

Ứng dụng dữ liệu lớn - thách thức đối với các ngân hàng thương mại Việt Nam

Đào Mỹ Hằng

Học viện Ngân hàng

Ngày nhận: 16/09/2020

Ngày nhận bản sửa: 06/10/2020

Ngày duyệt đăng: 21/10/2020

Đặng Thu Hoài

Ngân hàng Đại chúng

Dữ liệu có tầm ảnh hưởng quan trọng đối với hoạt động kinh doanh của mỗi ngân hàng thương mại (NHTM)- lĩnh vực kinh doanh ngành dịch vụ đặc thù dựa trên cơ sở thông tin. Sự xuất hiện của dữ liệu lớn (Big Data) trong bối cảnh kỷ nguyên số đang được NHTM kỳ vọng như một nguồn tài nguyên khổng lồ có thể mang đến cái nhìn đa chiều giúp ngân hàng đưa ra các quyết định sáng suốt. Tuy nhiên, việc ứng dụng Big Data vẫn đang là bài toán khó với nhiều thách thức đối với các NHTM tại Việt Nam. Nghiên cứu đề cập đến thực trạng ứng dụng và những thách thức cơ bản của Big Data đối với các NHTM tại Việt Nam và đưa ra các giải pháp khuyến nghị theo hướng tích hợp dữ liệu có cấu trúc truyền thống từ kho dữ liệu với các hệ thống nội bộ và dữ liệu phi cấu trúc trên Big Data.

Từ khóa: Dữ liệu lớn (Big Data), thách thức, ngân hàng thương mại.

Challenges in applying Big Data for the Vietnam commercial banks

Abstract: Data has an important influence on the business operations of each commercial bank. Commercial banks expect that the appearance of Big Data in the context of the digital era is a huge resource that can provide a multidimensional view to help banks make informed decisions. However, applying Big Data is difficult for commercial banks in Vietnam. The study discusses the application status and fundamental challenges of Big Data to commercial banks in Vietnam and recommends solutions.

Keywords: Big Data, Challenges, Commercial Banks

Hang My Dao

Email: myhang@hvn.edu.vn
Banking Academy of Vietnam

Hoai Thu Dang

Email: danghoai.bav@gmail.com
PvcomBank

1. Giới thiệu

Trong bối cảnh cạnh tranh khốc liệt của thị trường tài chính- ngân hàng (TC- NH), các NHTM đang từng bước ứng dụng các chuyển đổi số với trọng tâm là khách hàng (KH) nhằm không ngừng mang đến những trải nghiệm tốt hơn cho KH. Để làm được điều đó, NHTM cần có những chiến lược để hiểu được từng nhóm nhu cầu KH, từ đó có những quyết định phù hợp trong chính sách quản trị và hoạt động. Tại Việt Nam, tính đến tháng 01/2020, có khoảng 68 triệu người (khoảng 70% dân số) dùng internet, 65 triệu người (67% dân số) dùng mạng xã hội và 146 triệu lượt kết nối điện thoại (khoảng 150% dân số) (theo We are Social, 2020, “Digital 2020 global digital yearbook”), mọi thông tin về người dùng internet đều được lưu giữ lại thành dữ liệu theo đơn vị bit. Báo cáo này nhận định, sự gia tăng của nguồn dữ liệu này hoàn toàn có thể làm rung chuyển ngành dịch vụ TC- NH theo hướng tích cực hơn khi dữ liệu được thu thập, phân tích, giải thích liên quan đến xu hướng tài chính, từ đó vẽ nên một bức tranh chi tiết về xu hướng tiêu dùng và điều mà KH đang tìm kiếm. Đây sẽ là cơ sở rất quan trọng để các NHTM có thể cá nhân hóa dịch vụ đến từng KH, đáp ứng mong muốn KH tối ưu hơn trong định hướng chính sách sản phẩm, đặc biệt là nhóm sản phẩm có tính số hóa cao, đồng thời có thể xác định rủi ro kịp thời ở mọi thời điểm và có những thay đổi phù hợp trong điều hành, quản trị các hoạt động kinh doanh của ngân hàng để đạt được những mục tiêu lớn hơn. Tuy nhiên, việc khai thác, ứng dụng triển khai Big Data đang đặt rất nhiều thách thức cho các nhà quản trị NHTM và bộ phận công nghệ thông tin (CNTT), đặc biệt với các NHTM với mức vốn nhỏ. Để Big Data trở thành nguồn lực của ngân hàng, các NHTM cần trả lời được

các câu hỏi: (i) Làm thế nào để có thể khai thác giá trị từ nguồn dữ liệu Big Data một cách hiệu quả? (ii) Dữ liệu này có tương thích và tích hợp được với những phân hệ hiện tại hay không? (iii) Thực trạng điều kiện ngân hàng đang ở đâu, có phù hợp để ứng dụng triển khai không?

Để trả lời các câu hỏi đó, các nội dung tiếp theo sẽ được trình bày theo trình tự Tổng quan về Big Data, thực trạng ứng dụng Big Data tại các NHTM Việt Nam, thách thức của Big Data và giải pháp khuyến nghị đối với các NHTM.

2. Tổng quan về Big Data

2.1. Khái niệm Big Data

Big Data dùng để chỉ tập dữ liệu với kích thước vượt quá khả năng lưu trữ, quản lý và phân tích của phần mềm cơ sở dữ liệu điển hình (Manyika, 2011).

Big Data được hình thành từ việc sử dụng ngày càng tăng các công cụ điện tử và hệ thống thông tin, do các tổ chức và cá nhân hình thành trong đời thường dưới những hình thức khác nhau, bao gồm 6 nguồn dữ liệu chủ yếu sau: (i) dữ liệu hành chính; (ii) dữ liệu về hoạt động thương mại; (iii) dữ liệu từ các thiết bị cảm biến như thiết bị chụp ảnh vệ tinh, cảm biến đường, cảm biến khí hậu; (iv) dữ liệu từ thiết bị theo dõi; (v) dữ liệu từ hành vi (như tìm kiếm trực tuyến); (vi) dữ liệu từ thông tin về ý kiến, quan điểm cá nhân trên các phương tiện thông tin xã hội.

Cụ thể, dữ liệu Big Data có thể chia thành 3 loại theo hình thức như sau:

- Dữ liệu có cấu trúc (Structured): là những dữ liệu có thể được lưu trữ, truy cập và xử lý theo hình thức cố định truyền thống, đây là loại dữ liệu phổ biến nhất hiện nay tại các tổ chức.

- Dữ liệu không có cấu trúc (Unstructured):

những dữ liệu không có khuôn mẫu hoặc cấu trúc cố định, không phù hợp lưu trữ và quản lý với cấu trúc thông thường bằng hệ quản trị cơ sở dữ liệu truyền thống. Loại dữ liệu này có khối lượng lớn, không có cấu trúc và thay đổi nhanh, đặt ra rất nhiều thách thức trong xử lý và quản trị dữ liệu để có thể khai thác được giá trị từ nó. Ví dụ như dữ liệu từ các mạng xã hội như Facebook, Twitter, Instagram,... đều là những nguồn dữ liệu không có cấu trúc kết hợp từ văn bản, ảnh, video...

