**CÔNG NGHỆ XỬ LÝ VÀ PHÂN TÍCH DỮ LIỆU LỚN TRONG QUỐC PHÒNG VÀ DÂN SỰ**

**BÙI TĂNG BẢO NGỌC (Trung tâm Chỉ huy Điều khiển)**

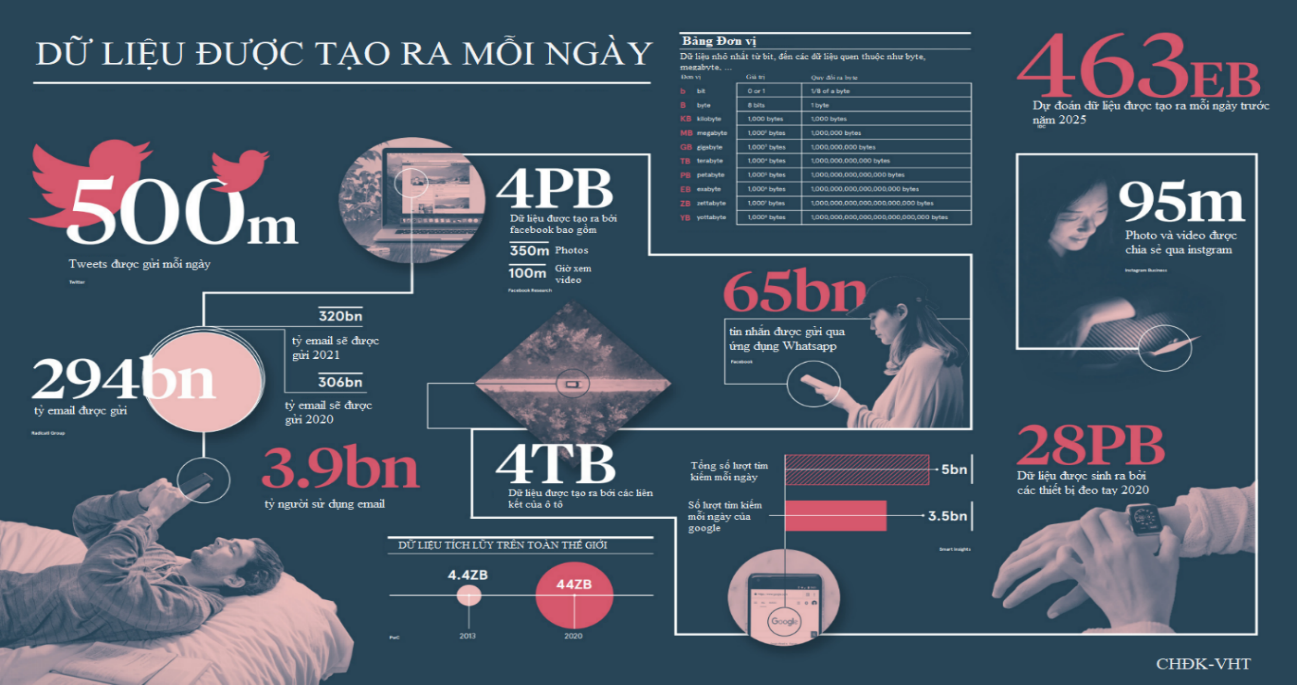
**TÓM TẮT** **Thế giới đang chứng kiến một cuộc cách mạng công nghiệp mới được thúc đẩy bởi các công nghệ Dữ liệu lớn, Internet kết nối vạn vật (IoT) và các ro bốt tự động hóa hay sự giao thoa giữa các xu hướng công nghệ và các vấn đề trong phát triển kinh tế - xã hội, bao gồm cả việc sử dụng Internet ngày càng tăng. Việc nắm bắt và làm chủ các công nghệ lõi trong việc xử lý và phân tích Dữ liệu lớn tại trung tâm Chỉ huy Điều khiển để dễ dàng chuyển dịch từ các nhiệm vụ quân sự sang lĩnh vực an ninh, dân sự theo kịp xu hướng phát triển trên thế giới và dẫn dắt công nghệ cùng lĩnh vực trong nước.**

