

# ỨNG DỤNG DATA LAKEHOUSE TRONG QUẢN TRỊ KHÁCH HÀNG TẠI NGÂN HÀNG

ThS. Ngô Thùy Linh

Học viện Ngân hàng

Tác giả liên hệ: linhnt@hvn.edu.vn

Ngày nhận: 13/6/2025

Ngày nhận bản sửa: 15/6/2025

Ngày duyệt đăng: 24/6/2025

DOI: <https://doi.org/10.71192/697251hbekkk>

## Tóm tắt

Trong bối cảnh ngành Ngân hàng đang chuyển mình mạnh mẽ theo xu hướng số hóa, nhu cầu xây dựng mối quan hệ khách hàng toàn diện, linh hoạt và cá nhân hóa ngày càng trở nên cấp thiết. Tuy nhiên, các hệ thống quản trị quan hệ khách hàng truyền thống đang gặp nhiều hạn chế về khả năng tích hợp dữ liệu, phân tích thời gian thực, và đáp ứng các yêu cầu quản trị dữ liệu hiện đại. Bài viết này đề xuất ứng dụng kiến trúc Data Lakehouse như một giải pháp công nghệ khả thi, kết hợp ưu điểm của Hồ dữ liệu (Data Lake) và Kho dữ liệu (Data Warehouse) nhằm hỗ trợ lưu trữ linh hoạt, truy vấn hiệu quả và tích hợp đa dạng dữ liệu từ các nguồn khác nhau trong hệ thống ngân hàng. Bài viết trình bày kiến trúc Lakehouse, phân tích các tầng chức năng trong mô hình tích hợp với việc quản trị quan hệ khách hàng. Qua đó, nghiên cứu khẳng định rằng Data Lakehouse không chỉ góp phần nâng cao năng lực phân tích và ra quyết định, mà còn là nền tảng chiến lược giúp các ngân hàng thích ứng với biến động thị trường và đáp ứng yêu cầu tuân thủ pháp lý ngày càng nghiêm ngặt.

**Từ khóa:** Data Lakehouse, kho dữ liệu, hồ dữ liệu, quản trị quan hệ khách hàng, ngân hàng.

## Utilizing Data Lakehouse technology in Banking Customer Relationship Management

Ngo Thuy Linh, MA.

Banking Academy of Viet Nam

Corresponding Author: linhnt@hvn.edu.vn

## Abstract

Amidst the rapid digital evolution reshaping the banking landscape, the demand for a comprehensive, flexible, and personalized customer relationship has become increasingly imperative. Yet, conventional Customer Relationship Management (CRM) systems encounter notable constraints in data amalgamation, real-time analytics, and adherence to contemporary data governance standards. This manuscript advocates for the implementation of the Data Lakehouse architecture as a viable technological solution that amalgamates the robust features of Data Lakes and Data Warehouses to facilitate scalable storage, streamlined querying, and seamless fusion of varied data origins within banking infrastructures. The study presents the fundamental structure of the Lakehouse architecture and analyzes its functional layers in relation to CRM integration. The findings suggest that Data Lakehouse not only enhances analytical and decision-making capabilities but also serves as a strategic foundation for banks to adapt to market volatility and comply with increasingly stringent regulatory requirements.

**Keywords:** Data Lakehouse, Data Warehouse, Data Lake, Customer Relationship Management (CRM), Banking.

## 1. Giới thiệu

Trong bối cảnh ngành Ngân hàng đang bước vào thời kỳ chuyển đổi số toàn diện, việc xây dựng và duy trì mối quan hệ bền vững với khách hàng đã trở thành một yếu tố chiến lược sống còn. Với sự cạnh tranh ngày càng khốc liệt và kỳ vọng ngày một cao từ phía khách hàng, các ngân hàng không thể chỉ dừng lại ở việc cung cấp sản phẩm tài

chính, mà còn phải cá nhân hóa trải nghiệm dịch vụ trên nhiều kênh giao tiếp (Maximizer, 2025). Hệ thống quản trị quan hệ khách hàng (CRM) được xem như một hệ thống trung tâm giúp ngân hàng hiểu rõ hơn về hành vi, nhu cầu và hành trình của từng khách hàng, thông qua việc tích hợp dữ liệu từ các điểm chạm như giao dịch tại quầy, ứng dụng di động, tổng đài và mạng xã hội.