- Dữ liệu bán cấu trúc (Semi-structured): Dữ liệu kiểu bán cấu trúc có thể bao gồm tất cả các mẫu dữ liệu. Chúng ta có thể thấy dữ liệu bán cấu trúc như một mẫu nhưng nó không được định nghĩa với bảng với mối quan hệ như hệ quản trị quan hệ dữ liệu quan hệ (DBMS- Relational Database Management System). Loại dữ liệu này thường được thể hiện trong tập tin ngôn ngữ đánh dấu mở rộng (XML- Extensible Markup Language), có chức năng truyền dữ liệu và mô tả nhiều loại dữ liệu khác nhau.

2.2. Đặc tính của Big Data

Big Data có 5 đặc trưng cơ bản như sau (mô hình 5Vs về Big Data- theo Issues with big data, Gartner, 2018):

- Khối lượng dữ liệu (Volume): Đây là đặc điểm tiêu biểu nhất của Big Data, đó là khối lượng hoặc kích thước dữ liệu rất lớn. Kích cỡ của Big Data ngày càng tăng lên, dự kiến đến năm 2025 IDC dự đoán số liệu này sẽ là 175 zettabytes (10^{21} bytes) tương ứng với khoảng 59% công suất của nền công nghiệp cung cấp ổ cứng (HDD) toàn cầu (theo David Reinsel và cộng sự, 2018), nên các tổ chức sẽ cần sử dụng công nghệ “đám mây” mới đáp ứng khả năng lưu trữ được Big Data.

- Tốc độ (Velocity): Tốc độ có thể hiểu theo hai khía cạnh: (i) Khối lượng dữ liệu

gia tăng rất nhanh; (ii) Xử lý dữ liệu nhanh ở mức thời gian thực (real time) tính bằng mili giây. Công nghệ xử lý Big Data ngày nay đã cho phép chúng ta xử lý tức thì trước khi chúng được lưu trữ vào cơ sở dữ liệu.

- Đa dạng (Variety): Big Data cho phép liên kết và phân tích nhiều dạng dữ liệu khác nhau, ví dụ như kết nối thông tin của người dùng trên các nền tảng mạng xã hội khác nhau.

- Độ tin cậy/ chính xác (Veracity): Đây là một trong những tính chất phức tạp nhất của Big Data bởi các phương tiện truyền thông xã hội, mạng xã hội ngày càng phổ biến với đa dạng người dùng tương tác và chia sẻ nên việc xác định độ tin cậy và chính xác của dữ liệu ngày một phức tạp hơn.

- Giá trị (Value): Giá trị là đặc điểm quan trọng nhất của Big Data, vì khi bắt đầu triển khai xây dựng Big Data thì việc đầu tiên phải làm là xác định được giá trị của thông tin mang lại như thế nào, bởi dữ liệu tự bản thân không sinh ra giá trị, chỉ khi được xử lý và khai phá hợp lý thì chúng mới phát huy được lợi ích. Dựa trên các cân đối về nguồn lực khi đó chúng ta mới quyết định triển khai Big Data hay không.

2.3. Tầm quan trọng của ứng dụng Big Data đối với các ngân hàng thương mại

Big Data với nguồn dữ liệu khổng lồ, cập nhật từng phút, đang có vai trò lớn trong lĩnh vực ngân hàng với các ứng dụng cụ thể như: phân tích, phân loại sự hài lòng và hành vi KH; phân tích phát hiện và cảnh báo, ngăn chặn các hành vi rủi ro, giả mạo; tối ưu hóa hoạt động xử lý dữ liệu trong quá trình vận hành, phân tích và hỗ trợ ra quyết định.

Tiếp cận, nghiên cứu và khai thác Big Data sẽ mang đến nhiều lợi ích cho ngân hàng trong kinh doanh như: tiết giảm chi phí; tăng thời gian phát triển và tối ưu hóa

sản phẩm; đồng thời hỗ trợ ban lãnh đạo, cán bộ ngân hàng đưa ra những quyết định đúng và hợp lý hơn; tiết kiệm thời gian xử lý thông tin của KH và phòng chống rủi ro gian lận... Nghiên cứu của Manyika, J. (2017) cho thấy, khoảng 30% tất cả các công việc trong NHTM có thể được tự động hóa thông qua công nghệ và chìa khóa cho vấn đề này nằm ở Big Data. Điều này sẽ giúp NHTM có thể tiết kiệm đáng kể chi phí và giảm rủi ro bằng cách loại bỏ sự chi phối của yếu tố con người trong một số quy trình quan trọng.

Thực tế, bản thân NHTM cũng có được nguồn dữ liệu khá lớn từ chính các KH của mình qua các giao dịch trong hệ thống ngân hàng lõi (corebanking system). Nhưng chỉ bó hẹp trong lịch sử giao dịch đó là chưa đủ để NHTM hiểu về nhu cầu của KH, bởi trong thời đại con người ngày càng phụ thuộc vào Internet, thì nguồn dữ liệu lớn nhất phản ánh rõ mong muốn, thói quen và hành vi KH lại nằm ở mạng xã hội như Facebook, Twitter, Instagram,... và thông tin sử dụng Internet. Theo Oracle (2016), 84% các giám đốc điều hành ngân hàng được khảo sát đồng ý rằng KH đang tìm kiếm trải nghiệm phù hợp, cá nhân hơn và nếu ngân hàng có thể cung cấp cho người dùng những gì họ cần, ngân hàng có thể thúc đẩy doanh thu hàng năm cao hơn 18%. Do vậy, vấn đề cấp thiết đặt ra là cần kết nối các dữ liệu này với nhau để hình thành một cái nhìn tổng quan về KH (360° view). Khi ứng dụng Big Data, NHTM có thể sử dụng dữ liệu từ các mạng xã hội để các ngân hàng đưa ra chiến lược kinh doanh sát với nhu cầu của nhóm KH mục tiêu hơn. Kết quả nghiên cứu của Analytics, M. (2016) cho thấy rằng việc sử dụng Big Data để đưa ra quyết định có thể tiết kiệm tới 15- 20% ngân sách tiếp thị của ngân hàng. Có thể nói, khai thác Big Data là cơ hội tuyệt vời để không chỉ tiết kiệm chi phí mà còn tạo

thêm doanh thu thông qua các chiến lược tiếp thị được nhắm đến đúng mục tiêu. Đồng thời, từ nguồn thông tin dồi dào từ Internet và các mạng xã hội, ngân hàng có thể tương tác chặt chẽ hơn với KH, cải thiện dịch vụ KH một cách nhanh chóng, giảm ảnh hưởng của độ trễ thời gian. Bởi hệ thống truyền thống nhận phản hồi của KH đang bộc lộ nhiều hạn chế, và cần được cải thiện dần bởi hệ thống mới thiết kế với công nghệ xử lý Big Data và công nghệ xử lý ngôn ngữ tự nhiên.

Hơn nữa, công nghệ Big Data có thể được sử dụng để lưu trữ các dữ liệu mới trước khi chúng được đưa vào các kho chứa dữ liệu (Data Warehouse). Sự kết hợp giữa công nghệ Big Data có tốc độ xử lý ở mức thời gian thực với kho dữ liệu giúp cho các ngân hàng thoát khỏi việc không khai thác hết giá trị của dữ liệu do độ trễ thời gian. Do Big Data được ứng dụng các công nghệ mới giúp kiểm soát chặt chẽ và cập nhật dữ liệu liên tục sẽ ưu việt hơn dữ liệu truyền thống ở 4 điểm: dữ liệu đa dạng hơn, lưu trữ dữ liệu lớn hơn, truy vấn nhanh hơn và độ chính xác cao hơn (theo Cornelia Hammer và cộng sự, 2017).

