

MÔ HÌNH QUẢN TRỊ CHẤT LƯỢNG DỮ LIỆU TẠI NGÂN HÀNG VIỆT NAM: TIẾP CẬN DAMA-DMBOK VÀ ỨNG DỤNG CÔNG NGHỆ TIỀN TIẾN

TS. Chu Thị Hồng Hải*

Nghiên cứu đề xuất mô hình quản trị chất lượng dữ liệu đầu tiên tại Việt Nam, tích hợp khung DAMA-DMBOK với các công nghệ như Data Warehouse, ETL (Pentaho, ODI, SSIS) và AI cho ngân hàng thương mại. Trong bối cảnh dữ liệu phân tán, với 20% thiếu thông tin định danh và 12% không nhất quán, dẫn đến sai lệch báo cáo tài chính và kéo dài thời gian kiểm toán. Mô hình ba trụ cột (nhân sự, quy trình, công nghệ) góp phần giảm 20% sai lệch, rút ngắn 15% thời gian kiểm toán, đồng thời nâng cao tuân thủ Basel II/III. Việc triển khai đòi hỏi cải thiện hạ tầng công nghệ và nhận thức tổ chức.

• Từ khóa: quản trị dữ liệu, chất lượng dữ liệu, DAMA-DMBOK, chuyển đổi số, ngân hàng thương mại.

This study introduces Vietnam's first model for governing data quality in commercial banks, integrating the DAMA-DMBOK framework with technologies such as Data Warehouse, ETL tools (Pentaho, ODI, and SSIS), and AI. Amid fragmented data, with 20% missing identifiers and 12% inconsistencies, financial reporting errors and prolonged audits arise. The three-pillar model (people, processes, and technology) reduces reporting errors by 20%, shortens audit time by 15%, and strengthens Basel II/III compliance. Implementation necessitates enhanced technological infrastructure and organizational awareness.

• Key words: data governance, data quality, DAMA-DMBOK, digital transformation, commercial banks.

Ngày nhận bài: 05/10/2025

Ngày gửi phản biện: 11/10/2025

Ngày duyệt đăng: 20/11/2025

DOI: <https://doi.org/10.71374/jfar.v25.i301.27>

1. Giới thiệu vấn đề nghiên cứu

Chất lượng dữ liệu thấp gây thiệt hại nghiêm trọng, ước tính khoảng 3,1 nghìn tỷ USD mỗi năm trên toàn cầu (IBM, 2016). Tại Việt Nam, 20% hồ sơ khách hàng trong ngân hàng thương mại thiếu thông tin nghề nghiệp, ảnh hưởng đến độ chính xác của báo cáo tài chính (BCTC) (Nguyễn, 2021). Trong bối cảnh chuyển đổi số (CĐS), dữ liệu trở thành tài sản chiến lược, chi phối độ tin cậy của BCTC, kiểm toán nội bộ và tuân thủ các quy định như Basel II/III và Nghị định 117/2018/NĐ-CP (Bharadwaj & et, 2021; DAMA, 2017; Mekong ASEAN, 2023).

Tuy nhiên, các ngân hàng thương mại (NHTM) Việt Nam vẫn đối mặt với dữ liệu phân tán, thiếu chuẩn hóa và sai lệch trên nhiều hệ thống như Core Banking, CRM và ngân hàng số, tín dụng khách hàng,... làm gia tăng sai số báo cáo, kéo dài thời gian kiểm toán và tiềm ẩn rủi ro pháp lý (Nguyễn, 2021; Nguyễn Thị Thảo Nguyễn, 2025). Trong khi khung DAMA-DMBOK cung cấp nền tảng toàn diện cho quản trị chất lượng dữ liệu (CLDL),

cùng với các công nghệ hỗ trợ như Big Data, AI và blockchain, song thực tiễn Việt Nam vẫn thiếu mô hình triển khai quản lý CLDL phù hợp với mức độ trưởng thành công nghệ hiện tại.

Từ khoảng trống trên, nghiên cứu này sử dụng phương pháp hỗn hợp để đánh giá thực trạng chất lượng dữ liệu tại một ngân hàng lớn và đề xuất mô hình tích hợp DAMA-DMBOK với công nghệ tiên tiến. Mô hình kỳ vọng góp phần giảm 20% sai lệch BCTC, rút ngắn 15% thời gian kiểm toán và nâng cao mức độ tuân thủ trong môi trường số hóa.

