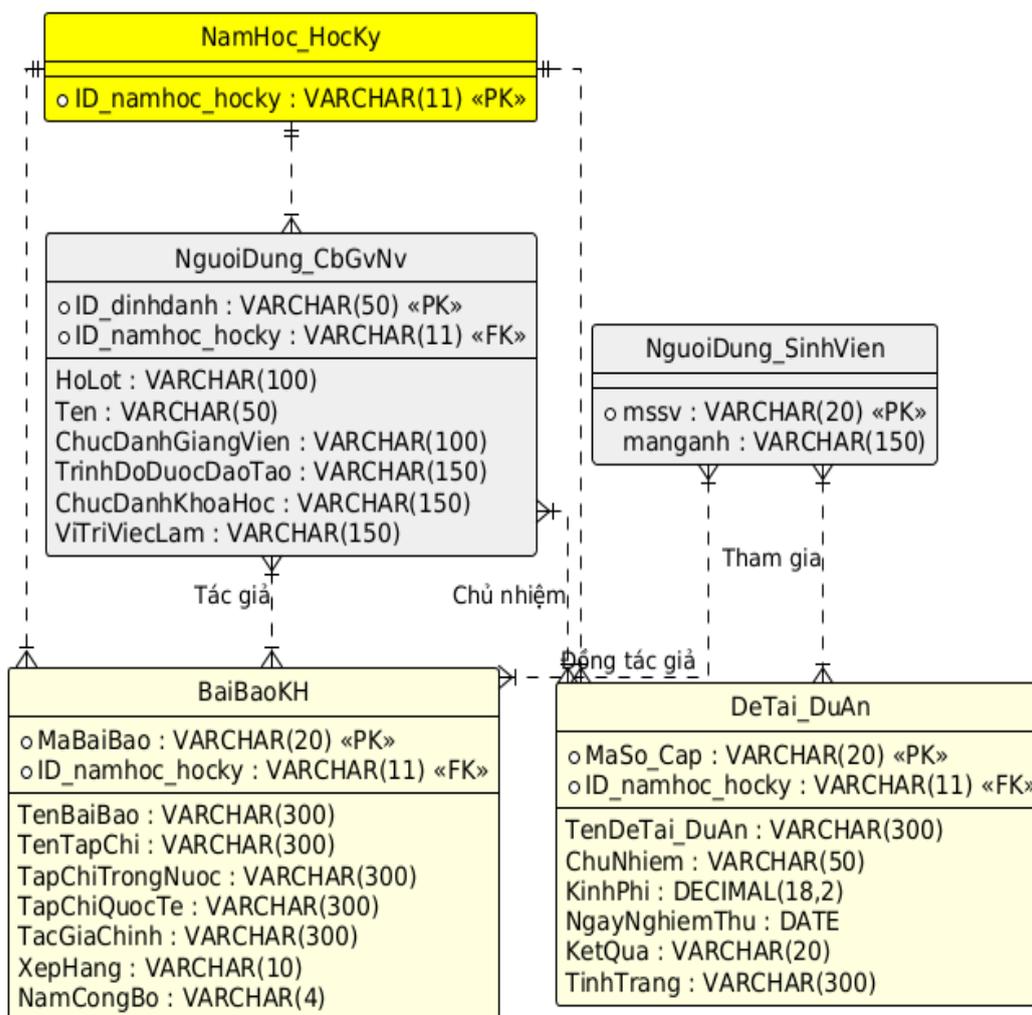
Mô hình mô tả hệ thống quản lý học tập và xét học bổng, bao gồm thông tin sinh viên, kết quả học tập, rèn luyện và phúc lợi. Quy trình xét học bổng dựa trên GPA, điểm rèn luyện, tín chỉ và nợ môn.



Hình 6: Mô hình thực thể kết hợp 02

Nguồn: Tác giả, 2025

Sơ đồ quản lý nghĩa vụ tài chính và hoạt động hỗ trợ của sinh viên, bao gồm quản lý công nợ, học phí, Bảo hiểm y tế, học tập trực tuyến và lưu thông thư viện. Hệ thống sử dụng cơ chế Mapping để đồng bộ hóa danh tính sinh viên.