Theo Veritis (2024), CRM còn đóng vai trò quan trọng trong việc chuẩn hóa và tập trung hóa dữ liệu khách hàng, tạo điều kiện thuận lợi cho việc phân tích xu hướng, dự đoán nhu cầu và tối ưu hóa chiến dịch tiếp thị. Bên cạnh đó, việc áp dụng CRM còn giúp nâng cao hiệu suất vận hành nội bộ thông qua tự động hóa các quy trình quản lý, từ theo dõi giao dịch, lịch sử liên lạc đến quản lý phản hồi và khiếu nại. Đặc biệt trong môi trường tuân thủ nghiêm ngặt về bảo mật thông tin và các quy định như GDPR (Quy định chung về bảo vệ dữ liệu của EU) hay CCPA (Đạo luật quyền riêng tư người tiêu dùng của bang California), CRM hiện đại cung cấp các tính năng mã hóa, kiểm soát truy cập và lưu vết audit nhằm đảm bảo an toàn dữ liệu và tuân thủ pháp lý (Maximizer, 2025). Từ đó, có thể thấy, CRM không chỉ đơn thuần là một giải pháp công nghệ hỗ trợ bán hàng, mà còn là nền tảng chiến lược giúp ngân hàng tăng cường năng lực cạnh tranh, nâng cao lòng trung thành và duy trì tính linh hoạt trong môi trường kinh doanh biến động (Veritis, 2024; Maximizer, 2025).

Một trong những thách thức phổ biến và nghiêm trọng nhất trong quản trị khách hàng tại ngân hàng là tình trạng dữ liệu bị phân mảnh và thiếu tích hợp. Dữ liệu khách hàng thường nằm rải rác trong các hệ thống riêng biệt như Core Banking, CRM, giao dịch điện tử, ứng dụng di động, và nền tảng mạng xã hội. Khi các hệ thống này không được đồng bộ, ngân hàng khó có thể xây dựng được một cái nhìn toàn diện (Customer 360) về khách hàng, từ đó, cản trở khả năng cá nhân hóa dịch vụ và tối ưu hóa hành trình trải nghiệm (Drenik, 2025). Trong thời đại số, khi khách hàng kỳ vọng sự liên mạch, thông minh và tự động hóa trong mọi điểm chạm, việc không tích hợp dữ liệu đồng nghĩa với việc đánh mất cơ hội xây dựng mối quan hệ khách hàng bền vững.

Bên cạnh vấn đề tích hợp, chất lượng dữ liệu cũng là một rào cản đáng kể. Dữ liệu khách hàng thường bị lỗi, thiếu, trùng lặp hoặc không cập nhật - một phần do nhập liệu thủ công, phần khác do không có hệ thống kiểm tra và làm sạch dữ liệu tự động. Theo NetSuite (2025), đây là nguyên nhân chính dẫn đến việc các báo cáo phân tích thiếu chính xác, ảnh hưởng trực tiếp đến quyết định chiến lược và chất lượng dịch vụ khách hàng. Việc thiếu chuẩn hóa dữ liệu giữa các phòng ban hoặc chi nhánh cũng tạo ra sự bất nhất trong nhận thức và xử lý thông tin khách hàng.

Dữ liệu khách hàng trong ngân hàng không chỉ đa dạng, mà còn ngày càng tăng

nhANH theo cấp số nhân. Khối lượng dữ liệu phát sinh từ các giao dịch, tương tác kỹ thuật số và cảm xúc trên mạng xã hội đặt ra yêu cầu xử lý trong thời gian thực. Tuy nhiên, nhiều hệ thống truyền thống chưa đủ khả năng mở rộng để đáp ứng hiệu năng cần thiết. SyncMatters (2024) chỉ ra rằng, nếu không có khả năng phản hồi thời gian thực, ngân hàng sẽ bỏ lỡ các tín hiệu quan trọng, đặc biệt trong các hoạt động như đề xuất sản phẩm tức thì hoặc cảnh báo gian lận.