**TỪ KHÓADữ liệu lớn, Phân tích dữ liệu lớn, Ứng dụng phân tích dữ liệu lớn, Dữ liệu lớn cho quốc phòng, an ninh, Dữ liệu lớn trong lĩnh vực dân sự.**

1. **GIỚI THIỆU CHUNG**

**1.1. TỔNG QUAN VỀ DỮ LIỆU LỚN (BIG DATA)**

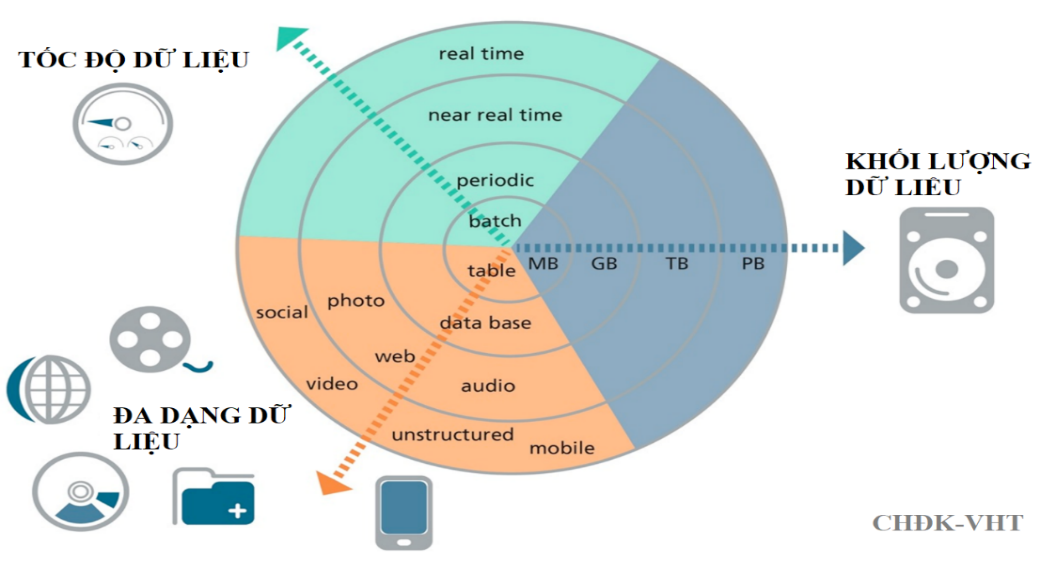
Hiện nay, nguồn tài nguyên giá trị nhất của thế giới không còn là dầu mỏ, mà là kho dữ liệu số đang tăng lên với cấp độ lũy thừa mỗi ngày.



**Hình 1: Dữ liệu lớn được tạo ra mỗi ngày trên thế giới**

Theo báo cáo của Công ty Nghiên cứu thị trường IDC (Hoa Kỳ) [1] được minh họa tóm tắt qua Hình 1 cho thấy: mỗi ngày có xấp xỉ 500 triệu tweet và 294 tỷ email được gửi, 4 Petabyte dữ liệu được tạo trên Facebook, 4 Terabyte dữ liệu từ mỗi chiếc xe ô tô được kết nối, 65 tỷ tin nhắn được gửi trên WhatsApp, 5 tỷ lượt tìm kiếm được thực hiện và đến năm 2025 ước tính rằng có 463 Exabyte dữ liệu sẽ được tạo ra mỗi ngày trên toàn cầu và những con số này vẫn không ngừng gia tăng. Về mặt thị trường, theo báo cáo tài chính từ Forbes [2], doanh thu từ thị trường dữ liệu lớn dự đoán sẽ tăng đến 99 tỷ USD vào năm 2022 với tốc độ tăng trưởng kép là 42%.