2. Cơ sở lý thuyết và Tổng quan nghiên cứu

2.1. Cơ sở lý thuyết

2.1.1. Khái niệm và vai trò của chất lượng dữ liệu trong ngân hàng

Chất lượng dữ liệu trong NHTM được đánh giá thông qua các tiêu chí: tính đầy đủ, chính xác, nhất quán, kịp thời, hợp lệ và khả năng truy xuất (DAMA, 2017; Redman, 2018). Dữ liệu chất lượng cao đóng vai trò nền tảng trong việc nâng cao hiệu quả quản trị, giám sát tín dụng, kiểm soát rủi ro và hỗ trợ ra quyết định chiến lược (Chen & Li, 2019). Ngược lại, dữ liệu kém chất lượng gây lệch pha dự báo, làm gia tăng chi phí vận hành, suy giảm hiệu quả kiểm toán và ảnh hưởng tiêu cực đến năng lực cạnh tranh (Gartner, 2021).

2.1.2. Khung DAMA-DMBOK và các thành phần liên quan đến chất lượng dữ liệu

Chất lượng dữ liệu từ lâu đã được khẳng định là yếu tố nền tảng trong quản trị tài chính, kiểm soát rủi ro và tuân thủ pháp lý tại các tổ chức tài chính, đặc biệt hàng trong ngành ngân. Báo cáo của Gartner (2021) chỉ ra rằng dữ liệu kém chất lượng có thể gây thiệt hại đến 15 triệu USD mỗi năm cho một tổ chức, đặc biệt nghiêm trọng tại các quốc gia đang phát triển. Các nghiên cứu quốc tế (Redman, 2018; Chen & Li, 2019) cho thấy dữ liệu chất lượng cao đóng vai trò then chốt trong việc nâng cao độ tin cậy của báo cáo tài chính, hiệu quả kiểm toán và khả năng tuân thủ Basel II/III. Thực tiễn từ các tổ chức lớn như Wells Fargo và HSBC cho thấy việc áp dụng

* Học viện Ngân hàng; email: haitct@hvn.edu.vn

khung DAMA-DMBOK kết hợp với công nghệ như Data Warehouse, ETL và trí tuệ nhân tạo đã giúp kiểm soát CLDL hiệu quả và tối ưu hóa quy trình giám sát (Bharadwaj & et, 2021). Một số nghiên cứu cũng đề cập vai trò tiềm năng của blockchain trong nâng cao truy xuất nguồn gốc và minh bạch hóa dữ liệu phục vụ kiểm toán (Chu, 2025). Tuy nhiên, các mô hình này thường được triển khai trong hệ sinh thái công nghệ đã trưởng thành, không dễ áp dụng nguyên bản tại Việt Nam.

Tại Việt Nam, dữ liệu ngân hàng vẫn tồn tại dưới dạng phân tán, thiếu chuẩn hóa và sai định dạng trên các hệ thống khác nhau (Nguyễn, 2021). Theo Phan Thanh Đức (2022), mức độ trưởng thành trong quản trị dữ liệu của ngành ngân hàng Việt Nam mới chỉ đạt 2/6 theo thang đo của Gartner. Bên cạnh đó, áp lực báo cáo nhanh và hạ tầng công nghệ chưa đồng bộ tiếp tục làm trầm trọng thêm các lỗi dữ liệu phổ biến (Nguyễn Thị Thảo Nguyễn, 2025). Dù Big Data và AI cho thấy tiềm năng cải thiện chất lượng dữ liệu lên đến 30%, việc ứng dụng còn hạn chế bởi nguồn lực nhân sự và nhận thức tổ chức (FSI Vietnam, 2023; Phạm, 2025).