Dữ liệu khách hàng trong ngành Tài chính có tính nhạy cảm cao và bị điều chỉnh bởi các quy định nghiêm ngặt như GDPR, CCPA, hoặc quy định nội bộ của ngân hàng trung ương. Các ngân hàng không chỉ phải lưu trữ và xử lý đúng cách, mà còn cần xây dựng hệ thống mã hóa, kiểm soát truy cập và nhật ký hoạt động đầy đủ. Nếu dữ liệu bị rò rỉ hoặc sử dụng sai mục đích, tổ chức có thể phải chịu hậu quả nghiêm trọng về pháp lý và uy tín. Các chuyên gia từ NetSuite (2025) và SyncMatters (2024) nhấn mạnh rằng, bảo mật dữ liệu không chỉ là yêu cầu công nghệ, mà còn là trách nhiệm quản trị mang tính chiến lược.

Cuối cùng, một thách thức ít được nhắc đến nhưng có ảnh hưởng sâu sắc là khoảng cách giữa hệ thống dữ liệu và quy trình vận hành thực tế. Các ngân hàng có thể đầu tư hàng triệu đô vào các nền tảng dữ liệu, nhưng nếu không đồng bộ với quy trình nghiệp vụ hiện hành, dữ liệu thu thập sẽ không thể phát huy hiệu quả. SyncMatters (2024) cho rằng, cần phải có một chiến lược tích hợp không chỉ về mặt công nghệ, mà cả ở khía cạnh tổ chức - tức là, đào tạo con người, điều chỉnh quy trình và thay đổi tư duy vận hành dựa trên dữ liệu.

Lakehouse là một kiến trúc kết hợp giữa Data Lake và Data Warehouse, với chi phí lưu trữ thấp, linh hoạt xử lý mọi dạng dữ liệu và hỗ trợ ACID, metadata, time-travel, đồng thời, tương thích BI/ML (Striim, 2021; IBM, 2024; Databricks, 2019). Như vậy, bài toán quản trị quan hệ khách hàng trong ngân hàng đang đối mặt với nhiều thách thức về dữ liệu - từ vấn đề tích hợp, chất lượng, bảo mật, đến khả năng phản hồi theo thời gian thực và đồng bộ với quy trình nghiệp vụ. Trong bối cảnh đó, kiến trúc Data Lakehouse nổi lên như một hướng tiếp cận mới, hội tụ được các đặc tính vượt trội của cả Data Lake và Data Warehouse. Với khả năng xử lý dữ liệu đa định dạng, hỗ trợ phân tích chuyên sâu và tích hợp linh hoạt, Lakehouse hứa hẹn sẽ là nền tảng hạ tầng dữ liệu tối ưu cho các hệ thống CRM hiện đại trong ngành Ngân hàng. Trên cơ sở đó, bài viết sẽ đi sâu phân tích vai

trò của Lakehouse trong việc giải quyết các vấn đề dữ liệu của CRM ngân hàng, đồng thời, đề xuất kiến trúc ứng dụng cụ thể và thảo luận các yếu tố then chốt trong triển khai thực tiễn.