Về bản chất, dữ liệu lớn là thuật ngữ dùng để chỉ một tập hợp dữ liệu rất lớn và phức tạp, khi mà các khả năng tính toán truyền thống (lưu trữ, phân tích, chuyển mạng và trực quan hóa) không còn có thể xử lý với số lượng, tốc độ, độ phức tạp. Chúng thường có 3 thuộc tính đặc trưng là: **Kích thước lớn (big volume)** dữ liệu có kích thước lớn có thể lên đến hàng ngàn tỉ Gigabyte hoặc thậm chí lớn hơn; **Tốc độ dữ liệu nhanh (velocity)** dữ liệu đảm bảo xử lý các thao tác như truy xuất, cập nhật, chỉnh sửa với tốc độ nhanh tới giây thậm chí là một vài mili giây. Sự đa dạng trong dữ liệu (variety): Dữ liệu không cần tuân theo một cấu trúc và có thể lưu trữ nhiều định dạng khác nhau như văn bản, hình ảnh, âm thanh video, …



**Hình 2: Các đặc trưng cơ bản của dữ liệu lớn**

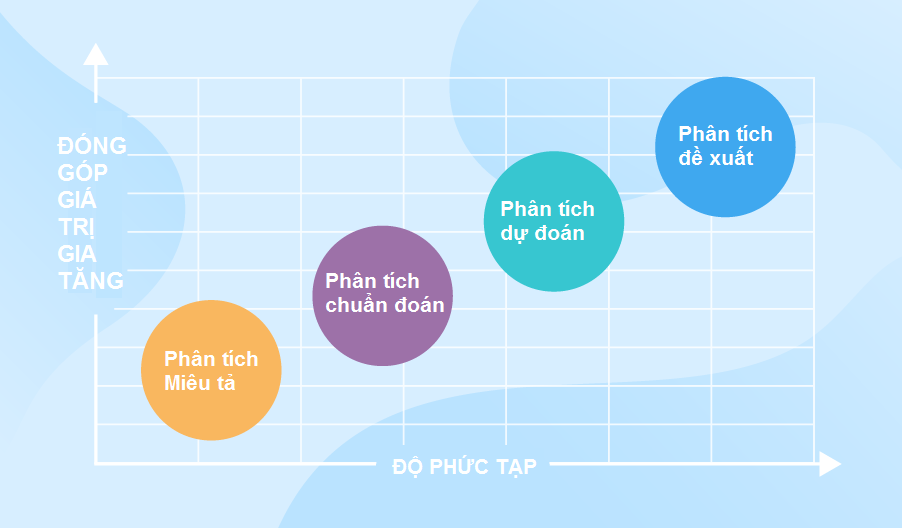
Hình 2 mô tả sự tương quan giữa ba đặc trưng chính của Dữ liệu lớn trong đó các nguồn chính tạo ra Dữ liệu lớn đó là: **Hộp đen dữ liệu** là dữ liệu được tạo ra bởi máy bay, bao gồm máy bay phản lực và trực thăng, dữ liệu này bao gồm thông tin tạo ra bởi giọng nói của phi hành đoàn, các bản thu âm và thông tin về chuyến bay. **Dữ liệu từ các kênh truyền thông xã hội** là dữ liệu được tạo ra và phát triển bởi như các trang web truyền thông xã hội như Twitter, Facebook, Instagram, Pinterest và Google. **Dữ liệu giao dịch chứng khoán** là số liệu từ thị trường chứng khoán đối với quyết định mua và bán cổ phiếu được thực hiện bởi khách hàng. **Dữ liệu giao thông là** dữ liệu này bao gồm sức chứa và các mẫu phương tiện giao thông, độ sẵn sàng và khoảng cách đã đi được của từng phương tiện giao thông. **Dữ liệu được tạo ra từ các công cụ tìm kiếm** và đây cũng là một trong nguồn dữ liệu lớn nhất của Big Data. Công cụ tìm kiếm có cơ sở dữ liệu cực kỳ rộng lớn, nơi họ có thể tìm thấy dữ liệu họ cần. **Dữ liệu từ các thiết bị radar, cảm biến**: Dữ liệu được thu thập từ các nguồn khác nhau từ cấp trạm gửi về trung tâm liên tục từng giây, hàng ngày ở các trung tâm sở chỉ huy điều khiển giám sát máy bay, hàng hải, hay các trung tâm sở chỉ huy quân sự,…