Tóm lại, Big Data là cơ hội đồng thời cũng là thách thức đặt ra cho NHTM trong thời đại số hiện nay. Một khi làm chủ được Big Data từ việc trích xuất thông tin chính xác hơn, hữu ích hơn với chi phí thấp hơn thì NHTM sẽ có cơ hội thành công lớn hơn trong bối cảnh cạnh tranh ngày càng khắc nghiệt và nhu cầu phát sinh đa dạng hơn từ KH.

2.4. Cách tiếp cận Big Data

Hiện nay có nhiều cách tiếp cận triển khai các giải pháp Big Data trong các tổ chức, tuy nhiên dựa trên cơ sở nền tảng công nghệ, có ba cách tiếp cận phổ biến được áp dụng là (i) cách tiếp cận truyền thống, (ii) cách tiếp cận tiến hóa và (iii) cách tiếp

cận theo phương pháp lai (Nasser & Tariq, 2015). Các ngân hàng cận nhắc tùy vào các điều kiện thực tế về quy mô, nguồn lực hiện tại, mục tiêu mà mỗi đơn vị có một chiến lược khác nhau. Cụ thể:

Thứ nhất là cách tiếp cận truyền thống mang tính cách mạng, nghĩa là ngân hàng thiết lập một môi trường tính toán Big Data mới và chuyển tất cả dữ liệu sang nền tảng mới, ví dụ như Hadoop. Vì vậy tất cả các quá trình xử lý, phân tích, báo cáo và lập mô hình được thực hiện trên môi trường mới với đầy đủ thông tin, dữ liệu thông qua các ứng dụng và công nghệ phân tích mới (Business Intelligence & Analytics) để chuyển dữ liệu thô thành những mảng thông tin hữu ích toàn cảnh về quá khứ, dự đoán tương lai của tổ chức hướng đến xây dựng các giải pháp kinh doanh thông minh. *Thứ hai* là cách tiếp cận tiến hóa trong đó Big Data được xử lý bằng nền tảng Business Intelligence (BI) truyền thống hiện tại. Dữ liệu được thu thập và phân tích thông qua các công cụ có cấu trúc và không có cấu trúc và sau đó đầu ra được chuyển tiếp đến kho dữ liệu. Các tiện ích báo cáo và lập mô hình truyền thống giờ đây có thể truy cập suy nghĩ và bản ghi trực tuyến từ các nguồn truyền thông xã hội. Tuy nhiên, ngay cả khi cách tiếp cận tiến hóa đáp ứng được nhiều yêu cầu của môi trường Big Data, nó vẫn có hầu hết các vấn đề của BI cổ điển, nó có thể trở thành một nút cổ chai giữa thông tin trực tuyến từ các nguồn Big Data và phân tích sức mạnh của biến đổi BI hoặc kho dữ liệu.

Thứ ba là phương pháp lai mà cả công nghệ Big Data truyền thống và mới được sử dụng và dữ liệu được phân phối giữa hai nền tảng, một ví dụ về cách tiếp cận như vậy là giải pháp Hana từ SAP (công ty dẫn đầu thị trường về phần mềm ứng dụng doanh nghiệp trên thế giới, có trụ sở chính tại Đức). Đây là một giải pháp phân tích tích hợp dữ liệu, tính toán, nền tảng và

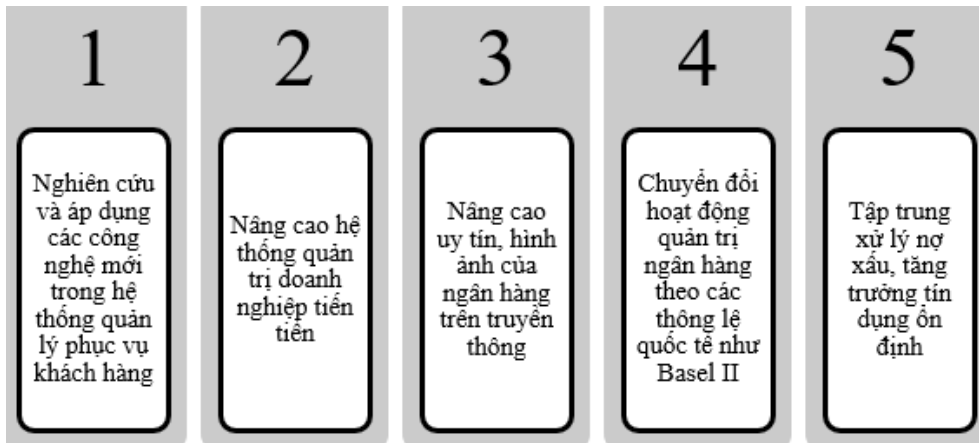
được chạy trên công nghệ điện toán bộ nhớ đệm (In-memory computing), công nghệ mới này có thể chạy đồng thời hệ thống xử lý giao dịch trực tuyến (OLTP- **Online Transaction Processing system**), hệ thống xử lý phân tích trực tuyến (OLAP- **Online Analytical Processing system**) và được xử lý trên bộ nhớ đệm máy tính, cho phép chuyển đổi các dòng dữ liệu thành các cột dữ liệu để gia tăng khả năng xử lý dữ liệu cho hệ thống, tối ưu hiệu suất phân tích dữ liệu và hiện nay được ứng dụng cho các tập đoàn lớn toàn cầu như Unilever, Colgate Palmolive...

3. Thực trạng ứng dụng Big Data tại các ngân hàng thương mại Việt Nam

Theo kết quả khảo sát của Vietnam Report (2019) với các NHTM đang hoạt động tại Việt Nam thì gần 3/4 (75%) số ngân hàng được hỏi cho biết sẽ ưu tiên cho hoạt động nghiên cứu và áp dụng công nghệ mới trong hệ thống quản lý, phục vụ KH; hơn 3/5 (60%) NHTM dự định đầu tư nâng cao hệ thống quản trị doanh nghiệp tiên tiến.

Trên thực tế, khu vực dịch vụ TC - NH ở Việt Nam đã có phản ứng nhanh so với các khu vực khác trong việc chủ động nghiên cứu và ứng dụng khoa học - công nghệ của Cách mạng công nghiệp 4.0 vào hoạt động kinh doanh và quản trị. Theo thống kê từ Vietnam Report (2019) thì 59% các doanh nghiệp đang hoặc dự kiến sẽ sử dụng công cụ Big Data để thúc đẩy chuyển đổi số hiệu quả. Đối với NHTM lớn có một số động thái tiêu biểu như:

- Năm 2020 TPBank chính thức đưa vào sử dụng hai giải pháp công nghệ Big Data do FIS triển khai bao gồm 2 cấu phần chính: Kho dữ liệu Data Lake được xây dựng dựa trên nền tảng mở Hortonworks Data Platform (HDP)- lưu trữ Big Data từ nhiều nguồn, bao gồm các nhóm dữ liệu thô, phi



Biểu đồ 1. Top 5 chiến lược ưu tiên của các NHTM trong năm 2019- 2020

Nguồn: Vietnam Report (2019), Khảo sát các NHTM Việt Nam

cấu trúc và Nền tảng xây dựng mô hình học máy Watson Studio Local, kết hợp thiết bị IBM Integrated Analytics System (IIAS) tối ưu cho việc phân tích dữ liệu với tốc độ cao, giảm thời gian huấn luyện mô hình (đây là hợp đồng đầu tiên hoàn thành về Big Data cho NHTM tại Việt Nam).