## 2. Kiến trúc Data Lakehouse

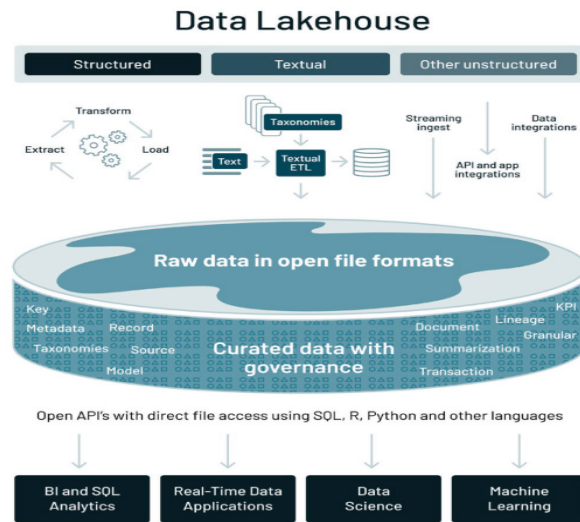
Data Lakehouse là một kiến trúc dữ liệu mở được thiết kế nhằm hỗ trợ cả dữ liệu thô và nhu cầu phân tích ở cấp độ doanh nghiệp. Kiến trúc này cho phép tổ chức lưu trữ dữ liệu ở định dạng gốc trong kho lưu trữ đối tượng (object storage), đồng thời, áp dụng các lớp xử lý như siêu dữ liệu (metadata), chỉ mục hóa (indexing), ràng buộc lược đồ (schema enforcement), và nhật ký giao dịch đảm bảo tính toàn vẹn (ACID) nhằm đảm bảo độ tin cậy và khả năng truy vấn nhất

quán (Databricks, 2025).

Theo Armbrust et al. (2021), kiến trúc Lakehouse dựa trên nguyên lý hợp nhất, giúp phá vỡ “data silos” và loại bỏ nhu cầu phải xây dựng hệ thống song song (ETL từ lake sang warehouse). Điều này đặc biệt hữu ích trong ngân hàng, nơi dữ liệu khách hàng được lưu trữ rải rác ở CRM, giao dịch, kênh số, mạng xã hội và hệ thống rủi ro.

Trong bối cảnh dữ liệu lớn và nhu cầu phân tích tăng cao, kiến trúc Lakehouse đã nổi lên như một giải pháp tích hợp các ưu điểm của mô hình data lake và data warehouse, nhằm đáp ứng đồng thời các yêu cầu về lưu trữ linh hoạt và phân tích mạnh mẽ (Armbrust et al., 2021; Databricks, n.d.).

Hình 1. Kiến trúc Data Lakehouse



Nguồn: Bill Inmon, 2021

Lớp lưu trữ (Storage Layer): Dữ liệu trong Lakehouse được lưu trữ trong các hệ thống lưu trữ đối tượng đám mây như Amazon S3 hoặc Azure Data Lake, sử dụng định dạng tệp tối ưu như Parquet hoặc Delta Lake. Việc sử dụng định dạng mở giúp tăng tính linh hoạt và khả năng tương thích với nhiều hệ thống tính toán khác nhau (Databricks, n.d.).

Lớp metadata (Metadata Layer): Đây là yếu tố tạo nên sự khác biệt quan trọng của Lakehouse. Thông qua cơ chế quản lý transaction log như Delta Lake, hệ thống có thể đảm bảo tính nhất quán, hỗ trợ các thao tác ACID và cho phép truy xuất dữ liệu lịch sử. Điều này đặc biệt quan trọng đối với các tổ chức cần đảm bảo kiểm toán và tuân thủ (Armbrust et al., 2020).

Lớp tính toán (Compute Layer): Lakehouse hỗ trợ cả xử lý batch và streaming bằng các framework như Apache Spark hoặc

Databricks Runtime. Điều này giúp triển khai các quy trình ETL/ELT linh hoạt cũng như các bài toán học máy (AI/ML) trên cùng một nền dữ liệu (Databricks, n.d.).

Lớp truy cập và phân tích (Access Layer): Người dùng cuối có thể truy cập dữ liệu thông qua SQL, REST API, hoặc công cụ phân tích như Power BI, Looker. Lớp này đóng vai trò kết nối giữa dữ liệu thô và các bảng biểu, dashboard phục vụ ra quyết định (Databricks, n.d.).

Một số đặc điểm nổi bật của kiến trúc Lakehouse bao gồm:

- Quản lý giao dịch (ACID transactions) cho phép quản lý dữ liệu đáng tin cậy, tránh tình trạng sai lệch khi nhiều tiến trình truy cập cùng lúc (Armbrust et al., 2020).