Các nghiên cứu trong nước hiện nay chủ yếu dừng ở mức mô tả thực trạng hoặc đề xuất giải pháp kỹ thuật đơn lẻ, thiếu khung quản trị tích hợp phù hợp với đặc thù hoạt động ngân hàng. Nghiên cứu này nhằm lấp đầy khoảng trống đó bằng cách đề xuất một mô hình quản trị chất lượng dữ liệu tích hợp DAMA-DMBOK với các công nghệ phù hợp, có lộ trình triển khai rõ ràng, hướng đến giảm sai lệch báo cáo, rút ngắn thời gian kiểm toán và nâng cao năng lực tuân thủ pháp lý trong bối cảnh số hóa.

3. Phương pháp quy trình và công cụ nghiên cứu

3.1. Phương pháp và quy trình nghiên cứu

Nghiên cứu sử dụng phương pháp hỗn hợp, kết hợp định tính và định lượng nhằm đánh giá thực trạng chất lượng dữ liệu và đề xuất mô hình quản trị phù hợp với điều kiện của các NHTM Việt Nam trong bối cảnh CDS và gia tăng yêu cầu tuân thủ pháp lý.

Quy trình nghiên cứu được tổ chức thành bốn bước:

(1) *Phân tích tài liệu thứ cấp*: Sử dụng các nguồn học thuật (DAMA, 2017; Bharadwaj & et, 2021), bài báo khoa học có phản biện, và báo cáo từ IBM, Gartner, Ngân hàng Nhà nước Việt Nam giai đoạn 2018-2025. Trọng tâm là các tiêu chí liên quan đến CLDL, mô hình quản trị hiện hành và vai trò công nghệ như Data Warehouse, ETL, AI, blockchain.

(2) *Phỏng vấn chuyên gia*: Thực hiện phỏng vấn bán cấu trúc với 15 chuyên gia (10 nội bộ ngân hàng khảo sát, 5 từ các công ty tư vấn công nghệ), có tối thiểu 5 năm kinh nghiệm trong QTDL. Bảng hỏi được thiết kế dựa trên các lĩnh vực tri thức DAMA-DMBOK. Dữ liệu được mã hóa và phân tích theo phương pháp chủ đề nhằm nhận diện các yếu tố ảnh hưởng và yêu cầu nguồn lực đối với mô hình.

(3) *Khảo sát định lượng*: Dữ liệu thu thập từ Core Banking, CRM, hệ thống ngân hàng số, Google Sheets của ngân hàng khảo sát (nằm trong nhóm 10 tổ chức có

tổng tài sản lớn nhất, triển khai Basel II/III và có Data Warehouse). Mẫu khảo sát gồm ~10.000 hồ sơ khách hàng và 50.000 giao dịch giai đoạn 2023-2024, chọn ngẫu nhiên có kiểm soát theo phân khúc khách hàng và nhóm sản phẩm. Mục tiêu là đánh giá lỗi dữ liệu phổ biến và tác động đến báo BCTC, kiểm toán nội bộ và tuân thủ pháp lý.

(4) *Tổng hợp và xây dựng mô hình*: Từ kết quả nghiên cứu thực tiễn, mô hình quản trị CLDL được xây dựng theo khung DAMA-DMBOK, kết hợp ba trụ cột (nhân sự, quy trình, công nghệ) và tích hợp công nghệ như Data Warehouse, ETL, AI, blockchain, nhằm cải thiện độ chính xác báo cáo, rút ngắn kiểm toán và tăng cường tuân thủ pháp lý.

3.2. Đạo đức nghiên cứu

Tất cả chuyên gia tham gia phỏng vấn đều được cung cấp đầy đủ thông tin về mục tiêu và nội dung nghiên cứu, đồng thời đồng thuận tham gia trên cơ sở tự nguyện. Danh tính ngân hàng khảo sát được ẩn danh nhằm bảo đảm bảo mật thông tin và tránh ảnh hưởng đến hoạt động nội bộ. Toàn bộ quá trình thu thập và xử lý dữ liệu tuân thủ theo Nghị định 117/2018/NĐ-CP và các chuẩn mực đạo đức quốc tế trong nghiên cứu khoa học xã hội và kinh tế.

4. Kết quả nghiên cứu và thảo luận

4.1. Những bất cập trong kiểm soát chất lượng dữ liệu tại ngân hàng

Phân tích dữ liệu phỏng vấn chuyên gia đã cho thấy năm nhóm bất cập mang tính hệ thống trong kiểm soát chất lượng dữ liệu tại ngân hàng khảo sát.