- Năm 2019 Vietcombank đã ký hợp tác với FIS triển khai thực hiện Dự án “Mua sắm phần mềm quản trị nhân sự mới (HRM)” gồm 6 phân hệ chính: Quản lý thông tin nhân sự, Quản lý chi phí tiền lương, Quản lý tuyển dụng, Quản lý đào tạo, Quản lý đánh giá cán bộ và Quản lý nhân tài.

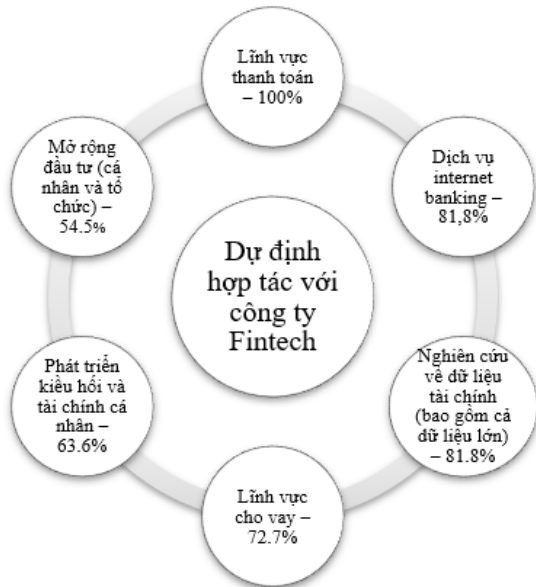
- VietinBank đã xây dựng chiến lược và lộ trình quản lý và khai thác tài sản dữ liệu trước khi chính thức ứng dụng Big Data thì chú trọng củng cố và xây dựng nền tảng hạ tầng về dữ liệu bao gồm dự án Kho dữ liệu doanh nghiệp (EDW- Enterprise Data Warehouse) triển khai xây dựng trong vòng 3 năm bao gồm tích hợp hơn 200 hệ thống nguồn, BI, Quản lý dữ liệu tập trung (MDM- Master Data Management).

- MB hợp tác với Infosys, Amigo triển khai kho dữ liệu và tập trung công cụ báo cáo quản trị...

Mặt khác, trong bối cảnh thời gian gần đây sự xâm nhập của công nghệ vào lĩnh vực

tài chính (Fintech- Công nghệ tài chính) ngày càng sâu rộng, ví dụ như ví Momo với lượng người dùng trong năm 2019 tăng từ 10 triệu (đầu 2019) lên hơn 15 triệu (cuối 2019), với hơn 100.000 điểm chấp nhận thanh toán (momo.vn) cho thấy sự tiện lợi đi kèm với xu hướng tiêu dùng, thanh toán online của KH sẽ tạo ra nguồn dữ liệu lớn và quan trọng tiết lộ về hành vi KH, nên 81,8% NHTM tại Việt Nam dự định sẽ hợp tác với các công ty Fintech trong lĩnh vực nghiên cứu về dữ liệu tài chính có bao gồm Big Data (Vietnam Report, 2019).

Cụ thể hơn về áp dụng Big Data trong chiến lược kinh doanh đặt KH là trọng tâm, trong nghiên cứu về “Ứng dụng dữ liệu lớn trong hoạt động quản trị quan hệ KH tại các NHTM Việt Nam” năm 2019, TS. Phan Thanh Đức và cộng sự đã sử dụng phương pháp BDMM (Big Data Maturity Model) và mô hình Hortonworks (Hortonworks, 2019) để đo lường và giám sát trạng thái dữ liệu, xác định mức độ trưởng thành Big Data của các NHTM Việt Nam. Thông qua khảo sát tại 36 NHTM, bộ câu hỏi được nhóm nghiên cứu sử dụng nhằm xác định mức độ trưởng thành Big Data tại các NHTM Việt Nam gồm 5 phần chính: (1) Định hướng chiến lược, (2) Dữ liệu và Phân tích dữ



Biểu đồ 2. Kế hoạch mở rộng cơ hội hợp tác với công ty Fintech của các NHTM Việt Nam 2019- 2020

Nguồn: Vietnam Report, (2019), Khảo sát các NHTM Việt Nam

liệu, (3) Công nghệ và Cơ sở hạ tầng, (4) Tổ chức và Kỹ năng & (5) Quản lý và Quy trình. Kết quả thu được: 72% các NHTM (26 ngân hàng) đang dừng ở mức độ 1- Nhận thức, 25% NHTM (9 ngân hàng) đã có mức độ trưởng thành dữ liệu lớn ở mức độ 2- Khai phá, 3% NHTM (1 ngân hàng) được đánh giá ở mức độ 3- Tối ưu, mức độ 4- Chuyển đổi thì hiện chưa có NHTM nào đạt được.

Trong đó, nhóm ngân hàng nhỏ chủ yếu nằm ở mức độ 1- Nhận thức: Ngân hàng có thể hiện sự quan tâm đến ứng dụng Big Data và bắt đầu có những nghiên cứu về phân tích dữ liệu, phân tích Big Data. Theo khảo sát sơ bộ từ một số cán bộ đang công tác tại NHTM nhỏ thì hiện tại công tác hiện đại hóa ngày càng được chú trọng hơn, thể hiện qua danh mục trong kế hoạch mua sắm CNTT, tỉ trọng chi phí phân bổ cho đầu tư các dự án CNTT- vận hành- hiện đại hóa ngân hàng trong chi phí hoạt động hàng năm cũng có xu hướng tăng. Các dự

án chú trọng xây dựng từng phần các hệ thống chuẩn hóa quy trình và nâng cao năng lực quản trị như hệ thống quản trị nội bộ (Management Information System- MIS) bao gồm các cấu phần về tình hình kinh doanh; tình hình nhân sự; quản trị rủi ro; hệ thống báo cáo thống kê; hệ thống quản lý KH... với kỳ vọng hướng đến trong tương lai sẽ tích hợp dữ liệu các hệ thống xây dựng thành kho dữ liệu có thể tích hợp ứng dụng Big Data. Tuy nhiên với hạn chế về quy mô, nguồn chi còn hạn hẹp và hạn chế từ cơ sở hạ tầng cũng như nhân lực triển khai nên các hệ thống dữ liệu tại NHTM nhỏ vẫn còn ít, lưu trữ độc lập, chưa có tính liên kết và chưa phủ đủ các mảng dữ liệu cần quản lý tự động, ví dụ như có NHTM thực hiện theo dõi nhân sự, tiền lương, KPI chỉ bằng file Excel tổng hợp từ các đơn vị gửi lên; chưa có đủ các ứng dụng ngân hàng điện tử cơ bản như Mobile banking; chức năng chuyển tiền nhanh 24/7 liên ngân hàng; chưa có hệ thống theo dõi thông tin KH sử dụng các tiện ích ngân hàng điện tử... Vì vậy để có thể hướng đến ứng dụng Big Data vẫn còn là một chặng đường rất dài với nhóm ngân hàng này.