- Các lược đồ (Schema enforcement & evolution), giúp duy trì tính toàn vẹn và thích nghi với thay đổi trong cấu trúc dữ liệu.

- Quản trị dữ liệu hợp nhất, cho phép kiểm soát quyền truy cập và bảo mật một cách tập trung trên toàn bộ chuỗi dữ liệu.

- Đạt hiệu năng cao nhờ áp dụng các kỹ thuật tối ưu như bộ nhớ đệm (caching), lập chỉ mục (indexing) và bỏ qua dữ liệu không liên quan (data skipping) (Databricks, n.d.).

So với kiến trúc Data Lake và Data Warehouse, Lakehouse được xem là một sự dung hòa hiệu quả giữa tính linh hoạt và khả năng phân tích cao. Trong khi Data Lake phù hợp cho lưu trữ thô và Data Warehouse phục vụ BI truyền thống, thì Lakehouse hợp nhất cả hai bằng cách cung cấp một nền tảng duy nhất cho cả lưu trữ và tính toán (Armbrust et al., 2021).

### **3. Ứng dụng cụ thể của Data Lakehouse trong quản trị khách hàng ngân hàng**

Trong bối cảnh ngành Ngân hàng đầy mạnh chuyển đổi số, nhu cầu hiểu khách hàng một cách toàn diện trở nên cấp thiết hơn bao giờ hết. Mô hình Customer 360, khi được triển khai trên kiến trúc Data Lakehouse, cho phép ngân hàng tổng hợp và phân tích dữ liệu khách hàng từ nhiều nguồn như core banking, ứng dụng số, mạng xã hội, và tương tác tại quầy - tất cả trong một nền tảng thống nhất. Theo Databricks, nền tảng Lakehouse hỗ trợ xây dựng hệ thống Customer 360 có thể giảm đáng kể thời gian triển khai so với hạ tầng truyền thống, nhờ tích hợp ingestion, quản lý metadata (Unity Catalog), phân tích với SQL/DataFrame, và tự động hóa với AutoML (Databricks, n.d.).

Tại một ngân hàng khu vực ở Mỹ, Công ty Tư vấn SunnyData đã triển khai thành công kiến trúc Delta Lakehouse, giúp ngân hàng này xây dựng hồ sơ khách hàng thống nhất, phân khúc theo hành vi và làm giàu dữ liệu giao dịch. Việc tích hợp Customer 360 trên Lakehouse đã giúp ngân hàng cải thiện khả năng nhận diện hành vi khách hàng, nâng cao hiệu quả chiến dịch tiếp thị và giảm thiểu tỷ lệ rời bỏ (SunnyData, 2023).

Như vậy, Data Lakehouse không chỉ cung cấp nền tảng dữ liệu mạnh mẽ và linh hoạt mà còn đóng vai trò chiến lược trong việc xây dựng tầm nhìn 360 độ về khách hàng - một trong những yếu tố cốt lõi giúp ngân hàng cá nhân hóa trải nghiệm, phát hiện sớm rủi ro, và gia tăng giá trị vòng đời khách hàng.

### **4. Lợi ích và thách thức khi triển khai Data Lakehouse trong quản trị khách hàng tại ngân hàng**

Trong quá trình chuyển đổi số, các ngân hàng đang đối mặt với nhu cầu xử lý và phân tích khối lượng lớn dữ liệu từ nhiều nguồn - từ giao dịch tài chính, dữ liệu khách hàng, đến phản hồi từ mạng xã hội. Kiến trúc Data

Lakehouse nổi lên như một giải pháp linh hoạt, kết hợp khả năng lưu trữ đa dạng và phân tích hiệu năng cao trong một nền tảng duy nhất. Theo IBM (2023), Lakehouse cung cấp khả năng phân tích dữ liệu thời gian thực, loại bỏ sự phân mảnh dữ liệu giữa các hệ thống, đồng thời, hỗ trợ truy vấn hiệu quả cho cả dữ liệu có cấu trúc và phi cấu trúc.

