

# Big data: TÁC ĐỘNG VÀ XU HƯỚNG PHÁT TRIỂN

Lò Thị Phương Nhung, Nguyễn Mai Phương

Viện Thông tin khoa học, Học viện Chính trị quốc gia Hồ Chí Minh

Cách mạng công nghiệp lần thứ tư đang làm thay đổi cách thức sản xuất, chế tạo và hình thành kết nối của hàng tỷ người trên thế giới thông qua các thiết bị di động và khả năng tiếp cận với cơ sở dữ liệu lớn (big data). Trong thế giới toàn cầu hóa ngày nay, các hoạt động hàng ngày của chúng ta đã tạo ra các dữ liệu một cách liên tục. Nhờ những công nghệ đột phá của cách mạng công nghiệp lần thứ tư, lượng dữ liệu này ngày một khổng lồ, liên tục thay đổi và có những tác động trở lại đến hầu hết các lĩnh vực và mọi mặt của đời sống xã hội.

## Khái quát chung về big data

Big data đã và đang là một trong những vấn đề trung tâm, nhận được nhiều sự quan tâm trong cuộc cách mạng công nghiệp lần thứ tư. Big data chính là cốt lõi để sử dụng, phát triển internet vạn vật (IoT) và trí tuệ nhân tạo (AI). Theo dự báo, cách mạng công nghiệp lần thứ tư sẽ tạo ra một lượng lớn dữ liệu, dự kiến đến năm 2020, lượng dữ liệu sẽ tăng gấp 50 lần hiện nay [1]. Thông qua thu thập, phân tích và xử lý lượng big data này sẽ tạo ra những tri thức mới, hỗ trợ việc đưa ra quyết định của các chủ thể trên thế giới (doanh nghiệp, chính phủ, người dân).

Thuật ngữ big data là để chỉ một tập hợp dữ liệu rất lớn, không đồng nhất, phức tạp đến mức những công cụ, ứng dụng xử lý dữ liệu truyền thống không thể nào đảm đương được. Sự tăng trưởng của big data thường được mô tả với 3 chữ V: Volume - Tăng về lượng, Velocity - Tăng về vận tốc và Variety - Tăng về chủng loại [2]. Trong bối cảnh cuộc cách mạng công nghiệp lần thứ tư và thế giới dịch vụ thông minh (smart service world), big data có thể được đề cập đến dưới dạng 6C: Connection - Kết nối (cảm biến và mạng); Cloud - Đám mây (tính toán và theo yêu cầu);

Cyber - Không gian ảo (mô hình và bộ nhớ); Content - Nội dung/Bối cảnh (ý nghĩa và tương quan); Community - Cộng đồng (chia sẻ và cộng tác); Custom - Tùy chỉnh (cá nhân hoá và giá trị) [3].

Kích cỡ của big data đang tăng lên từng ngày, trong khoảng vài chục terabyte cho đến nhiều petabyte (1 petabyte = 1.024 terabyte) chỉ cho một tập hợp dữ liệu. Nếu như các dữ liệu truyền thống có thể lưu trữ trên các thiết bị đĩa mềm, đĩa cứng thì cách mạng công nghiệp lần thứ tư hiện nay hướng đến sử dụng công nghệ “đám mây” để lưu trữ big data này. Mặc dù lượng thông tin lưu trữ ngày càng lớn, khối lượng dữ liệu gia tăng rất nhanh, nhưng chi phí lưu trữ thì ngày càng giảm. Ngày nay, lưu trữ 1 GB có chi phí trung bình ít hơn 0,03 USD/năm, so với hơn 10.000 USD vào thời điểm cách đây 20 năm. Chỉ đơn giản với một thiết bị như một máy tính bảng, chúng ta có thể đọc sách, “lướt” web và trao đổi thông tin liên lạc, sở hữu khả năng xử lý tương đương với 5.000 máy tính để bàn của 30 năm trước.