**1.2. PHÂN TÍCH DỮ LIỆU LỚN (BIG DATA ANALYTICS)**

Tầm quan trọng của Dữ liệu lớn không nằm ở lượng dữ liệu mà chúng ta có, nó nằm ở việc chúng ta làm gì với những dữ liệu đó. Đối với các doanh nghiệp các lợi ích của phân tích Dữ liệu lớn đem lại như: **Các quyết định nhanh hơn và sáng suốt hơn** do khả năng xử lý và phân tích dữ liệu do sử dụng Dữ liệu lớn có khả năng tạo ra trong thời gian thực có nghĩa là các công ty có thể hành động ngay lập tức để giải quyết vấn đề, điều chỉnh chiến lược hoặc giải mã xu hướng thị trường. **Hoạt động hiệu quả** hiện tại **n**hiều công ty sử dụng phân tích Dữ liệu lớn để có những cái nhìn sâu sắc về các mắc xích cung ứng nội bộ hoặc dịch vụ cho phép họ thực hiện thay đổi và hợp lý hóa các hoạt động dựa trên thông tin cập nhật từng phút. **Giảm chi phí** không chỉ các công ty có thể giảm chi phí bằng cách tăng hiệu quả hoạt động, mà ngày nay các cơ sở hạ tầng phân tích Dữ liệu lớn có chi phí thấp hơn nhiều so với các hệ thống dữ liệu trước đây. **Cải thiện phát triển sản phẩm hoặc dịch vụ** Hiểu biết về thị trường, khách hàng hoặc ngành thời gian thực có thể giúp các công ty xây dựng sản phẩm tuyệt vời tiếp theo hoặc tạo ra các dịch vụ mà khách hàng rất cần. Thông tin này có thể mang lại cho các công ty sự tự tin để có những bước nhảy vọt lớn hơn trong nghiên cứu và phát triển mà không có cùng rủi ro.

Đối với các nhà nước, các tổ chức chính phủ lợi ích của việc ứng dụng phân tích Dữ liệu lớn mang lại đó là: **Giáo dục công**: Bộ Giáo dục có thể ứng dụng dữ liệu lớn để cải thiện phương pháp giảng dạy và học tập của học sinh [3]. Các tổ chức giáo dục đại học áp dụng phân tích Dữ liệu lớn để cải thiện các dịch vụ nhằm giúp việc học tập của học sinh, sinh viên hiệu quả và giữ được các học sinh, sinh viên theo học. **Quy luật kinh tế:** Phân tích dữ liệu lớn giúp tạo ra các mô hình tài chính từ các dữ liệu kinh tế lịch sử để xây dựng các chính sách trong tương lai. Từ đó các tổ chức kinh tế như Ủy ban Chứng khoán và Giao dịch có thể sử dụng những phân tích Dữ liệu lớn này để điều chỉnh hoạt động tài chính, phát hiện các tác nhân xấu và gian lận tài chính [4]. **Bảo vệ môi trường:** Trong hơn hai thập kỷ, NASA và Bộ Năng lượng Hoa Kỳ đã sử dụng phân tích dữ liệu trong các nghiên cứu của họ để dự đoán tốt hơn các kiểu thời tiết [5], cháy rừng và các tác động rủi ro môi trường khác.

Trong kỹ thuật phân tích dữ liệu lớn có bốn loại thành phần chính: **Phân tích dữ liệu miêu tả (descriptive)**: mô tả những gì xảy ra ở quá khứ dựa trên những báo cáo bằng biểu đồ, hình ảnh, nhưng không giải thích tại sao và tiên đoán những gì sẽ xảy ra. **Phân tích dữ liệu chuẩn đoán (diagnostic)**: giải thích những gì đã xảy ra ở quá khứ. **Phân tích dữ liệu dự đoán (predictive)**: tiên đoán những gì có thể diễn ra. **Phân tích dữ liệu đề xuất (prescriptive)**: hệ thống quyết định phân tích và tiên đoán dữ liệu và đưa ra đề xuất hành động thích hợp và dự đoán những kết quả có thể xảy ra dựa trên những kết quả phân tích được. Mức độ đóng góp giá trị cũng như độ phức tạp tính toán của bốn loại hình phân tích dữ liệu lớn này được mô ta như Hình 3 dưới đây