Một trong những lợi ích lớn nhất của Lakehouse là khả năng loại bỏ silo dữ liệu - một thách thức phổ biến trong hệ thống ngân hàng truyền thống. Dữ liệu từ core banking, tổng đài, ứng dụng số, và hệ thống CRM có thể được tích hợp vào một kiến trúc thống nhất, hỗ trợ xây dựng hồ sơ khách hàng 360 độ (Lumenalta, 2023). Ngoài ra, kiến trúc Lakehouse giúp giảm chi phí lưu trữ và bảo trì, vì sử dụng nền tảng lưu trữ linh hoạt giống như data lake, nhưng vẫn đảm bảo tính toàn vẹn dữ liệu thông qua các giao dịch ACID và cơ chế kiểm soát schema như trong Data Warehouse (Cloudian, 2022).

Lakehouse cũng tăng cường năng lực phân tích nâng cao, đặc biệt là trong các tác vụ sử dụng AI/ML như phân khúc khách hàng, dự báo rủi ro tín dụng, hoặc phát hiện gian lận. Theo Dremio (2024), các ngân hàng có thể triển khai các mô hình học máy trực tiếp trên nền tảng Lakehouse mà không cần di chuyển dữ liệu, từ đó giảm độ trễ và tăng tốc độ ra quyết định.

Tuy nhiên, triển khai Data Lakehouse trong thực tế không đơn giản. Một trong những thách thức lớn là việc tích hợp hệ thống cũ (legacy systems) với nền tảng hiện đại này. Ngân hàng thường có các hệ thống kế thừa phức tạp, yêu cầu thiết kế các quy trình ETL chuyên biệt để đảm bảo tương thích và chất lượng dữ liệu đầu vào (Dataversity, 2024). Bên cạnh đó, việc quản trị dữ liệu và tuân thủ quy định như GDPR, Basel III cũng đòi hỏi hệ thống Lakehouse phải có cơ chế quản lý metadata, lineage và quyền truy cập chi tiết - nếu không sẽ dễ phát sinh rủi ro vận hành và pháp lý (IBM, 2023; Lumenalta, 2023).

Ngoài ra, kiến trúc Lakehouse vẫn còn tương đối mới mẻ so với các giải pháp truyền thống, nên việc lựa chọn công cụ phù hợp, đào tạo đội ngũ và vận hành ổn định vẫn là những rào cản về năng lực tổ chức mà các ngân hàng cần cân nhắc kỹ trước khi triển khai quy mô lớn (Dremio, 2024).

### **5. Đề xuất ứng dụng Data Lakehouse vào quản trị quan hệ khách hàng tại ngân hàng**

Trong bối cảnh chuyển đổi số đang diễn ra mạnh mẽ, các tổ chức tài chính, đặc biệt là ngân hàng, ngày càng cần những giải pháp dữ liệu tiên tiến để nâng cao hiệu quả trong việc quản trị quan hệ khách hàng

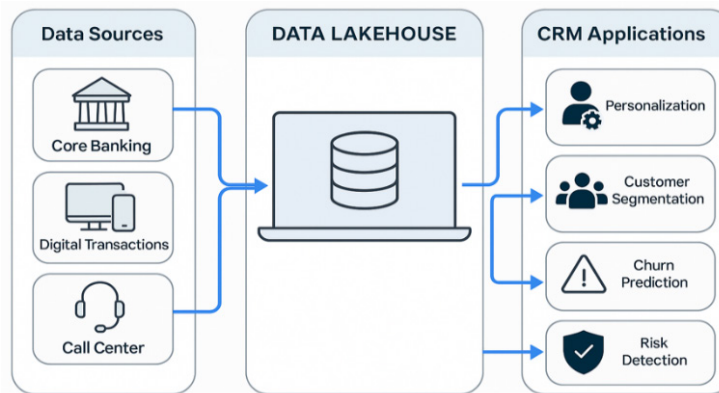
(CRM). Tuy nhiên, hệ thống CRM truyền thống thường gặp phải các hạn chế như dữ liệu phân mảnh giữa nhiều kênh (core banking, giao dịch số, tổng đài), thiếu khả năng phân tích thời gian thực, và khó khăn trong việc cá nhân hóa trải nghiệm khách hàng ở quy mô lớn (MacDonald, 2024).