Hình 7: Mô hình thực thể kết hợp 03

Nguồn: Tác giả, 2025

Mô hình quản lý nghiên cứu khoa học đại học kết nối giảng viên, sinh viên với sản phẩm học thuật. Mô hình vận hành hai quy trình: công bố bài báo khoa học và triển khai đề tài/dự án, với dữ liệu được lưu trữ theo học kỳ/năm học.

4. Thực nghiệm và đánh giá kết quả

Hệ thống đã được triển khai thực tế tại Trường Đại học Thủ Dầu Một với các phân hệ chức năng chính sau:

4.1. Khả năng Tích hợp và Quản lý dữ liệu (ETL)

Phân hệ quản lý kho dữ liệu đóng vai trò cốt lõi trong việc đảm bảo tính toàn vẹn của dữ liệu đầu vào. Hệ thống hỗ trợ cơ chế nạp dữ liệu linh hoạt từ nhiều định dạng phổ biến như CSV, Excel, cho phép tích hợp dữ liệu lịch sử từ các hệ thống cũ.

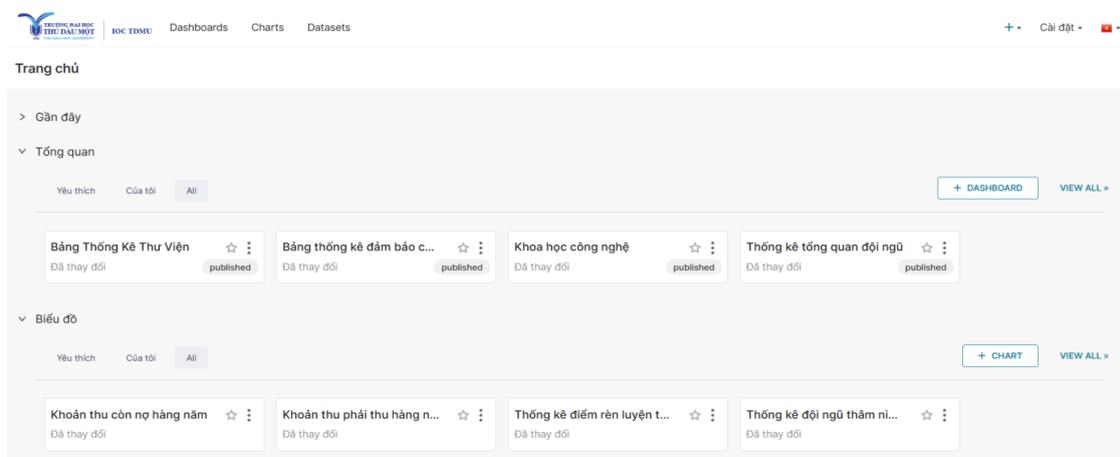
Quản trị viên hệ thống sẽ đăng nhập vào giao diện kho dữ liệu này bằng tài khoản và mật khẩu riêng để quản lý. Hệ thống kho dữ liệu cho phép nạp dữ liệu từ các tệp có định dạng thông dụng, đặc biệt là CSV.

Thay vì các thao tác thủ công rời rạc, quản trị viên có thể kiểm soát quy trình dòng chảy dữ liệu (Data Flow) thông qua giao diện quản trị tập trung. Dữ liệu sau khi được trích xuất (Extract) sẽ trải qua quá trình làm sạch và chuyển đổi (Transform) trước khi được nạp (Load) vào Apache Doris. Kiến trúc này giúp chuẩn hóa dữ liệu từ các nguồn khác nhau về một định dạng thống nhất, sẵn sàng cho việc phân tích. Trước khi tải tệp lên hệ thống, tệp dữ liệu cần phải chuẩn hóa về định dạng file CSV.

4.2. Hệ thống Báo cáo thông minh và Trực quan hóa

Dựa trên nền tảng công nghệ trực quan hóa hiện đại, hệ thống cung cấp khả năng tạo lập các báo cáo động (Dynamic Reports). Người dùng có thể tương tác trực tiếp với dữ liệu thông qua các thao tác lọc (Filter), xem chi tiết (Drill-down) hoặc xuất dữ liệu ra các định dạng PDF/Image để phục vụ báo cáo văn bản.

Tính năng phân quyền (Role-based Access Control) được áp dụng chặt chẽ, đảm bảo rằng từng nhóm người dùng (Ban Giám hiệu, Phòng Đào tạo, Khoa/Viện) chỉ tiếp cận được những dữ liệu phù hợp với chức năng nhiệm vụ.