Như vậy có thể nhận thấy, so với quy mô tổng thể của cả hệ thống ngân hàng Việt Nam, việc nghiên cứu, chủ động ứng dụng Big Data vẫn còn nhiều khó khăn và tiềm năng ứng dụng chủ yếu vẫn ở NHTM lớn- là nhóm có thể mạnh sẵn có về khoa học- công nghệ. Còn nhóm NHTM nhỏ với mô hình hoạt động vẫn theo truyền thống là chủ đạo cùng với các hạn chế về quy mô nên còn đang trong giai đoạn nghiên cứu sơ khai.

4. Thách thức trong ứng dụng Big Data tại ngân hàng thương mại Việt Nam

Về nguồn lực tài chính

Theo Vietnam Report (2019), việc yêu cầu nguồn vốn đầu tư lớn là khó khăn lớn nhất

đối với các tổ chức khi tiếp cận và ứng dụng công nghệ.

Hầu hết các doanh nghiệp Việt Nam muốn sử dụng Big Data đều phải thông qua một công ty cung cấp dịch vụ nước ngoài với chi phí bản quyền và đội ngũ tư vấn chuyên nghiệp dẫn đến những chi phí khá cao. Tính riêng về cơ sở hạ tầng - chi phí trả trước, chi phí của các nền tảng Big Data như Hadoop và Spark sẽ được chia tỷ lệ tương ứng với lượng lưu trữ, tính toán và sức mạnh xử lý mà NHTM sử dụng, ví dụ mỗi cụm Hadoop được khuyến nghị là ít nhất là một máy chủ Intel tầm trung, có giá từ 4.000 - 6.000 USD cho không gian đĩa 3TB và 6TB, một cụm Hadoop petabyte sẽ có giá khoảng 1 triệu USD. Chi phí nhân sự triển khai, ví dụ chi trả cho 1 chuyên gia Hadoop toàn thời gian tiêu tốn từ 70.000 - 150.000 USD/năm, mức giá nếu sử dụng nhân sự thuê ngoài là trung bình từ 81 - 100 USD/giờ. Ngoài ra còn có các chi phí khác như chi phí quản lý và bảo trì, chi phí kết nối mạng, chi phí cho nhà cung cấp proxy, chi phí bảo quản dữ liệu... Các NHTM lớn mặc dù có đủ vốn đầu tư, song họ cũng phải cân nhắc khi đặt trong bài toán chi phí, lợi nhuận và những ưu tiên chiến lược trong kế hoạch phát triển từng năm. Còn đối với các NHTM nhỏ thì việc đầu tư còn rất thận trọng để cân đối giữa các chi phí hoạt động

vận hành và đầu tư dài hạn cho hiện đại hóa hướng đến Big Data.

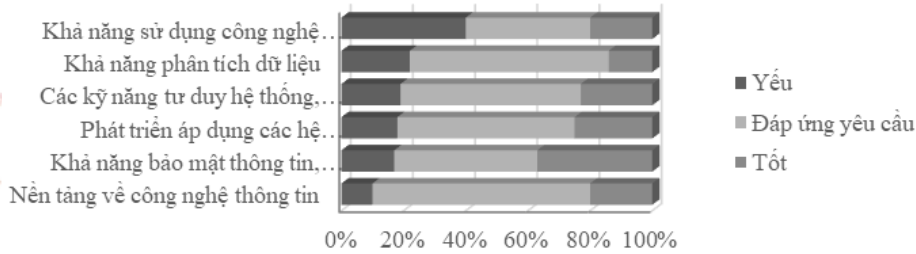
Về nguồn nhân lực

Hiện nay để triển khai ứng dụng và khai phá Big Data tạo nên các giá trị từ dữ liệu yêu cầu nhân sự có chất lượng cao và việc tuyển dụng rất cạnh tranh, không chỉ ở Việt Nam mà NHTM trong khu vực Asean đều đang cạnh tranh để chiêu mộ các tài năng nên nguồn nhân lực khá khan hiếm đối với các NHTM và đặc biệt khó khăn với NHTM nhỏ, nơi mà sức hấp dẫn về quy mô, thương hiệu còn thấp. Các NHTM đối mặt với vấn đề thiếu người tài để tận dụng sức mạnh của Big Data, thiếu kiến thức về thống kê, khai phá dữ liệu và các công nghệ sử dụng trong Big Data hiện nay như: hệ xử lý phân tán, công nghệ ảo hóa, công nghệ điện toán đám mây, cơ sở dữ liệu bên trong bộ nhớ (IMDS - In-memory Database system). Mặt khác, các nhà quản lý NHTM hiện nay thường xuất phát từ nghiệp vụ, phát triển hệ thống từng phần riêng biệt theo nhu cầu phát sinh, chủ yếu là các vấn đề liên quan đến kinh doanh, nên cần có sự tư vấn từ bộ phận CNTT để có thể xác định đủ định hướng sâu sắc trong xây dựng chiến lược phát triển đồng bộ cả hệ thống công nghệ số hóa đầy đủ trong dài hạn. Vì vậy, muốn có thể triển khai ứng dụng và tận dụng được một cách tối ưu hóa các tiềm



Biểu đồ 3. Top 3 khó khăn doanh nghiệp gặp phải khi tiếp cận và ứng dụng công nghệ

Nguồn: Vietnam Report, (2019), Báo cáo Vietnam CEO Insight 2019



Biểu đồ 4. Đánh giá kỹ năng nhân viên trong thời đại số

Nguồn: Vietnam Report, (2019), Báo cáo Vietnam CEO Insight 2019

năng của Big Data thì các nhà quản trị cần có những bước chuẩn bị thật kỹ lưỡng về chiến lược phát triển và đặc biệt là yếu tố con người, bởi dù công nghệ có tiên tiến đến đâu thì con người vẫn là yếu tố cốt lõi trong vận hành không thể thay thế được.

Về công nghệ

Với khối lượng dữ liệu khổng lồ như vậy, rõ ràng việc lưu trữ và xử lý Big Data là một thách thức không hề nhỏ. Hạ tầng CNTT của các doanh nghiệp Việt Nam nói chung và NHTM nói riêng còn chưa đủ mạnh để có thể khai thác một cách tối ưu hết các tiện ích của Big Data. Ngoài ra, Big Data có nhiều đặc điểm tạo nên những thách thức về mặt công nghệ và kỹ thuật, bao gồm thu thập dữ liệu, hiệu chỉnh, lọc nhiễu, tích hợp dữ liệu yêu cầu một thiết kế lưu trữ thông minh, xử lý truy vấn với tốc độ nhanh nhất, mô hình hóa và phân tích đầu ra với những yêu cầu dành cho phân tích theo nhu cầu (báo cáo ad-Hoc) và các báo cáo chuyên sâu ngày càng phức tạp, đáp ứng nhu cầu kịp thời cho các nhà quản trị.