Để giải quyết những vấn đề này, việc tích hợp kiến trúc Data Lakehouse vào hệ thống CRM được đề xuất như một hướng đi khả thi. Data Lakehouse kết hợp khả năng lưu trữ linh hoạt và chi phí thấp của data lake với các tính năng phân tích, truy vấn, và đảm bảo chất lượng dữ liệu của Data Warehouse (OWOX, 2025). Nền tảng này hỗ trợ lưu trữ và xử lý mọi loại dữ liệu - từ dữ liệu giao dịch có cấu trúc đến phản hồi từ khách hàng, dữ liệu cảm xúc từ mạng xã hội, và hành vi người dùng trên nền tảng

số - trong cùng một hệ thống lưu trữ định dạng mở như Parquet hoặc Delta Lake (Databricks, n.d.).

Về mặt kiến trúc, hệ thống CRM tích hợp Lakehouse bao gồm năm tầng chính: (1) tầng thu thập dữ liệu (ingestion) từ nhiều nguồn như CRM hiện tại, web/app, call center; (2) tầng lưu trữ đối tượng (storage layer) hỗ trợ mở rộng linh hoạt; (3) tầng metadata với khả năng quản lý schema, versioning, lineage, và ACID transaction để đảm bảo tính toàn vẹn và truy vết dữ liệu; (4) tầng truy vấn và phân tích (API/query engine) hỗ trợ SQL và DataFrame API cho các workload BI và học máy; và (5) tầng tiêu thụ dữ liệu (consumption layer) phục vụ các ứng dụng BI, mô hình AI, và dashboard phân tích hành vi khách hàng (MacDonald, 2024; OWOX, 2025).

**Hình 2.** Sơ đồ kiến trúc đề xuất mô hình Data Lakehouse quản trị quan hệ khách hàng



Nguồn: Tác giả đề xuất

Lợi ích khi tích hợp kiến trúc Lakehouse vào hệ thống quản trị quan hệ khách hàng (CRM) là rất rõ ràng. Hệ thống này hỗ trợ xây dựng hồ sơ khách hàng toàn diện 360 độ, nâng cao khả năng cá nhân hóa các khuyến nghị theo thời gian thực, dự báo khả năng rời bỏ của khách hàng, và nhận diện các phân khúc khách hàng tiềm năng dựa trên hành vi trong quá khứ. Bên cạnh đó, kiến trúc Lakehouse còn giúp đơn giản hóa quá trình kiểm toán và tuân thủ quy định (như Quy định bảo vệ dữ liệu chung - GDPR hoặc các chuẩn nội bộ trong ngành Ngân hàng) thông qua các cơ chế như lưu vết kiểm toán (audit trail), khôi phục dữ liệu (rollback) và quản lý truy cập theo vai trò người dùng (Databricks, n.d.; Fivetran, 2023). Với những ưu điểm về khả năng mở rộng, tích hợp và tối ưu vận hành, Data Lakehouse hứa hẹn sẽ trở thành hạ tầng dữ liệu cốt lõi trong chiến lược hiện đại hóa CRM của các ngân hàng thương mại.

## 6. Kết luận

Việc ứng dụng kiến trúc Data Lakehouse vào quản trị khách hàng ngân hàng mang lại nhiều lợi ích rõ rệt: từ khả năng tích hợp dữ liệu toàn diện (Customer 360), hỗ trợ phân tích nâng cao với AI/ML, đến tối ưu hóa hiệu quả vận hành và tuân thủ quy định. Mô hình này cho phép ngân hàng vượt qua những hạn chế cố hữu của CRM truyền thống như dữ liệu phân mảnh, thiếu chuẩn hóa và khó mở rộng theo thời gian thực. Tuy nhiên, để triển khai hiệu quả, các tổ chức tài chính cần chuẩn bị kỹ lưỡng về mặt kỹ thuật, hạ tầng và chiến lược vận hành, đặc biệt là trong việc tích hợp với hệ thống kế thừa, đảm bảo chất lượng dữ liệu, và xây dựng năng lực quản trị phù hợp. Trong tương lai, Data Lakehouse không chỉ là một công nghệ, mà là nền tảng quan trọng trong chiến lược số hóa của ngành ngân hàng, góp phần thúc đẩy khả năng cạnh tranh, cải thiện trải nghiệm khách hàng, và phát triển bền vững trong kỷ nguyên dữ liệu.