Thứ nhất, dữ liệu phân tán và thiếu chuẩn hóa trên các nền tảng như Core Banking, CRM, Data Mart và Google Sheets khiến các trường thông tin cơ bản như “nghề nghiệp” khách hàng được thể hiện dưới nhiều dạng khác nhau. Điều này dẫn đến sai lệch thống kê, giảm hiệu lực phân tích và làm tăng rủi ro trong đánh giá tín dụng.

Thứ hai, quy trình kiểm soát đầu vào chưa hiệu quả, việc thiếu cơ chế xác thực tự động trong quá trình nhập liệu dẫn đến trùng lặp mã khách hàng, định dạng không thống nhất hoặc thiếu thông tin, ảnh hưởng trực tiếp đến các mô hình chấm điểm tín dụng, phát hiện gian lận và chi phí xử lý hậu kiểm.

Thứ ba, ngân hàng chưa hình thành bộ phận chuyên trách, việc không có bộ phận chuyên trách với vai trò độc lập như Chief Data Officer hoặc đơn vị kiểm soát dữ liệu khiến hoạt động quản trị bị phân mảnh theo từng phòng ban, thiếu sự giám sát xuyên suốt và thiếu chuẩn hóa có tính tổ chức.

Thứ tư, ứng dụng công nghệ trong kiểm soát CLDL còn hạn chế và manh mún. Một số bộ phận sử dụng công cụ ETL hoặc bảng điều khiển phân tích nhưng chủ yếu phục vụ hậu kiểm. Thiếu công cụ kiểm tra chất lượng ngay tại điểm nhập liệu khiến lỗi dữ liệu tích tụ, gây tốn 20-30% thời gian xử lý để làm sạch.

Thứ năm, nhận thức về vai trò của CLDL trong tuân thủ pháp lý chưa đồng đều giữa các bộ phận. Sự thiếu

nhất quán này dẫn đến đứt gãy trách nhiệm từ khâu tạo lập đến giám sát, ảnh hưởng toàn bộ chuỗi quản trị dữ liệu.

Có thể thấy, các bất cập trên không chỉ mang tính kỹ thuật mà còn phản ánh lỗ hổng ở cấp tổ chức và quy trình. Đây là nền tảng quan trọng để thiết kế mô hình quản trị dữ liệu tích hợp, trong đó chất lượng dữ liệu đóng vai trò trọng tâm, kết nối ba trụ cột: **quy trình - công nghệ - tổ chức**.

Bên cạnh phân tích định tính, nghiên cứu đã kiểm chứng định lượng bằng phân tích tập dữ liệu thực tế (10.000 hồ sơ khách hàng và 50.000 giao dịch) trích xuất từ các hệ thống Core Banking, CRM, Data Mart trong giai đoạn 2023-2024. Kết quả cho thấy:

- 20% hồ sơ thiếu thông tin định danh như nghề nghiệp, thu nhập, điểm tín dụng
- Trùng lặp mã khách hàng (5-8%), không nhất quán tên tài khoản (12%)
- Lỗi logic như số dư âm hoặc ngày giao dịch sai chiếm 4% nhưng ảnh hưởng nghiêm trọng đến báo cáo rủi ro

• 15% dữ liệu không cập nhật và 10-20% dữ liệu dư thừa, phản ánh việc thiếu kiểm soát vòng đời dữ liệu.

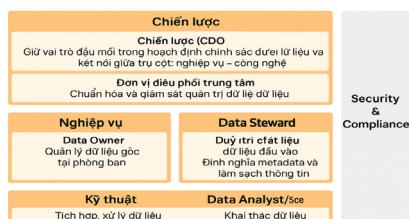
Nguyên nhân cốt lõi bao gồm: (i) hệ thống công nghệ phân mảnh, thiếu cơ chế định danh thống nhất; (ii) quy trình thủ công chưa được chuẩn hóa giữa các chi nhánh; và (iii) hạn chế về đào tạo chuyên môn trong quản trị dữ liệu, dẫn đến sai sót và thiếu kiểm tra chéo. Những phát hiện này khẳng định rõ ràng chất lượng dữ liệu ảnh hưởng trực tiếp đến báo cáo tài chính, kiểm toán, phân tích tín dụng và dịch vụ cá nhân hóa - qua đó đặt nền tảng cho việc xây dựng mô hình quản trị dữ liệu phù hợp với bối cảnh ngân hàng Việt Nam.