Hình 8: Giao diện Trực quan hóa dữ liệu

Nguồn: Tác giả, 2025

Giao diện này là công cụ trực quan hóa dữ liệu, cho phép người dùng tạo ra các biểu đồ, đồ thị và bảng điều khiển trực quan để phân tích và hiểu rõ hơn về dữ liệu. Công cụ này giúp biến dữ liệu thành thông tin có thể dễ dàng diễn giải và sử dụng.

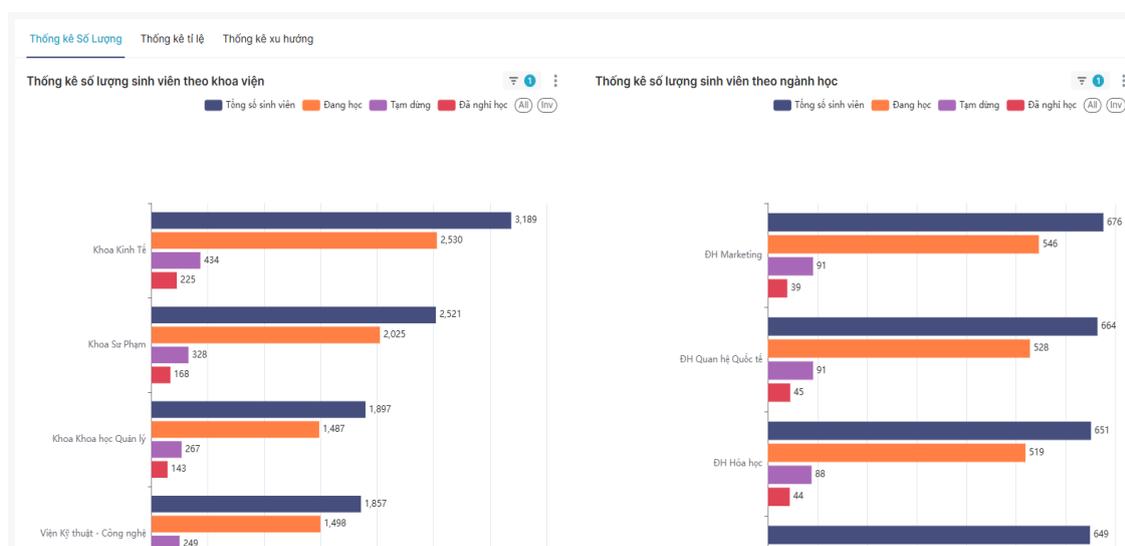
Bên cạnh khả năng tương tác trực tuyến, hệ thống cung cấp công cụ xuất báo cáo linh hoạt, cho phép chuyển đổi các dashboard phân tích động thành các định dạng tài liệu tĩnh tiêu chuẩn (PDF, hình ảnh). Tính năng này đáp ứng nhu cầu thiết yếu trong quy trình hành chính của nhà trường, hỗ trợ cán bộ quản lý dễ dàng trích xuất số liệu cho các báo cáo định kỳ, hồ sơ kiểm định chất lượng hoặc tài liệu trình chiếu tại các cuộc họp. Việc tích hợp này đảm bảo tính nhất quán và chính xác giữa dữ liệu trên hệ thống và các văn bản báo cáo được ban hành.

4.3. Phân tích dữ liệu thực tế tại Trường Đại học Thủ Dầu Một

Kết quả vận hành hệ thống đã cung cấp những thông tin quản trị giá trị thông qua các nhóm dữ liệu chính:

4.3.1. Phân tích dữ liệu Tuyển sinh

Dashboard tuyển sinh cung cấp bức tranh toàn cảnh về quy mô và chất lượng đầu vào. Các biểu đồ phân tích sâu (Hình 9) cho thấy sự phân bố sinh viên theo Khoa/Viện và Ngành học. Dữ liệu này giúp nhà trường nhận diện rõ các ngành thu hút người học (như Khối ngành Kinh tế, Kỹ thuật - Công nghệ) và các ngành cần có chiến lược tuyển sinh mới. Bên cạnh đó, các thống kê về nhân khẩu học (giới tính, vùng miền) cũng là cơ sở quan trọng cho công tác công tác sinh viên.