**Hình 3: Bốn loại phân tích dữ liệu lớn và độ phức tạp tính toán**

1. **ỨNG DỤNG PHÂN TÍCH DỮ LIỆU LỚN VÀO CÁC LĨNH VỰC CUỘC SỐNG**

Việc ứng dụng Dữ liệu lớn vào các hoạt động nghiên cứu, phát triển sản phẩm, dịch vụ không còn là hoạt động quá xa lạ đối với các doanh nghiệp trên thế giới. Roll-Royces sử dụng công nghệ này trong cả ba khâu: Thiết kế, sản xuất và hỗ trợ bán hàng. Netflix cũng áp dụng thành công Phân tích Dữ liệu lớn khi đưa ra khuyến nghị cho từng người dùng theo thời gian thực, mang về 75% lượng truy cập cho hãng. Các công ty sử dụng Dữ liệu lớn để tìm hiểu hành vi tiêu dùng của khách hàng từ đó đưa ra các đề xuất mua sắm cá nhân hóa cho từng đối tượng dựa trên thông tin thu thập được từ họ (Ebay, Facebook, Google…). Các cơ sở nghiên cứu khoa học sử dụng Dữ liệu lớn để tìm ra những khám phá khoa học mới cho nhân loại, ví dụ như xây dựng bản đồ gene của con người, hay tiêu biểu như các nghiên cứu tại trong máy gia tốc Hadron của Tổ chức Nghiên cứu Nguyên tử Châu Âu, với luồng dữ liệu sẽ trở nên vô cùng lớn, có thể đạt đến 150 triệu petabyte mỗi năm, gấp 200 lần so với tất cả các nguồn dữ liệu khác trên thế giới gộp lại. Dưới đây bài viết điểm qua các ứng dụng chính của phân tích dữ liệu lớn vào các lĩnh vực khác nhau trong đời sống:

**Phân tích mục tiêu của khách hàng:** Đây là một trong những lĩnh vực lớn nhất của việc ứng dụng công nghệ Dữ liệu lớn, nó giúp các cơ quan, tổ chức hiểu rõ hơn về nhu cầu khách hàng cũng như hành vi và sở thích của họ để có bức tranh hoàn chỉnh, toàn diện hơn về khách hàng. Ví dụ: nhà bán lẻ Target của Hoa Kỳ có thể dự đoán gần như chính xác thời điểm một khách hàng của họ mong muốn sinh em bé, để giới thiệu, cung cấp các sản phẩm phù hợp, các công ty viễn thông có thể dự đoán tốt hơn việc thay đổi mong muốn của khách hàng. Ngay cả các chiến dịch bầu cử của chính phủ cũng có thể được tối ưu hóa bằng cách sử dụng phân tích Dữ liệu lớn như chiến thắng của cựu tổng thống Barack Obama vào năm 2012 là nhờ khả năng vượt trội của các công nghệ Dữ liệu lớn.

**Hiểu và tối ưu hóa sản phẩm, quy trình kinh doanh:** Dữ liệu lớn ngày càng được sử dụng nhiều trong việc tối ưu hóa quy trình kinh doanh. Các nhà bán lẻ có thể tối ưu hóa cổ phiếu của họ dựa trên các dự đoán được tạo ra từ dữ liệu truyền thông xã hội, xu hướng tìm kiếm trên web. Các nhà bán lẻ sử dụng rất nhiều phân tích dữ liệu lớn để tối ưu hóa chuỗi cung ứng, cũng như tuyến đường phân phối. Ví dụ, Công ty Sociometric Solutions (Hoa Kỳ) đặt cảm biến vào tấm huy hiệu tên nhân viên để có thể phát hiện các động lực xã hội tại nơi làm việc. Các cảm biến sẽ báo cáo về cách nhân viên di chuyển xung quanh nơi làm việc, nói chuyện. Mọi người có thể gắn thẻ RFID lên điện thoại, chìa khóa hoặc kính của bạn, để xác định vị trí của chúng, tránh bị mất.