Những công nghệ và công cụ hỗ trợ cần thiết, bao gồm: hệ sinh thái Hadoop; cơ sở dữ liệu NoSQL (Not only SQL) có thể được thu nhỏ theo chiều ngang trên hàng trăm hoặc hàng nghìn máy chủ; cơ sở dữ liệu bộ nhớ trong (IMDB) với ưu điểm nhanh hơn so với cơ sở dữ liệu trên đĩa được tối ưu hóa, một lựa chọn quan trọng cho việc sử dụng phân tích Big Data tạo kho dữ liệu. Đồng thời, việc xử lý kỹ thuật, lựa chọn

công nghệ, cập nhật thuật toán trong đảm bảo tính riêng tư, bảo mật dữ liệu với khối lượng lớn trong xu hướng lưu trữ “đám mây” ngày càng tăng cũng là một vấn đề hóc búa với bộ phận CNTT của các NHTM.

Về dữ liệu

Thực tế hiện nay dữ liệu ở NHTM Việt Nam còn khá manh mún và chưa được làm sạch một cách đạt yêu cầu, cũng chưa có sự phân loại rõ ràng về dữ liệu có cấu trúc và phi cấu trúc nên việc tích hợp dữ liệu đòi hỏi làm sao chuẩn hoá các dữ liệu cả đầu vào và đầu ra, ví dụ như trường hợp NHTM triển khai các hệ thống từ các đối tác CNTT khác nhau cũng chưa có chuẩn mực ngay từ đầu sẽ dẫn đến khó khăn hơn cho việc hợp nhất dữ liệu nếu sử dụng các nền tảng khác nhau như Oracle và SQL. Ngoài ra, theo các nghiên cứu trong nước như “Tương lai nền kinh tế số Việt Nam hướng tới năm 2030 và 2045” (CSIRO, 2019), “Dữ liệu lớn: Cách thức khai thác cơ hội từ dữ liệu? Trường hợp Amazon” (Nguyễn Anh Duy, 2019),... thực trạng tâm lý và hành vi người tiêu dùng Việt Nam là hay thay đổi và thường có sở thích, xu hướng ngắn hạn chi phối bởi các cộng đồng số và những người có tầm ảnh hưởng (Influencer), do vậy, việc phân tích nhu cầu KH để đưa vào áp dụng là thách thức với các NHTM, vì nếu nhận định sai xu hướng nhu cầu thị trường Big Data có thể trở thành một con dao hai lưỡi. Do đó, NHTM còn đối mặt với việc xác định thời gian dữ liệu hợp lệ,

trả lời câu hỏi trong bao lâu chúng ta nên giữ thông tin đó trong cơ sở dữ liệu và thời điểm nào thì dữ liệu không còn phù hợp với phân tích để đảm bảo được chất lượng thông tin đầu ra. Mặt khác, NHTM cần lưu ý trong khai thác dữ liệu KH cần tuân thủ các quy định của pháp luật, ví dụ theo Nghị định số 117/2018/NĐ-CP “Về việc giữ bí mật, cung cấp thông tin KH của tổ chức tín dụng, chi nhánh ngân hàng nước ngoài” có hiệu lực từ 01/11/2018. Đây là một thách thức quan trọng khi việc quan tâm đến bảo mật thông tin của KH ngày càng tăng, mà ngân hàng lại đang có xu hướng sử dụng lưu trữ đám mây và hợp tác nhiều hơn với các tổ chức fintech. Không chỉ vậy, ngân hàng với mạng lưới các điểm giao dịch lớn trải dài trên khắp đất nước và có một số ngân hàng có chi nhánh tại nước ngoài có quy mô nhân viên tương đối lớn, thực hiện nhiều loại giao dịch với khách hàng mỗi ngày nên ngân hàng cần chú ý vấn đề xây dựng quy trình kiểm soát, phân quyền truy cập hệ thống và đào tạo nhân sự nội bộ bài bản về nhận diện và bảo mật thông tin khách hàng. Từ những phân tích trên cùng với thực tế tội phạm an ninh mạng ngày càng nhiều, với trình độ tinh vi đến từ khắp nơi trên thế giới chỉ với kết nối mạng thông thường, các ngân hàng cần rất thận trọng trong việc đảm bảo các yếu tố bảo mật dữ liệu từ các hệ thống đến nhân sự, đối tác, nếu không đảm bảo được yếu tố này, ngân hàng rất có thể đối mặt với rủi ro về khủng hoảng truyền thông, hình ảnh thương hiệu cũng như hoạt động kinh doanh.

Tóm lại, bốn nhóm thách thức chính nêu trên phản ánh thực tế rằng các NHTM Việt Nam cần thời gian để chuẩn bị và tìm kiếm giải pháp phù hợp để có thể ứng dụng Big Data. Nếu như việc áp dụng triển khai Big Data không có lộ trình rõ ràng, thực hiện không đúng cách, người triển khai không đủ năng lực sẽ có thể nảy sinh nhiều vấn đề

từ thiết kế, lưu trữ đến quản lý truy cập dữ liệu, an toàn bảo mật.

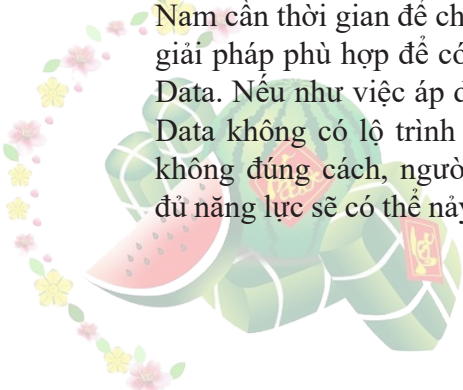
5. Khuyến nghị và giải pháp đối với các ngân hàng thương mại Việt Nam

Để có thể ứng dụng Big Data trong NHTM với một mục tiêu dài hạn cần chuẩn bị kỹ càng nhiều yếu tố, cần có quan điểm quản trị đúng đắn, xác định rõ nhu cầu, mục tiêu cùng với cân nhắc các nguồn lực, từ đó có lộ trình với những mức độ ưu tiên riêng cho từng cấu phần trong chiến lược tổng thể. Thông qua trao đổi với các chuyên gia có kinh nghiệm thực tế trong việc triển khai các dự án về Kho dữ liệu Doanh nghiệp, chuyển đổi số hướng đến ứng dụng Big Data và Trí tuệ nhân tạo từ các ngân hàng Vietinbank, BIDV, Công ty ví điện tử Momo, Công ty Trusting Social, Vintech và những chia sẻ khác tại Hội thảo Quản lý dữ liệu thông minh trong lĩnh vực ngân hàng, tài chính tháng 9/2020, nhóm tác giả khuyến nghị NHTM áp dụng theo 7 bước ở cấp độ “High level” trong phân tích cân nhắc lợi ích trước khi triển khai Big Data, góp phần đảm bảo dự án Big Data đi đúng hướng:

Bước 1. Xác định các tình huống kinh doanh thực tế cần cải thiện: NHTM cần đặt ra những câu hỏi chi tiết trong các vấn đề để đảm bảo các ưu tiên cho các vấn đề cần cải thiện đưa ra hợp lý nhất với điều kiện thực tại.

Bước 2. Hiểu quy trình kinh doanh và các sản phẩm: Cách thức hoạt động, quy trình và đặc trưng sản phẩm tại mỗi NHTM sẽ có sự khác biệt, việc nắm rõ là rất quan trọng trong triển khai dự án.