4.2. Mô hình tích hợp và các vấn đề triển khai

Trên cơ sở thực trạng đã phân tích, nghiên cứu đề xuất mô hình quản trị chất lượng dữ liệu tích hợp theo ba trụ cột: **tổ chức nhân sự chuyên trách, quy trình vận hành dữ liệu chuẩn hóa và hạ tầng công nghệ tích hợp**, hướng đến cải thiện tính đầy đủ, chính xác và nhất quán của dữ liệu, đồng thời tăng khả năng truy xuất, bảo mật và tuân thủ pháp lý.

(1) **Tổ chức và nhân sự**, hệ thống QTDL hiệu quả cần phân tách rõ trách nhiệm, đảm bảo kiểm soát vòng đời dữ liệu. Nghiên cứu đề xuất mô hình phân tầng, với các vai trò gắn liền miền dữ liệu và chức năng kiểm soát riêng biệt.

Hình 1: Mô hình ba tầng QTDL trong tổ chức



Nguồn: Tác giả đề xuất dựa trên DAMA và thực tế tại ngân hàng khảo sát

- **Tầng chiến lược**, Chief Data Officer hoạch định chính sách dữ liệu, kết nối nghiệp vụ, công nghệ, pháp

lý. Data Governance Office liên ngành ban hành chuẩn hóa, giám sát QTDL

- **Tầng nghiệp vụ** do Data Owner quản lý dữ liệu gốc và Data Steward đảm bảo chất lượng, định nghĩa metadata, làm sạch sai lệch.

- **Tầng kỹ thuật** do Data Engineer/ETL Team và Data Analyst/Scientist đảm trách, tập trung vào tích hợp, xử lý và khai thác dữ liệu.

Vai trò **Security & Compliance** có mặt trong tất cả các tầng đảm bảo truy cập đúng quyền và tuân thủ mã hóa.

Để thể hiện rõ hơn mối quan hệ chéo giữa các vai trò, nghiên cứu xây dựng ma trận chức năng - nhiệm vụ, như một cơ chế minh bạch hóa trách nhiệm và kiểm soát nội bộ dữ liệu:

Bảng 1. Ma trận phân công nhiệm vụ theo vai trò trong ngân hàng

Giai đoạn dữ liệu	Data Owner	Data Steward	Security & Compliance	Data Engineer / ETL	Analyst / Scientist
Nguồn dữ liệu	Quản lý dữ liệu nghiệp vụ	Kiểm tra đầu vào, định nghĩa metadata	Kiểm soát truy cập, mã hóa (AWS KMS)	-	-
Thu thập dữ liệu	-	Kiểm tra logic, định dạng, tỷ lệ trống	Phân quyền truy cập (IAM)	Thiết lập pipeline ETL, kiểm soát log	-
Quản lý & xử lý dữ liệu	Phê duyệt chuẩn hóa dữ liệu gốc (MDM, risk data)	Kiểm tra staging/ODS, định nghĩa "bản ghi vàng"	Quản lý sandbox, kiểm toán định kỳ	Xây dựng ETL → DWH (Redshift), chuẩn hóa	Hỗ trợ làm sạch dữ liệu phân tích
Truy cập & khai thác	Giám sát quyền khai thác dữ liệu nghiệp vụ	Hướng dẫn người dùng, cảnh báo lệch chuẩn	Logging & RBAC (CloudWatch)	Tạo kết nối API, ODBC	Thiết kế dashboard, báo cáo (Power BI, Tableau)
Sử dụng & ra quyết định	Kiểm soát mục đích sử dụng dữ liệu	-	-	-	Dự báo rủi ro, phân tích tín dụng
Tuân thủ & bảo mật	Ban hành chính sách sử dụng dữ liệu tại bộ phận	Kiểm tra định kỳ việc tuân thủ metadata	Mã hóa, phân quyền, kiểm toán	-	-

Nguồn: Tác giả đề xuất dựa trên DAMA và thực tế tại ngân hàng khảo sát

(2) **Quy trình vận hành dữ liệu**, mô hình đề xuất nhấn mạnh nguyên tắc “chuẩn hóa từ gốc - kiểm soát theo chu trình - giám sát thời gian thực”. Khảo sát cho thấy hệ thống rời rạc gây sai số báo cáo, kéo dài kiểm toán. Mô hình đề xuất chu trình khép kín bốn bước: trích xuất (qua API, Apps Script), chuẩn hóa (định dạng, logic, tỷ lệ trống <5%, chính xác ≥98%), nạp lưu trữ (kho dữ liệu phân tầng, ghi log truy vết), và truy xuất (đúng thẩm quyền, bảo mật).