### Tài liệu tham khảo

- Armbrust, M., Das, T., Zhu, S., Yavuz, B., Huber, M., Liu, A., ... & Xin, R. (2020). Delta Lake: High-performance ACID table storage over cloud object stores. *Proceedings of the VLDB Endowment*, 13(12), 3411–3424. <https://doi.org/10.14778/3415478.3415560>
- Armbrust, M., Ghodsi, A., Xin, R. S., Zaharia, M., et al. (2021). *Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics*. CIDR 2021.
- Armbrust, M., Xin, R., Lian, C., Huai, Y., Liu, D., Bradley, J. K., ... & Zaharia, M. (2021). *Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics* [Whitepaper]. Databricks. <https://www.databricks.com/blog/2021/01/13/what-is-a-data-lakehouse.html>
- Cloudian. (2022). *Data Lakehouse: Is it the right choice for you?* <https://cloudian.com>
- Databricks. (2019). *What is a data lakehouse?* <https://databricks.com/glossary/what-is-a-data-lakehouse>
- Databricks. (2025). *The data lakehouse platform for dummies* [Whitepaper].
- Databricks. (n.d.). *Databricks solution accelerator: Customer 360*. Retrieved June 13, 2025, from <https://www.databricks.com/solutions/accelerators/customer-360>
- Dataversity. (2024). *Data lake strategy: Its benefits, challenges, and implementation*. <https://www.dataversity.net>
- Dremio. (2024). *The Lakehouse: Benefits, implementation challenges, and practical solutions*. <https://www.dremio.com>
- Drenik, G. (2025, April 24). How data integration transforms customer insights for banks. *Forbes*. <https://www.forbes.com/sites/garydrenik/2025/04/24/how-data-integration-transforms-customer-insights-for-banks>
- Fivetran. (2023). *The future of data warehousing: The rise of the lakehouse*. <https://www.fivetran.com/blog/future-of-data-warehousing>
- IBM. (2024, November 20). *Data warehouses vs. data lakes vs. data lakehouses*. <https://www.ibm.com/blogs/data-warehouses-vs-lakes-vs-lakehouses>
- Inmon, W. H., Levins, M., & Srivastava, R. (2021). *Building the data lakehouse*. Technics Publications.
- Lumenalta. (2024). *How a data lakehouse can reshape financial services*. <https://www.lumenalta.com/resources/data-lakehouse-financial-services>
- MacDonald, L. (2024). 5 layers of data lakehouse architecture explained. *Monte Carlo Data*. <https://www.montecarlodata.com/blog-data-lakehouse-architecture-5-layers>
- Maximizer. (2025). *Customer relationship management in banking: A 2025 guide*. <https://www.maximizer.com/blog/customer-relationship-management-in-banking>
- NetSuite. (2025, February 19). *The 8 top data challenges in financial services (with solutions)*. <https://www.netsuite.com/portal/resource/articles/financial-management/data-challenges-financial-services.shtml>
- OWOX. (2025). *Data lakehouse architecture and use cases*. <https://www.owox.com/blog/articles/data-lakehouse-architecture>
- SunnyData. (2023). *Empowering business insights at a regional bank using Delta Lake*. <https://www.sunnydata.io/blog/customer-360-delta-lakehouse>
- SyncMatters. (2024). *The biggest data integration challenges and solutions*. <https://syncmatters.com/blog/data-integration-challenges-and-solutions>
- Veritis. (2024). *Overcoming challenges in implementing CRM in banking*. <https://www.veritis.com/blog/overcoming-challenges-in-implementing-crm-in-banking>