Big data cần đến các kỹ thuật khai thác thông tin rất đặc biệt do tính chất khổng lồ và phức tạp của nó. Năm 2011, Tập đoàn Phân tích McKinsey (Mỹ) đề xuất những

công nghệ có thể dùng với big data bao gồm crowdsourcing (tận dụng nguồn lực từ nhiều thiết bị điện toán trên toàn cầu để cùng nhau xử lý dữ liệu), các thuật toán về gen và di truyền, những biện pháp machine learning (các hệ thống có khả năng học hỏi từ dữ liệu, một nhánh của trí tuệ nhân tạo), xử lý ngôn ngữ tự nhiên (giống như Siri hay Google Voice Search, nhưng cao cấp hơn), xử lý tín hiệu, mô phỏng, phân tích chuỗi thời gian, mô hình hóa, kết hợp các server mạnh lại với nhau... Các cơ sở dữ liệu hỗ trợ xử lý dữ liệu song song, ứng dụng dựa trên hoạt động tìm kiếm, file system dạng rời rạc, các hệ thống điện toán đám mây (bao gồm ứng dụng, nguồn lực tính toán cũng như không gian lưu trữ) và internet cũng là công cụ đắc lực phục vụ cho công tác nghiên cứu và trích xuất thông tin từ big data.

## Tác động và xu hướng phát triển của big data

### Nhu cầu của con người với big data

Erik Swan - Đồng sáng lập kiêm Giám đốc công nghệ của Công ty Spunk (Mỹ) dự đoán rằng, sự thay đổi nhất trong big data chính là thái độ của mọi người đối với nó [4]. Việc sử dụng big data sẽ dần phổ biến

hơn với mọi người, từ những nông dân cho đến kỹ sư. Mọi người sẽ mặc định sử dụng dữ liệu để phân tích mọi thứ trong vòng 10 năm tới. Tất nhiên là kỹ thuật và công nghệ cũng cần phải phát triển theo thì điều này mới có thể trở thành hiện thực. Bên cạnh đó, hoàn cảnh phát sinh dữ liệu sẽ trở nên quan trọng hơn. Ankur Jain, nhà sáng lập và CEO của Humin (một công ty công nghệ của Mỹ) cho rằng, chúng ta sẽ bắt đầu định tuyến dữ liệu vào các đối tượng, sự vật, sự việc trong đời thực và chuyện đó giúp chúng ta xử lý công việc tốt hơn.

Như vậy, big data giúp các công ty phân tích và dự báo tốt nhu cầu của khách hàng. Với big data, các công ty có thể dự báo sản phẩm nào có nhu cầu cao để thay đổi chiến lược tiếp thị của họ, lực lượng bán hàng của họ, và các quy trình chế tạo để tạo ra chúng một cách nhanh chóng. Vì doanh nghiệp của họ thay đổi, đối thủ cạnh tranh sẽ không có khả năng đáp ứng đủ nhanh để cạnh tranh. Bằng việc phân tích big data, Tập đoàn Samsung đã dự báo tính năng nào mà khách hàng muốn có trong điện thoại di động để phát triển điện thoại của họ (loạt Galaxy), trong khi Motorola, Sony, Nokia vẫn còn phụ thuộc vào ý kiến của người quản lý bán hàng của họ. Nhờ vậy mà Samsung đã thu tóm phần lớn thị trường di động, thậm chí vượt Apple, trong khi Nokia, Motorola và Sony mất hầu hết thị trường và có thể phải ngừng kinh doanh mặt hàng này.

### Công cụ cạnh tranh chiến lược

Trong thị trường toàn cầu được “dẫn lái” bởi công nghệ, mọi nhà quản lý đều cần hiểu sức mạnh của công nghệ thông tin nói chung và big data nói riêng. Big data là nhu cầu đang tăng trưởng lớn đến nỗi Software AG, Oracle, IBM, Microsoft, SAP, EMC, HP và Dell

đã chi hơn 15 tỷ USD cho các công ty chuyên về quản lý và phân tích dữ liệu. Năm 2016, ngành công nghiệp big data có giá trị hơn 190 tỷ USD và đang tăng nhanh với tốc độ 10% mỗi năm, nhanh gấp đôi so với tổng ngành phần mềm nói chung [4]. Theo các nghiên cứu về công nghiệp năm 2015, big data có thể giúp cải tiến việc ra quyết định quản lý và có tác động lớn đến khả năng cạnh tranh và ngăn chặn rủi ro của công ty. Gần 91% số công ty của Mỹ được điều tra đều báo cáo rằng, việc dùng tốt hơn big data sẽ dẫn tới việc ra quyết định tốt hơn và giúp làm tăng tính cạnh tranh toàn cầu của công ty; 65% đồng ý rằng, công nghệ big data sẽ là quan trọng để nhận diện rủi ro và bảo vệ công ty khỏi bị tấn công “cyber”.