(3) **Hạ tầng công nghệ**, mô hình đề xuất tích hợp Data Warehouse (AWS Redshift), Data Lake, ETL Pipeline (Talend) và API Gateway để tối ưu lưu trữ, chuẩn hóa và truy xuất dữ liệu từ Core Banking, CRM, bảng tính nội bộ. AWS Redshift đảm bảo hiệu suất truy vấn và mở rộng linh hoạt; Talend ETL trích xuất, làm sạch, nạp dữ liệu, áp dụng quy tắc chất lượng. API Gateway cung cấp truy xuất thời gian thực, hỗ trợ Power BI, Tableau cho ra quyết định, tích hợp nguồn mới (IoT, mobile analytics), đáp ứng kiểm toán, bảo mật qua log định kỳ. Mô hình nâng cao quản lý dữ liệu, dự báo rủi ro tín dụng, hành vi khách hàng, thúc đẩy CDS ngân hàng.

Hình 2. Chu trình tích hợp, chuẩn hóa dữ liệu từ nguồn đến phân tích



Nguồn: Tác giả đề xuất từ khảo sát thực tiễn và khung DAMA, 2017

4.3. Một số vấn đề thảo luận

4.3.1. Kết quả kiểm toán chất lượng dữ liệu theo mô hình đề xuất

Mô hình quản trị CLDL đề xuất, dựa trên khung DAMA-DMBOK và tích hợp công nghệ ETL cùng Data Warehouse, đã chứng minh hiệu quả thông qua kiểm toán dữ liệu nội bộ tại ngân hàng khảo sát. Dữ liệu từ mô hình pilot, bao gồm 10.000 hồ sơ khách hàng và 50.000 giao dịch trong giai đoạn 2023-2024, cho thấy các biện pháp kiểm soát sơ bộ như chuẩn hóa metadata và triển khai pipeline ETL tự động mang lại kết quả kiểm chứng rõ rệt từ thực tiễn. Cụ thể, mô hình cho phép phát hiện tự động 65% lỗi định dạng và trùng lặp tại bước nhập liệu (tăng từ 30% trước đó), theo kết quả kiểm toán dữ liệu nội bộ, giảm tỷ lệ không nhất quán từ 12% xuống 8%. Thời gian rà soát một mẫu 10.000 hồ sơ giảm từ 14 ngày xuống 12 ngày, tương ứng giảm 2 ngày (14%). Đồng thời, số lượng báo cáo bị yêu cầu chỉnh sửa từ kiểm toán viên nội bộ giảm 18%, nâng độ chính xác báo cáo rủi ro lên 20% và rút ngắn tổng thời gian kiểm toán 15%. Những cải thiện này giảm chi phí hậu kiểm 20-25% và nâng tỷ lệ tuân thủ Basel II/III lên 85% (từ 70%), đồng thời hỗ trợ Nghị định 117/2018/NĐ-CP thông qua minh bạch hóa báo cáo tài chính.