**Cải thiện an ninh và thực thi pháp luật:** Phân tích Dữ liệu lớn được áp dụng rất nhiều trong việc cải thiện bảo mật và cho phép thực thi pháp luật. Ví dụ Cơ quan An ninh quốc gia Hoa Kỳ (NSA) sử dụng các phân tích dữ liệu lớn dự đoán sớm về các cuộc khủng bố; một số đơn vị sử dụng công nghệ Big Data để phát hiện và ngăn chặn các cuộc tấn công trên mạng; lực lượng cảnh sát sử dụng các công cụ Big Data để bắt tội phạm và dự đoán hoạt động của các tổ chức tội phạm; các công ty tín dụng sử dụng Big Data để phát hiện các giao dịch gian lận. Thông qua phân tích dữ liệu có thể giúp các tổ chức doanh nghiệp xác định các hoạt động khả nghi chỉ ra hành vi gian lận để giải quyết.

**Xây dựng thành phố thông minh:** Một số thành phố hiện đang thử nghiệm các mô hình Dữ liệu lớn với mục đích chuyển đổi sang thành phố thông minh, nơi mà tất cả các quy trình, cơ sở hạ tầng giao thông và tiện ích đều được kết nối với nhau. Thành phố Long Beach, California (Hoa Kỳ) đang sử dụng đồng hồ nước thông minh để phát hiện mất trộm nước theo thời gian thực, giúp một số chủ nhà cắt giảm tới 80% lượng nước sử dụng. Los Angeles sử dụng dữ liệu từ cảm biến và 4.500 camera lắp đặt trên các đường phố để kiểm soát đèn giao thông, giúp giảm tắc nghẽn giao thông lên đến 16%.

**Hoạt động tài chính - ngân hàng:** Sự gia tăng của hoạt động giao dịch tần suất cao (High frequency trading - HFT) theo thời gian thực đã khiến các nhà quản lý trong lĩnh vực chứng khoán và thị trường chứng khoán phải làm việc rất nhiều để theo kịp sự mở rộng của dữ liệu cấu trúc và phi cấu trúc. Bằng cách sử dụng những công cụ và kỹ thuật mới của Big Data, những công ty hoạt động trong lĩnh vực thị trường vốn có thể cải thiện hoạt động giám sát giao dịch và phát hiện được những hành vi giao dịch bất thường và hành vi phạm pháp trước khi chúng được thực hiện.

**Quốc phòng và An ninh:** Đối với các ngành quân sự như Không quân, Quân đội, Hải quân và Thủy quân lục chiến dữ liệu là sự khởi đầu của một hoạt động quân sự. Cụ thể việc sử dụng dữ liệu lớn để giám sát, xử lý thông tin tình báo, theo dõi các mục tiêu, tham gia, đánh giá thiệt hại trước và sau tấn công hay công tác hậu cần quân sự và quản lý nguồn nhân lực. Một trong những nhiệm vụ chính là sự chuẩn bị đầy đủ của vật tư khi hoạt động trên các chiến trường, khả năng phản ứng quân đội cho các tình huống khác nhau. Thách thức lớn nhất ở đây là xử lý dữ liệu khổng lồ trong hệ thống cơ sở dữ liệu của quân đội, tất cả các quyết định phải đưa ra tức thời theo thời gian thực trong phân tích tối ưu chi phí, dự đoán các khả năng rủi ro, bất ngờ xảy ra hay một khía cạnh quan trọng khác là phân tích các hoạt động của kẻ địch để phát hiện các quy luất từ đó có thể giúp xác định các hoạt động bất thường trong tương lai, một trong những ứng dụng dự đoán chính trong phòng thủ. Một vài ví dụ dẫn chứng: Hiện tại Quân đội Mỹ đã sử dụng hệ thống giám sát mặt đất ARGUS gắn trên máy bay không người lái UAV thu thập hơn 40 gigabyte thông tin mỗi giây (ARGUS gồm 365 FPA (Focal Plane Array) chồng chéo, chụp ảnh một khu vực rộng liên tục ở tần số 10 Hz). Hay trong lĩnh vực xử lý ảnh vệ tinh, một vệ tinh thu thập trung bình khoảng 2 Petabyte dữ liệu hàng năm. Các công ty sẽ duy trì 40 PetaByte trên ổ cứng và 20 PB trên các băng (tape) và sẽ có một bản sao ở trung tâm dữ liệu dự phòng.