Bước 3. Hiểu cách phân loại và mức độ phân tích cần thiết: Để NHTM có cái nhìn hoàn chỉnh về thị trường và cách cạnh tranh hiệu quả đòi hỏi một khung phân tích mạnh mẽ bao gồm phân tích mô tả, phân tích dự đoán và phân tích theo quy định.



Bước 4. Tiến hành khám phá chi tiết: Các NHTM cần tìm ra quá trình nào bị ảnh hưởng bởi dữ liệu và thực hiện bắt đầu với các quy trình kinh doanh liên quan trực tiếp đến việc tạo doanh thu, sau đó là quy trình hỗ trợ tối ưu hóa.

Bước 5. Xác định các yêu cầu chính về dữ liệu: Dựa trên nhu cầu và tình hình hoạt động kinh doanh và quản trị, các điểm dữ liệu chính, số liệu và điểm đánh dấu cần được xác định.

Bước 6. Thiết kế kiến trúc dữ liệu mạnh mẽ: Kiểu kiến trúc dữ liệu tốt nhất là kiến trúc có khả năng mở rộng và giải quyết các yêu cầu của NHTM ít nhất là trong thập kỷ tới, khi dữ liệu tăng dần lên và yêu cầu phân tích thời gian thực cùng Dashboard sáng tạo.

Bước 7. Triển khai ứng dụng: Khi một tầm nhìn rõ ràng đã được đặt ra với một cam kết thực hiện từ các bên có bao gồm đối tác triển khai thì sẽ giúp giảm bớt các rào cản khi cùng hướng đến mục tiêu chung dựa trên nguyên tắc.

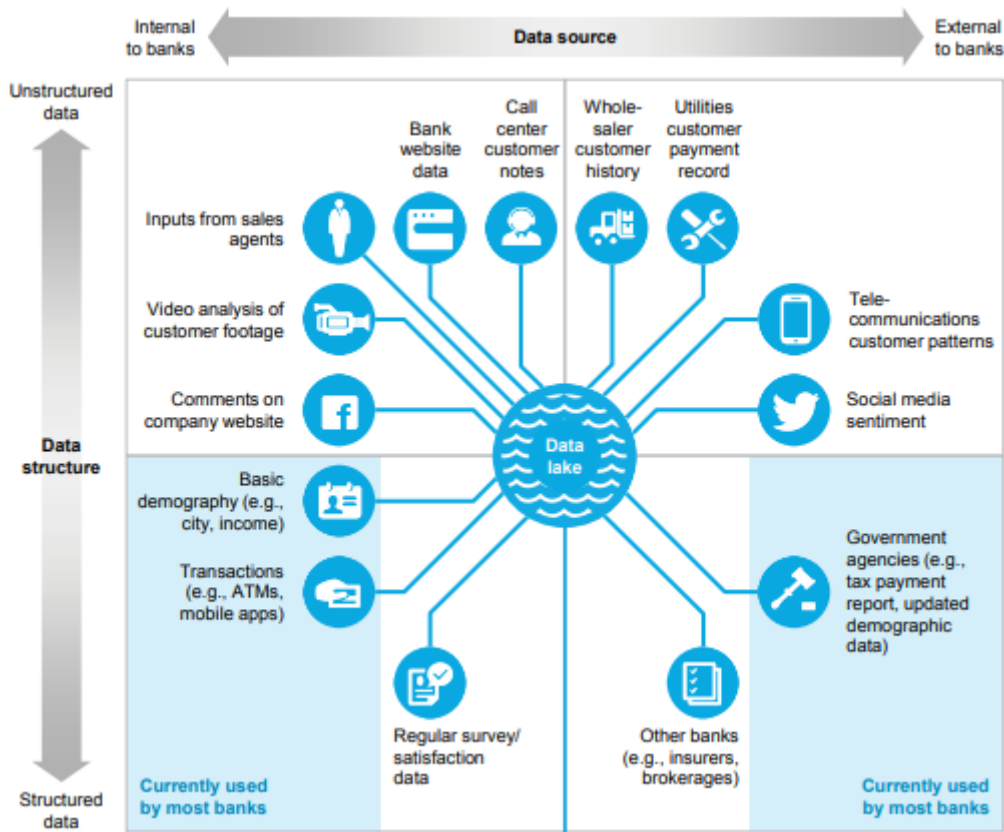
Ngoài ra, để chuẩn bị nền tảng và đối mặt với những thách thức đã phân tích, bài viết đưa ra bốn nhóm giải pháp như sau:

Một là, giải pháp về tài chính: Để phù hợp với nguồn lực tài chính, NHTM nên tận dụng sự hỗ trợ của công nghệ điện toán đám mây. Hiện nay, các nhà cung cấp giải pháp Big Data trên đám mây đã có thể xây dựng sẵn các mô hình phân tích và mang lại khả năng các tổ chức có thể chi trả cho các nguồn lực siêu tính toán theo phương thức chi tiêu tùy theo khả năng. Ngoài ra, để giảm thiểu một phần chi phí cho việc thu thập và làm sạch dữ liệu, NHTM nên đẩy mạnh hợp tác với các công ty tài chính công nghệ fintech, hướng đến xây dựng mô hình chuỗi cung ứng mới gắn kết chặt chẽ hơn với nhu cầu của KH. Khi đó, chuỗi cung ứng dịch vụ ưu việt hơn sẽ tạo ra một cơ sở dữ liệu thống nhất, minh bạch và hiệu

quả hơn ở mọi giai đoạn, từ khi KH phát sinh nhu cầu cho đến khi hoàn tất thực hiện cung ứng dịch vụ.

Hai là, giải pháp về công nghệ bao gồm kết cấu hạ tầng thông tin nhằm đáp ứng nền tảng về công nghệ để ứng dụng Big Data. Muốn sử dụng và khai thác Big Data một cách hiệu quả thì điều cần thiết là phải xây dựng cơ sở hạ tầng đủ để thu thập và lưu trữ dữ liệu, cung cấp quyền truy cập và bảo mật thông tin trong khi lưu trữ và chuyển tiếp, bao gồm: các hệ thống lưu trữ và máy chủ; phần mềm quản lý, tích hợp dữ liệu; phần mềm phân tích dữ liệu và một số thành phần khác. NHTM nên lựa chọn các giải pháp Big Data có nền tảng kết nối và tận dụng được dữ liệu từ các hệ thống hiện tại để không lãng phí cơ sở hạ tầng, dữ liệu và nguồn lực đã thực hiện trước đó.