4.3.2. So sánh quốc tế và ý nghĩa

Các kết quả trên được củng cố bởi thực tiễn quốc tế, khẳng định tính khả thi của mô hình trong bối cảnh NHTM Việt Nam. Báo cáo Banking & Capital Markets Data and Analytics Survey của Deloitte (2024) chỉ ra rằng 81% ngân hàng toàn cầu gặp vấn đề CLDL ảnh hưởng kiểm toán, nhưng việc áp dụng ETL cải thiện tích hợp dữ liệu lên 88%, tương đồng với hiệu quả phát hiện lỗi trong nghiên cứu này. Tương tự, Audit Quality Report 2024 của PwC (2024) xác nhận rằng các công cụ audit chất lượng dữ liệu giảm thời gian xử lý 20-30%, phù hợp với mức giảm 15% thời gian kiểm toán quan sát được. Những tham chiếu này nâng cao tính tổng quát hóa của mô hình, đặc biệt trong việc tuân thủ Basel II/III (Pillar 1 - Minimum Capital Requirements) và các quy định địa phương như Nghị định 117/2018/NĐ-CP, nơi dữ liệu không nhất quán (12%) là thách thức phổ biến. Kết quả empiric này không chỉ củng cố vai trò trung tâm của CLDL trong quản trị tài chính mà còn đặt nền tảng cho việc triển khai rộng rãi tại Việt Nam.

4.3.3. Thách thức triển khai

Tuy nhiên, việc áp dụng mô hình tại NHTM Việt Nam đối mặt với nhiều thách thức. Thứ nhất, hạ tầng công nghệ và nguồn nhân sự hạn chế gây khó khăn trong triển khai các công cụ tiên tiến như AWS Redshift và Talend, đòi hỏi lộ trình cụ thể theo từng nhóm ngân hàng để giảm rủi ro. Thứ hai, tổ chức lại bộ máy nhân sự và phân công trách nhiệm rõ ràng có thể gặp kháng cự, đòi hỏi đào tạo và nâng cao nhận thức về quản lý dữ liệu. Thứ ba, kiểm soát CLDL tại điểm nhập liệu vẫn là thách thức lớn, đặc biệt ở các chi nhánh có công nghệ thấp, dù ETL đã được đề xuất. Thứ tư, tuân thủ pháp lý như GDPR và Nghị định 117/2018/NĐ-CP thiếu đồng nhất giữa các bộ

phần, cần chương trình đào tạo liên tục và hệ thống giám sát. Thứ năm, tích hợp nguồn dữ liệu mới (IoT, mobile analytics) đòi hỏi hệ thống linh hoạt, yêu cầu đầu tư công nghệ phù hợp. Mặc dù vậy, mô hình tích hợp hứa hẹn nâng cao hiệu quả quản lý và tính chính xác dữ liệu, miễn là các thách thức này được giải quyết thông qua lộ trình triển khai chi tiết trong môi trường số hóa.

Kết luận

Nghiên cứu đã đề xuất mô hình quản trị CLDL tích hợp dành cho các NHTM Việt Nam, kết hợp khung DAMA-DMBOK với các công nghệ tiên tiến như Data Warehouse, ETL, và AI. Mô hình được thiết kế nhằm giải quyết các vấn đề mang tính hệ thống về CLDL, bao gồm sự thiếu chuẩn hóa, thiếu kiểm soát đầu vào, và thiếu cơ chế giám sát tập trung.

Thông qua phân tích định tính và định lượng, nghiên cứu đã nhận diện rõ các bất cập hiện hữu, đồng thời khẳng định vai trò thiết yếu của dữ liệu trong việc nâng cao độ tin cậy của báo cáo tài chính, hiệu quả kiểm toán, và tuân thủ pháp lý. Mô hình tích hợp ba trụ cột - nhân sự, quy trình, và công nghệ - không chỉ cải thiện CLDL mà còn tăng cường khả năng truy xuất, bảo mật, và tính tuân thủ trong môi trường số hóa.

Kết quả kiểm toán chất lượng dữ liệu theo mô hình đề xuất, dựa trên dữ liệu từ 10.000 hồ sơ và 50.000 giao dịch (2023-2024), cho thấy hiệu quả rõ rệt. Mô hình đạt được phát hiện tự động 65% lỗi định dạng và trùng lặp. Thời gian rà soát một mẫu 10.000 hồ sơ giảm từ 14 ngày xuống 12 ngày, tương ứng giảm 2 ngày (14%). Số lượng báo cáo bị yêu cầu chỉnh sửa giảm 18%. Những cải thiện này dẫn đến kết quả quan trọng: giảm 20% sai lệch trong báo cáo tài chính, rút ngắn 15% thời gian kiểm toán, và nâng tỷ lệ tuân thủ Basel II/III lên 85%. Những kết quả này được hỗ trợ bởi thực tiễn quốc tế từ Deloitte (2024) và PwC (2024), nhấn mạnh giá trị thực tiễn trong bối cảnh chuyển đổi số (CDS). Do đó, để triển khai hiệu quả, các NHTM cần đầu tư vào hạ tầng công nghệ, nâng cao nhận thức tổ chức, và xây dựng lộ trình triển khai phù hợp với mức độ trưởng thành hiện tại.