1. **ỨNG DỤNG PHÂN TÍCH DỮ LIỆU LỚN TẠI TRUNG TÂM CHỈ HUY ĐIỀU KHIỂN**

Thực tế dữ liệu thu thập tại các Trung tâm chỉ huy điều khiển ở Việt nam: Hiện tại, hằng ngày Sở chỉ huy Hệ thống Thông tin và Tác chiến điện tử vẫn phải lưu 40gb/ngày (tương đương 15TB mỗi năm) dữ liệu chỉ riêng về các thông tin tín hiệu từ các radar, âm thanh. Hay tại Sở chỉ huy Phòng không Không quân, dữ liệu hơn 200 quỹ đạo/1s, và kích thước khoảng 3gb/ngày (tương đương 1Tb mỗi năm). Dựa trên những phân tích trên đây, với những cơ sở hạ tầng và công nghệ như hiện tại không cho phép chúng ta có thể lưu trữ và ứng dụng các phân tích dữ liệu từ 10 năm trở lên, hoặc khi số lượng mục tiêu theo dõi tăng lên rất nhiều thì các thuật toán bắt bám mục tiêu truyền thống không thể đáp ứng tốc độ xử lý. Điều đó dẫn tới việc cần thiết nhìn trước phải đầu tư xây dựng và phát triển một hệ thống Dữ liệu lớn để có thể đáp ứng được những yêu cầu sắp tới.

**3.1 PHƯƠNG HƯỚNG XÂY DỰNG VÀ PHÁT TRIỂN**

**Xây dựng cơ sở hạ tầng hỗ trợ Big Data**: Muốn sử dụng và khai thác Big Data một cách hiệu quả đối với doanh nghiệp thương mại điện tử 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. Điều này có thể được hiện thực hóa bởi các trung tâm dữ liệu (data center) lớn, các dịch vụ điện toán đám mây cũng góp phần giải quyết bài toán này. Ngoài ra, còn cần xây dựng cơ sở hạ tầng bảo mật, bao gồm việc mã hóa dữ liệu, xác thực người dùng và các quyền truy cập, giám sát hệ thống, xây dựng tường lửa, quản lý doanh nghiệp và các sản phẩm khác để bảo vệ hệ thống và dữ liệu. Ngoài ra, để có thể khai thác hiệu quả Big Data, cần có những công nghệ và công cụ hỗ trợ cần thiết như là Hệ sinh thái Hadoop, Cơ sở dữ liệu theo hướng NoSQL, Cơ sở dữ liệu bộ nhớ trong,...

**Phân tích dữ liệu Big Data Khi đã có dữ liệu**: điều quan trọng nhất là phải phân tích những dữ liệu đó để chúng thực sự mang lại lợi ích như tăng doanh thu, cải thiện dịch vụ khách hàng, nâng cao hiệu suất và tăng sức cạnh tranh tổng thể. Việc phân tích dữ liệu bao gồm việc kiểm tra các tập dữ liệu, từ đó đưa thông tin chi tiết hoặc rút ra kết luận về xu hướng và dự đoán về hoạt động trong tương lai. Nhờ đó, dữ liệu có thể áp dụng vào công việc như lựa chọn thời gian, địa điểm để quảng cáo cho sản phẩm và dịch vụ, điều rất có ích cho các công ty cần những chiến dịch quảng bá sản phẩm theo thời điểm. Phân tích dữ liệu bao gồm: Phân tích dữ liệu thăm dò (để xác định các mẫu và mối quan hệ trong dữ liệu); Phân tích dữ liệu xác nhận (áp dụng các kỹ thuật thống kê để tìm hiểu xem giả định về một tập dữ liệu cụ thể có đúng không); Phân tích dữ liệu định lượng (so sánh thống kê); Phân tích dữ liệu định tính (tập trung vào dữ liệu phi số như video, hình ảnh và văn bản).

