

## VAI TRÒ CỦA BIG DATA TRONG LĨNH VỰC KẾ TOÁN

HỒ THỊ VÂN ANH, PHẠM TÚ ANH

*Khoa Kế toán – Kiểm toán, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh  
hothivananh@iuh.edu.vn*

**Tóm tắt.** Big data có ý nghĩa ngày càng quan trọng đối với lĩnh vực kinh tế nói chung và lĩnh vực kế toán nói riêng vì thông qua Big data có thể giúp cho các hoạt động kế toán được cải thiện. Bài viết này nghiên cứu tổng quan về việc xem xét sự cần thiết sử dụng Big data trong kế toán, ngoài việc cung cấp thông tin kế toán có chất lượng phục vụ cho việc giám sát, công khai tài chính và phân tích, tham mưu cho nhà quản trị, thì còn nhằm giúp cho học giả kế toán, người hành nghề kế toán và sinh viên kế toán có cơ hội tìm hiểu sâu hơn về những lợi ích tiềm năng của Big data, cũng như những thách thức và trở ngại khi sử dụng dạng dữ liệu này.

**Từ khóa.** Big data, ứng dụng Big data trong kế toán.

### THE ROLE OF BIG DATA IN THE FIELD OF ACCOUNTING

**Abstract.** Big data plays a more important role in economic area in general and accounting area in particular due to the fact that Big data improves the quality of accounting. This paper discusses the necessity of using Big data in accounting not only in terms of providing high quality of information needed for controlling and analysis purposes and presenting financial results but also regarding providing insights to academics, accounting professionals and students about potential benefits of Big data. Additionally, this paper presents the challenges of deploying Big data.

**Keywords.** Big data, applying Big data in accounting.

#### 1 GIỚI THIỆU

Ghi nhận kế toán về bản chất là tập hợp các dữ liệu tài chính trong quá khứ được tổng hợp và sử dụng để lập báo cáo tài chính cho người dùng nội bộ (nhà quản lý) và người dùng bên ngoài (nhà đầu tư và chủ nợ). Trước đây việc ghi nhận các sự kiện kinh tế phát sinh vào chứng từ, sổ sách và báo cáo kế toán được thực hiện thủ công bởi người làm kế toán thì bây giờ chúng gần như được số hóa. Ví dụ, năm 2000 có khoảng 25% tất cả thông tin được lưu trữ là dưới dạng công nghệ số thì hiện nay con số này là hơn 98% (Cukier và Mayer-Schonberger 2013).

Hiện nay, các hệ thống cơ sở dữ liệu hiện hành đều có khả năng thu thập và đánh giá cả hai loại dữ liệu tài chính và phi tài chính. Tuy nhiên, Big data cho thấy một sự tối ưu hơn liên quan đến việc cung cấp các bộ dữ liệu đa dạng, không lồ và các phân tích chi tiết. Điều này hàm ý rằng, tầm quan trọng ngày càng tăng của Big data sẽ ảnh hưởng lớn đến nghề kế toán. Cụ thể, sự ưu việt của Big data sẽ được phản ánh qua cách tích lũy và ghi nhận dữ liệu, cách sử dụng dữ liệu để đạt được các mục tiêu của tổ chức, cách xử lý và tập hợp các nghiệp vụ kinh tế phát sinh để lập báo cáo tài chính.

Cùng với xu hướng kỹ nguyên công nghệ số, các tổ chức hiện nay đã thu thập được nhiều dữ liệu hơn so với vài thập kỷ trước (Syed, Gillela và Venugopal, 2013). Phần lớn các dữ liệu này là phi cấu trúc và thường bắt nguồn từ các thiết bị cảm biến và phương tiện truyền thông xã hội. Các tổ chức tận dụng những dữ liệu này để cải thiện hiệu quả kinh doanh và lợi nhuận. Thực tế, nhiều tổ chức đã kết hợp phân tích dữ liệu với phân tích kinh doanh nhằm làm tăng năng suất (Brynjolfsson, Hammerbacher và Stevens, 2011). Big data có thể được xem như một tài sản tương tự như thương hiệu của công ty (Brown, Chui và Manyika, 2011). Do đó, các hoạt động đánh giá Big data đang nhanh chóng trở thành yếu tố quan trọng trong việc thiết lập và duy trì lợi thế cạnh tranh cho chính các công ty (Bughin, Livingston và Marwaha, 2011).

Thu thập và phân tích Big data đang được nhân rộng theo cấp số nhân. Hiện tượng nhân rộng này có ý nghĩa đối với sự phát triển của các hoạt động kế toán. Trong các phần sau, nghiên cứu thảo luận về khái niệm về Big data, ý nghĩa của Big data, tại sao Big data là vấn đề cần được quan tâm?, việc sử dụng Big data vào kế toán, trí tuệ kinh doanh (BI - Business intelligence) và phân tích dữ liệu, hạn chế và rủi ro khi sử dụng Big data, kỳ vọng về tương lai Big data, cuối cùng là kết luận.