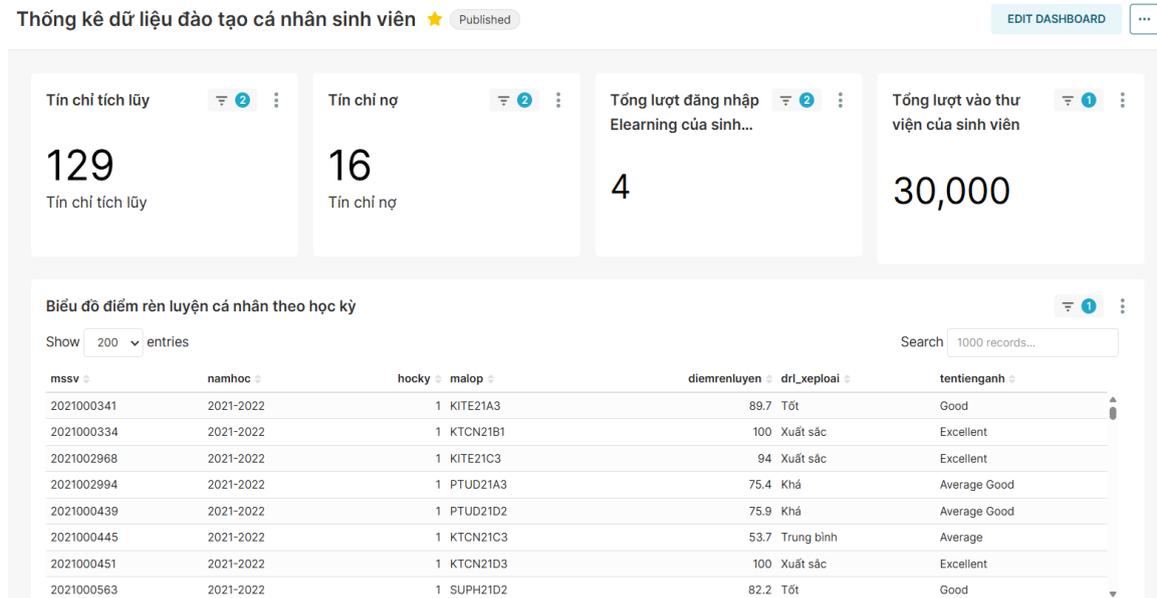
Hình 9: Thống kê tuyển sinh theo số lượng

Nguồn: Tác giả, 2025

4.3.2. Phân tích hiệu quả Đào tạo

Hệ thống cho phép theo dõi sát sao hoạt động học tập của sinh viên. Đặc biệt trong bối cảnh chuyển đổi số, dữ liệu từ hệ thống E-learning được tích hợp để đánh giá mức độ tham gia của người học. Biểu đồ xu hướng truy cập (Hình 10) cho thấy tần suất tương tác của sinh viên với hệ thống học tập trực tuyến, giúp giảng viên và nhà quản lý kịp thời điều chỉnh phương pháp giảng dạy. Ngoài ra, hệ thống còn cung cấp hồ sơ học tập cá nhân hóa:

tổng hợp từ tín chỉ tích lũy, điểm rèn luyện đến tình trạng nợ môn, giúp cố vấn học tập hỗ trợ sinh viên kịp thời.