### Vấn đề việc làm

Theo dự báo, riêng ở Mỹ, big data sẽ tạo ra trên 2 triệu việc làm mới. Tuy nhiên, cung cấp của toàn thế giới cho lĩnh vực này chỉ có thể tạo ra được không đầy 1/3 nhu cầu [5]. Trong cách mạng công nghiệp lần thứ tư, chúng ta sẽ còn tiếp tục chứng kiến sự tăng trưởng của big data. Dữ liệu từ IoT thực chất được thu thập từ một mạng lưới rất nhiều cảm biến và thiết bị điện tử, sẽ là một trong những nguồn quan trọng của big data. Lượng dữ liệu khổng lồ này có thể cho các nhà nghiên cứu biết được hành vi tiêu dùng của khách hàng, từ đó tinh chỉnh những thiết bị IoT cho phù hợp hơn, bắt chúng phục vụ đời sống hàng ngày của chúng ta một cách hiệu quả hơn.

\*  
\* \*

Có thể nói, big data đang ngày càng được ứng dụng rộng rãi trong các cơ quan, tổ chức, trong các lĩnh vực khác nhau của đời sống kinh tế - xã hội. Các phân tích trên big data góp phần cải tiến và tối ưu hóa

quá trình ra quyết định, giảm thiểu rủi ro, tạo ra những giá trị gia tăng cho doanh nghiệp, hỗ trợ tổ chức trong việc quản lý các hoạt động hàng ngày cũng như ra quyết định. Với tác động và xu hướng phát triển chung của big data, các cơ quan quản lý và doanh nghiệp của Việt Nam không thể đứng ngoài cuộc. Đối với cơ quan quản lý, việc phân tích big data sẽ giúp các quyết sách của họ phù hợp với thực tiễn và đối tượng quản lý; góp phần xây dựng Chính phủ liêm chính, kiến tạo, hành động và phục vụ. Đối với doanh nghiệp, việc phân tích big data và ứng dụng các công nghệ khác có liên quan sẽ giúp doanh nghiệp hiểu biết đầy đủ về nhu cầu của khách hàng, từ đó đưa ra các quyết định đúng đắn, tạo lợi thế cạnh tranh...

Tuy nhiên, việc ứng dụng, khai thác big data đã, đang và sẽ tạo ra thách thức lớn về sở hữu trí tuệ khi tham gia thương mại quốc tế, về nguồn nhân lực, cơ sở hạ tầng... Việt Nam cần thực sự quan tâm và đầu tư vào việc bảo hộ cơ sở dữ liệu nói chung và big data nói riêng, về việc đào tạo nguồn nhân lực, phát triển cơ sở hạ tầng... để tạo ra môi trường thuận lợi, lành mạnh, phù hợp với quốc tế, góp phần thúc đẩy phát triển mạnh mẽ thương mại điện tử cũng như các lĩnh vực kinh tế - xã hội khác ✍

### TÀI LIỆU THAM KHẢO

- [1] Marek Obitko, *Industry 4.0 and Big Data*, [http://www.stech.cz/Portals/0/Konference/2015/03%20Industry-/PDF/03\\_obitko.pdf](http://www.stech.cz/Portals/0/Konference/2015/03%20Industry-/PDF/03_obitko.pdf).
- [2] Thomas Roca (2015), “Le Big Data au service des plus pauvres”, *Idées pour le développement*.
- [3] Nilesh Jaiswal (2015), “Handling of Big Data by using Big Table”, *International Journal of Computer Science and Information Technology*.
- [4] John Vu, *Carnegie Mellon University*, <http://science-technology.vn>.
- [5] S. David (2016), “Economic Forecast in the age of Big Data”, *Working paper*.