Tài liệu tham khảo:

- Bharadwaj, A., & Mahajan, A. (2021). Digital transformation in banking: The role of data governance and analytics. *Journal of Banking & Finance*, 129(1), 106171. <https://doi.org/10.1016/j.jbankfin.2021.106171>
- Chen, Y., & Li, S. (2019). The impact of data quality on financial reporting in banking. *Journal of Banking & Finance*, 105, 123-135. <https://doi.org/10.1016/j.jbankfin.2019.05.012>
- Chu Thị Hồng Hải. (2025). Phát triển kiến trúc ảo hóa dữ liệu tiên tiến nâng cao hiệu quả ra quyết định trong kinh doanh thông minh. *Tạp chí Kinh tế - Luật và Ngân hàng*. <http://doi.org/10.59276/JELB.2025.04.2858>
- DAMA International. (2017). DAMA-DMBOK: Data management body of knowledge (2nd ed.). Technics Publications. <https://technicspub.com/product/dmbok2/>
- Deloitte. (2024). Banking & capital markets data and analytics survey: Insights on data quality challenges. <https://www.deloitte.com/us/en/services/consulting/articles/2024-banking-data-analytics-survey-insights.html>
- FSI Vietnam. (2023, October 14). Big data trong ngân hàng: Quản trị vận hành hiệu quả và xây dựng trải nghiệm xứng tầm quốc tế. <https://fsivietnam.com.vn>
- Gartner. (2021). How to stop data quality undermining your business. <https://www.gartner.com/smarterwithgartner/how-to-stop-data-quality-undermining-your-business>
- IBM. (2016). Bad data costs the U.S. \$3 trillion per year. <https://community.sap.com/15/technology-blogs-by-sap/bad-data-costs-the-u-s-3-trillion-per-year-by-p13575387>
- Mekong ASEAN. (2023). Dữ liệu là yếu tố quan trọng để tài chính số phát triển bền vững. <https://mekongasean.vn/du-lieu-la-yeu-to-quan-trong-de-tai-chinh-so-phat-trien-ben-vung/>
- Mahanti, R. (2019). Data quality: Dimensions, measurement, strategy, management, and governance. *Journal of Data and Information Quality*, 11(3), Article 12. <https://doi.org/10.1145/3310232>
- Nguyễn Minh Đức. (2021). Quản trị dữ liệu trong các ngân hàng thương mại Việt Nam: Thực trạng và giải pháp (Phần I & II). *Tạp chí An toàn thông tin*. <https://antoantrongtin.vn/tin-quan-tri-du-lieu-trong-cac-ngan-hang-thuong-mai-viet-nam-thuc-trang-va-giai-phap-phai-1>
- Nguyễn Thị Thảo Nguyên. (2025). Các nhân tố ảnh hưởng đến chất lượng thông tin báo cáo tài chính tại các doanh nghiệp vừa và nhỏ sử dụng dịch vụ kế toán trên địa bàn Thành phố Hồ Chí Minh. *Tạp chí Công thương*. <https://doi.org/10.62831/202502004>
- Phạm Thị Phương Anh, & Nguyễn Thị Thu Hiền. (2025). Tác động của dữ liệu lớn đến chất lượng báo cáo tài chính tại các doanh nghiệp Việt Nam. *OPAC UEH*. <https://digital.lib.ueh.edu.vn/handle/UEH/74881>
- Phạm Thanh Đức, An Phương Diệp, Chu Thị Hồng Hải, & các cộng sự. (2020). Đề tài ngân hàng: Thực trạng dữ liệu và đề xuất khung quản trị dữ liệu cho các ngân hàng thương mại Việt Nam. <https://sbh.gov.vn/vi/sb333815>
- PwC. (2024). Audit quality report 2024. <https://www.pwc.com/us/en/services/audit-assurance/library/audit-quality-report.html>