### Khái niệm Big data

Theo Wikipedia thì: “Big data là thuật ngữ đề cập đến tập hợp các tập dữ liệu được thu thập quá lớn và phức tạp đến nỗi khó xử lý khi sử dụng các công cụ quản lý cơ sở dữ liệu bằng tay hoặc các ứng dụng xử lý dữ liệu truyền thống. Các thách thức bao gồm thu thập, lựa chọn, lưu trữ, tìm kiếm, chia sẻ, chuyển giao, phân tích và trực quan hóa. Xu hướng tập hợp dữ liệu lớn hơn xuất phát là do thông tin bổ sung từ việc phân tích các bộ dữ liệu riêng lẻ có liên quan với nhau, tuy nhiên về bản chất thì tổng khối lượng dữ liệu là như nhau trong khi trước đây chúng ta lại tách các tập dữ liệu này thành các phần nhỏ hơn để phân tích thì bây giờ Big data gộp tất cả chúng lại với nhau, từ đó cho phép tìm ra mối tương quan để phát hiện xu hướng kinh doanh, xác định chất lượng nghiên cứu, ngăn ngừa bệnh tật, liên kết các trích dẫn pháp lý, chống tội phạm và xác định các điều kiện giao thông đường bộ với thời gian thực”.

Ngoài định nghĩa của Wikipedia vừa nêu trên thì vẫn còn nhiều định nghĩa khác nhau về Big data. Ví dụ, Alles và Gray (2016) xem xét các định nghĩa khác nhau về Big data từ nhiều trang web khác nhau thấy rằng, các định nghĩa của Big data tập trung vào hai hướng (i) đặc điểm của Big data, (ii) nội dung của Big data.

Về mặt đặc điểm, Big data được định nghĩa qua 4 đặc tính cụ thể như khối lượng, tốc độ, tính đa dạng và tính xác thực (gọi là 4Vs). Khối lượng đề cập đến tổng dung lượng dữ liệu có trong bộ dữ liệu Big Data. Tốc độ là tần suất dữ liệu thay đổi. Sự đa dạng là phạm vi rộng lớn của dữ liệu mà các tổ chức đang thu thập. Tính xác thực liên quan đến tính đầy đủ của dữ liệu.

Về mặt nội dung, Big data được định nghĩa là bộ dữ liệu có thể bao gồm một sự pha trộn giữa dữ liệu tài chính và phi tài chính có cấu trúc truyền thống (NFD), dữ liệu hỗ trợ được thu thập từ nhiều nguồn khác nhau, dữ liệu cảm biến, email, cuộc gọi điện thoại, dữ liệu truyền thông xã hội, blog, cũng như các dữ liệu bên trong và bên ngoài khác.

### Ý nghĩa của Big data

Ý nghĩa của Big data sẽ khác nhau ở từng lĩnh vực khác nhau. Một bộ dữ liệu cụ thể được xem lớn hay không lớn được xác định thông qua việc dữ liệu này có đầy khả năng giới hạn của việc xử lý hệ thống thông tin khi làm việc với những dữ liệu này hay không. Hơn nữa, việc xem xét Big data trong nghiên cứu này chỉ tập trung vào quan điểm của các công ty và nghiên cứu kế toán có liên quan. Do vậy, việc sử dụng thuật ngữ Big data trong nghiên cứu này hàm ý rằng lượng dữ liệu ở mức giới hạn hoặc vượt quá mức giới hạn những gì mà hệ thống thông tin đang sử dụng tại tổ chức có thể lưu trữ và/hoặc xử lý.

Thuật ngữ Big Data được sử dụng từ những năm 1990 và thực sự bùng nổ trong khoảng 10 năm trở lại đây. Theo định nghĩa của Wikipedia - Big data là một thuật ngữ cho việc xử lý một tập hợp dữ liệu rất lớn và phức tạp mà các ứng dụng xử lý dữ liệu truyền thống không xử lý được. Hay nói cách khác, Big data là tập hợp dữ liệu có dung lượng vượt khả năng hoạt động của các ứng dụng và công cụ truyền thống, và nó được mô tả thông qua 4Vs (Dung lượng – “Volume”, tốc độ – “Velocity”, tính đa dạng – “Variety”, và tính xác thực – “Veracity”), cụ thể là: dung lượng lớn, tốc độ các dữ liệu được tạo ra và xử lý cao, tính đa dạng cao và tính xác thực (IBM 2016). Do dung lượng lớn và tốc độ cao, việc áp dụng vào kế toán và kiểm toán liên tục (CA - Continuous Auditing) ngày càng trở nên phù hợp cho phép phân tích tự động và phân tích thời gian thực của dữ liệu (Vasarhelyi và cộng sự, 2010). Tuy nhiên, dung lượng lớn và tốc độ cao cũng có thể tạo ra khoảng cách giữa các phân tích kiểm toán hiện tại và các yêu cầu đối với phân tích Big data. Mặt khác, tính đa dạng cao và tính xác thực có thể đưa ra những thách thức vượt quá khả năng của các hệ thống xử lý dữ liệu hàng loạt hiện nay trong CA.