**3.2 HỢP TÁC CHUYỂN GIAO CÔNG NGHỆ**

Như đã nêu ở trên để xây dựng một hệ thống dữ liệu lớn tại trung tâm chỉ huy điều khiển phục vụ cho mảng Quân sự và Dân sự chúng ta cần tập trung vào hai khía cạnh chính. Một là xây dựng cơ sở hạ tầng: Điều này có thể kết hợp bên cạnh việc đầu tư mới thì cũng tận dụng những cơ sở hạ tầng hiện có, sử dụng các công nghệ ảo hóa, tận dụng tài nguyên điện toán đám mây,… tuy nhiên, phải đi kèm với đó là đáp ứng chính xác các nguyên tắc bảo mật nhằm trách việc rò rỉ, lộ thông tin. Bên cạnh đó, sẵn sàng chuẩn bị cho việc đầu từ chuyển giao các công nghệ và công cụ hỗ trợ nền tảng, sẵn sàng làm chủ công nghệ cốt lõi quan trọng như : **Nền tảng hệ thống lữ trữ file phân tán (Hadoop Distributed File System)** cung cấp quyền truy cập thông lượng cao vào dữ liệu ứng dụng; **Hadoop YARN** một nền tảng cho việc lập kế hoạch công việc và quản lý tài nguyên cụm; **Hadoop MapReduce** một hệ thống dựa trên YARN để xử lý song song các tập dữ liệu lớn; **Apache Spark** là một nền tảng hệ thống tính toán mã nguồn mở phục vụ như một công cụ để xử lý Big Data trong Hadoop. Spark đã trở thành một trong những nền tảng xử lý phân tán dữ liệu quan trọng và có thể được triển khai theo nhiều cách khác nhau. **Cơ sở dữ liệu NoSQL** lưu trữ và quản lý dữ liệu theo cách linh hoạt và xử lý tốc độ cao. Không giống như cơ sở dữ liệu SQL, nhiều cơ sở dữ liệu NoSQL có thể được thu nhỏ theo chiều ngang trên hàng trăm hoặc hàng nghìn máy chủ.

**TÀI LIỆU THAM KHẢO**

1. IDC report, https://www.idc.com/getdoc.jsp?containerId=US44803719
2. Forbes report, https://images.forbes.com/forbesinsights/StudyPDFs/RocketFuel\_BigData\_REPORT.pdf
3. Marie Bienkowski, Mingyu Feng, Barbara Means. *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief*, U.S. Department of Education Office of Educational Technology, <https://tech.ed.gov/wp-content/uploads/2014/03/edm-la-brief.pdf>.
4. https://www.sec.gov/news/speech/bauguess-big-data-ai.
5. https://www.esri.com/about/newsroom/arcuser/forecasting-weather-with-big-data-in-the-cloud/.

**VỀ TÁC GIẢ**

|  |  |
| --- | --- |
|  | Bùi Tăng Bảo Ngọc sinh năm 1983, nhận bằng Kỹ sư lĩnh vực Toán Tin Ứng dụng, tại trường Đại học Bách Khoa Hà Nội, năm 2006, bằng Tiến sĩ tại Đức 2010-2014 và chuyên gia sau tiến sĩ tại Mỹ 2014-2017. Hiện nay, Ngọc là kỹ sư chính nghiên cứu thuật toán tại Trung tâm Chỉ huy và Điều khiển. Lĩnh vực nghiên cứu yêu thích: Học máy và xử lý ngôn ngữ tự nhiên.  Bùi Tăng Bảo Ngọc nhận giải thưởng và học bổng của các nước Nhật, Đức, Mỹ trong quá trình học và nghiên cứu đồng thời tham gia giảng dạy tại Đại học Bách khoa Hà nội cũng như một số trường đại học tại các nước Đức, Ý, Mỹ. |