Ba là, nhóm giải pháp về nguồn nhân lực có đủ năng lực sáng tạo, nền tảng công nghệ của Cách mạng công nghiệp 4.0 nói chung và ứng dụng Big Data nói riêng. Các nhà quản trị cần tham khảo các mô hình thành công trên thế giới và có hoạch định rõ về chiến lược phát triển dài hạn hiện đại hóa NHTM. Ngoài việc tuyển dụng mới nhân sự có kinh nghiệm và trình độ về hiện đại hóa ngân hàng, ứng dụng Big Data thì cần ưu tiên chiến lược phát triển nhân sự nội bộ để xây dựng nguồn nhân lực mạnh với chi phí hợp lý, như phổ cập tham gia các khóa đào tạo về phân tích dữ liệu, lập trình và cơ sở dữ liệu. Đối với bộ phận nghiệp vụ, cần tiếp cận thêm các ngôn ngữ lập trình ứng dụng trong xử lý phân tích dữ liệu như R, Python, SQL để chủ động trong xử lý dữ liệu và có cái nhìn bao quát hơn về hệ thống dữ liệu. Đối với bộ phận CNTT, yêu cầu đặt ra có thể tiếp cận và nhanh chóng tiếp thu các xu hướng công nghệ mới của thế giới trong thiết kế cơ sở dữ liệu, xử lý truy vấn, tích hợp dữ liệu ví dụ như công cụ quản lý dữ liệu phân tán (Hadoop, Dryad),



Biểu đồ 5: Giải pháp Data Lake kết nối dữ liệu truyền thống và dữ liệu phi cấu trúc của ngân hàng

Nguồn: Analytics, M. (2016), The age of analytics: Competing in a data-driven world

hệ cơ sở dữ liệu không cấu trúc NoQuery, các công cụ xử lý (MapReduce)...

Bốn là, giải pháp về dữ liệu. Tuy các nhà cung cấp đám mây hiện nay mang lại một số lợi ích an ninh hơn so với trung tâm dữ liệu thông thường của NHTM như bảo vệ và giám sát dữ liệu tập trung hơn, tuy nhiên nó cũng mang lại rủi ro khi mà dữ liệu có thêm một tổ chức trung gian nắm giữ. Vì vậy, các NHTM cần áp dụng các tiêu chuẩn cao về đối tác, tuyển dụng và quản lý con người, hệ thống; đặc biệt quan tâm đến việc xây dựng trung tâm dự phòng dữ liệu (khôi phục dữ liệu sau thảm họa); nâng cấp hệ thống an ninh, bảo mật ở mức cao; đảm bảo việc mở rộng phạm vi hoạt động (nếu có) được ổn định, an toàn, mang lại hiệu quả lâu dài.

6. Kết luận

Nhờ khả năng kết hợp nhiều công nghệ, kỹ thuật và phương pháp hiện đại, tối ưu, Big Data đang dần thể hiện rõ tầm quan trọng của mình đối với NHTM hiện tại và trong tương lai. Tuy nhiên NHTM cần hết sức thận trọng trong việc ứng dụng nó để phù hợp với thực tiễn, khắc phục các hạn chế, thách thức và đảm bảo an toàn thông tin thông qua xây dựng một lộ trình triển khai phù hợp dựa trên khuyến nghị tiếp cận theo 7 bước ở cấp độ “High level”, và một số giải pháp được đúc kết dựa trên thực tế từ các chuyên gia đã triển khai và ứng dụng thành công vào các ngân hàng, fintech. Một khi có thể làm chủ, tận dụng, kết hợp giữa hệ thống dữ liệu hiện tại và Big Data thì

NHTM sẽ có cơ hội thành công lớn hơn trong bối cảnh cạnh tranh ngày nay và hưởng lợi nhiều hơn từ việc trích xuất thông tin một cách chính xác, hữu ích. Từ những nền tảng ban đầu trong phân tích này, nhóm nghiên cứu sẽ hướng đến nghiên cứu cụ thể theo các nhóm ngân hàng tương đồng về quy mô và ứng dụng công nghệ mới, từ đó đưa ra những giải pháp hữu ích hơn để xây dựng nền kinh tế số, giúp các tổ chức ngân

hàng có thể mang đến trải nghiệm dịch vụ tốt nhất tới khách hàng ■

Tài liệu tham khảo

- Analytics, M. (2016). The age of analytics: competing in a data-driven world. McKinsey Global Institute Research Banking Academy, (2017), Big Data for Banking and Financial Industry, Science and technics publishing house, ISBN: 9786046709305.*
- Cornelia Hammer, Diane C Kostroch, Gabriel Quiros, (2017), Big Data: Potential, Challenges, and Statistical Implications, ISBN/ISSN:9781484310908.*
- Cơ quan Nghiên cứu Khoa học và Công nghiệp Khó thịnh vượng chung Úc (CSIRO), (2019), Tương lai nền kinh tế số Việt Nam hướng tới năm 2030 và 2045.*
- Cục Thông tin Khoa học và Công nghệ Quốc gia, (2015), Tổng luận Dữ liệu lớn và xu hướng đổi mới sáng tạo dựa trên dữ liệu số.*
- Domo, (2020), Data never sleeps 7.0 (www.domo.com).*
- Đào Văn Hùng, (2019), Phát triển khu vực tài chính- ngân hàng trong bối cảnh Cách mạng công nghiệp lần thứ tư, Tạp chí Cộng sản số 6/2019.*
- Fred Zimmerman, (2015), Hadoop Business Case: A Cost Effective Queryable Data Archive/Storage Platform.*
- Gärtner, B., & Hiebl, M. R. (2018). Issues with big data. The Routledge Companion to Accounting Information Systems (S. 161-172). New York: Routledge.*
- Gordon, J., Perrey, J., & Spillecke, D. (2013). Big data, analytics and the future of marketing and sales. McKinsey: Digital Advantage.*
- Hortonworks, (2019), Data Strategy Scorecard Survey.*
- Học viện Ngân hàng, (2019), Kỳ yếu Hội thảo Hệ thống thông tin quản lý trong kỷ nguyên số, NXB Thanh Niên, ISBN: 9786049756801.*
- Mary Aleksandrova, (2020) Big Data in the Banking Industry: The Main Challenges and Use Cases. . McKinsey Global Institute Research*
- Manyika, J. (2011). Big data: The next frontier for innovation, competition, and productivity. . McKinsey Global Institute Research*
- Manyika, J. (2017). A future that works: AI, automation, employment, and productivity. McKinsey Global Institute Research, Tech. Rep, 60.*
- Nasser T & Tariq RS, (2015), Big Data Challenges, Journal of Computer Engineering & Information Technology Volume 4 Issue 3.*
- Ngô Kim Thanh, (2020), Ứng dụng dữ liệu lớn trong nền kinh tế số.*
- Nghị định số 117/2018/NĐ-CP Về việc giữ bí mật, cung cấp thông tin khách hàng của tổ chức tín dụng, chi nhánh ngân hàng nước ngoài, ban hành ngày 11/09/2018.*
- Nguyễn Anh Duy, Nguyễn Phúc Quỳnh, (2019), Dữ liệu lớn: Cách thức khai thác cơ hội từ dữ liệu? Trường hợp Amazon, Tạp chí Phát triển & Hội nhập số 46.*
- Oracle, (2016), The Era I Enterprise: Ready for Anything, Financial Services- Report*
- Phan Thanh Đức và cộng sự, (2019), Ứng dụng dữ liệu lớn trong hoạt động quản trị quan hệ khách hàng tại các NHTM Việt Nam, Tạp chí Khoa học & Đào tạo Ngân hàng số 203.*
- Philip Russom, (2015), Hadoop for the Enterprise: Making DataManagement Massively Scalable, Agile, Feature-Rich, and Cost-Effective.*
- Reinsel, D., Gantz, J., Rydning, J., (2018), Data Age 2025: The Digitization of the World from Edge to Core, IDC Vietnam Report, (2019), Báo cáo Vietnam CEO insight 2019: chuyển đổi số và cơ hội của các doanh nghiệp Việt Nam*
- We are Social, (2020), Digital 2020 global digital yearbook.*