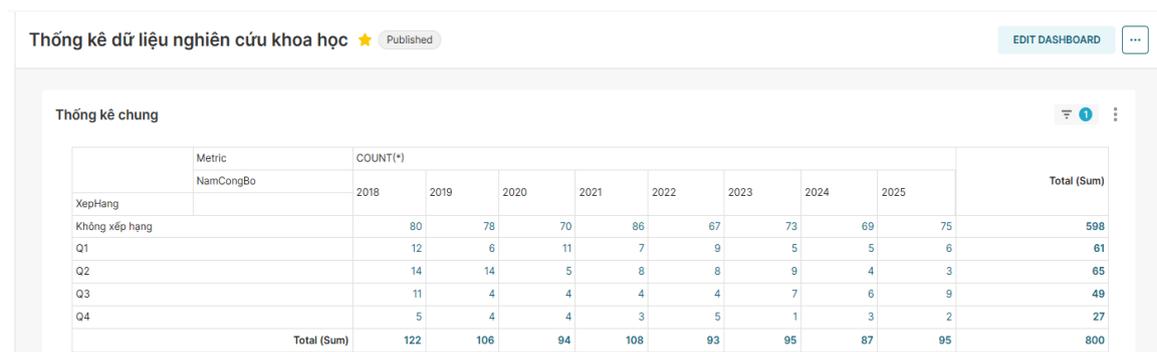


Hình 10: Thống kê dữ liệu đào tạo cá nhân sinh viên

Nguồn: Tác giả, 2025

4.3.3. Quản trị Nghiên cứu khoa học

Dữ liệu về công bố khoa học được thống kê chi tiết theo từng năm và phân loại theo xếp hạng tạp chí (Q1, Q2...). Biểu đồ thống kê (Hình 11) cho thấy xu hướng phát triển năng lực nghiên cứu của nhà trường, hỗ trợ việc khen thưởng và hoạch định chính sách phát triển đội ngũ giảng viên.



Hình 11: Thống kê dữ liệu nghiên cứu khoa học

Nguồn: Tác giả, 2025

5. Kết luận

Bài báo đã trình bày thành công quy trình xây dựng và triển khai hệ thống kho dữ liệu tích hợp tại Trường Đại học Thủ Dầu Một. Việc áp dụng mô hình kiến trúc hiện đại và các công nghệ mã nguồn mở (Apache Doris, Superset) đã giải quyết hiệu quả bài

toán phân mảnh dữ liệu, mang lại công cụ hỗ trợ ra quyết định mạnh mẽ cho lãnh đạo nhà trường.

Điểm đóng góp nổi bật của nghiên cứu là việc đề xuất kiến trúc Lakehouse chi phí thấp phù hợp với bối cảnh các trường đại học Việt Nam. Khác với các mô hình Kho dữ liệu (DWH) truyền thống thường đòi hỏi chi phí bản quyền cao và cấu trúc cứng nhắc, giải pháp đề xuất sử dụng công nghệ mã nguồn mở (Apache Doris, Superset) giúp tiết kiệm ngân sách đầu tư nhưng vẫn đảm bảo khả năng xử lý linh hoạt cả dữ liệu có cấu trúc và bán cấu trúc. Hướng tiếp cận này mở ra khả năng nhân rộng mô hình cho các cơ sở giáo dục khác đang trong tiến trình chuyển đổi số.

Hệ thống không chỉ giúp chuẩn hóa quy trình quản lý dữ liệu mà còn cung cấp cái nhìn sâu sắc về mọi mặt hoạt động của nhà trường thông qua các bảng điều khiển trực quan. Trong tương lai, nhóm nghiên cứu sẽ tiếp tục mở rộng hệ thống theo hướng tích hợp các thuật toán Trí tuệ nhân tạo (AI) và Học máy (Machine Learning) để xây dựng các mô hình dự báo (như dự báo tỷ lệ thôi học, dự báo xu hướng tuyển sinh), từ đó nâng cao hơn nữa năng lực cạnh tranh và vị thế của nhà trường trong kỷ nguyên số.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Trường Đại học Thủ Dầu Một trong đề tài mã số DT.24.2-185.

TÀI LIỆU THAM KHẢO

- [1] Huisman, J., Tight, A., & Vukasovic, L. (2024). Turning universities into data-driven organisations: Seven dimensions of datafication challenges and strategies. *Higher Education*, 87, 1205–1228. <https://doi.org/10.1007/s10734-024-01277-z>
- [2] Inmon, W. H. (2005). *Building the data warehouse* (4th ed.). Wiley.
- [3] Ismaili, B., & Besimi, A. (2024). A data warehousing framework for predictive analytics in higher education: A focus on student at-risk identification. *SEEU Review*, 19(2), 43–57.
- [4] Rongnaimuang, J., Janruang, J., & Karnka, S. (2024). Development of data warehouse using data integration techniques to reduce data redundancy for workload evaluation system. *Journal of Engineering and Digital Technology*, 12(2).
- [5] Rossi, A., Vassiliadis, M., & Sellis, T. (2025). Data integration and storage strategies in heterogeneous analytical architectures. *Information*, 16(11), 932. <https://doi.org/10.3390/info16110932>
- [6] Sharma, A. K., & Singh, R. (2023). Implementation of data warehouse technology in academic data management. *International Journal of Innovative Science and Research Technology*, 8(6).
- [7] Tamm, H. C., & Nikiforova, A. (2024). Towards augmented data quality management: Automation of data quality rule definition in data warehouses. *arXiv preprint arXiv:2406.10940*.