Từ những lập luận trên, một câu hỏi có thể được đặt ra là liệu một tập dữ liệu có dung lượng lớn có phụ thuộc vào khả năng của hệ thống thông tin không?. Và những khả năng này thường được phân loại theo tiêu thức gì (kích thước lưu trữ, khả năng xử lý, hay bao gồm cả hai)?.

### Kích thước lưu trữ

Kích cỡ Big Data ngày càng lớn hơn và quan trọng hơn trong thế giới hiện đại và đặc biệt là trong nền công nghiệp 4.0 hiện nay, tính đến năm 2012 nó có thể nằm trong khoảng vài chục terabyte cho đến nhiều petabyte (1 petabyte = 1024 terabyte). Nhưng khả năng lưu trữ của các hệ thống thông tin hiện đại lại rất khác nhau. Các thiết bị lưu trữ thế hệ sau nhanh hơn và đáng tin cậy hơn nhiều, nhưng chúng có thể đắt hơn cho mỗi byte lưu trữ. Các máy chủ bộ phận thường kết hợp nhiều ổ cứng thường trong cùng một hệ thống lưu trữ dự phòng (RAID) để tăng độ tin cậy và tốc độ xử lý, và tăng gấp ba lần dung lượng lưu trữ nhưng chúng thì khá đắt. Các máy chủ cấp doanh nghiệp thường dựa vào các hệ thống lưu trữ gắn liền với mạng (NAS - network attached storage) với dung lượng lưu trữ từ hàng chục đến hàng trăm terabyte. Do đó, tùy

thuộc vào loại hệ thống thông tin có sẵn, theo quan điểm dung lượng lưu trữ của các doanh nghiệp, thì "Big Dataset" có thể dao động từ một đến vài trăm Terabyte.

### **Khả năng xử lý**

Từ quan điểm khả năng xử lý, việc sử dụng Big data có sự khác biệt cho từng loại công việc. Nguyên nhân sâu xa của sự khác biệt này là sự đa dạng của các công việc tính toán trong lĩnh vực kế toán. Ví dụ, các công việc đơn giản (so sánh các yếu tố dữ liệu, tính tổng, tính tỷ lệ phần trăm, v.v...) có tỷ lệ tương ứng với kích thước của tập dữ liệu, nên quan điểm "lớn" được hiểu là khả năng xử lý về cơ bản được xác định bởi khả năng lưu trữ của hệ thống nhưng khi ước tính các mô hình phân tích phức tạp như giá trị thị trường của các công cụ tài chính được giao dịch hoặc phân tích mối quan hệ kế toán cấu trúc giữa các tài khoản và các giao dịch có thể yêu cầu khối lượng tính toán tăng cấp số nhân theo quy mô của bộ dữ liệu. Tuy nhiên, khi có sự gia tăng yêu cầu xử lý theo cấp số nhân nhưng chỉ là một đa thức mức độ thấp thì các nỗ lực sử dụng tất cả Big data được lưu trữ có thể dẫn đến sự không hiệu quả trong việc sử dụng bộ dữ liệu Big Data. Như vậy, độ phức tạp trong tính toán theo yêu cầu của công việc kế toán càng lớn thì quan điểm "lớn hơn" của một bộ dữ liệu cụ thể sẽ như thế nào. Các công việc mà kế toán viên có thể gặp phải khi thực hiện mô hình hóa hoặc các ước tính phức tạp trên các bộ dữ liệu chỉ vài trăm megabyte để xử lý công việc có thể cần nhiều năm tính toán. Từ quan điểm khả năng xử lý, một bộ dữ liệu phục vụ cho việc xử lý công việc không chỉ đơn giản là "lớn" mà là "rất lớn". Ví dụ, theo Arora, Barak, Brunnermeier và Ge (2011) có một số vấn đề nhất định trong việc định giá các công cụ tài chính phái sinh có thể khó tính toán, và việc giải quyết các vấn đề đó đòi hỏi một khối lượng thời gian tính toán rất lớn ngay cả đối với các vấn đề chỉ cần lượng dữ liệu tương đối nhỏ. Cách tiếp cận thông thường trong các trường hợp như vậy là thiết kế ra một mẫu dữ liệu mà độ phức tạp tính toán của mô hình hóa có thể chấp nhận được. Mặc dù, lấy mẫu là một cách để xử lý dữ liệu quá lớn, nhưng quá trình lấy mẫu sẽ làm giảm lợi ích vốn có trong dữ liệu hoàn chỉnh. Do đó, các công việc tính toán tạo thành một hệ thống phân cấp tự nhiên, với các công việc đơn giản nhất là sử dụng mọi thứ có thể được lưu trữ và khi độ phức tạp của tính toán tăng lên, tỷ lệ dữ liệu được sử dụng trong phân tích sẽ ngày càng nhỏ dần.

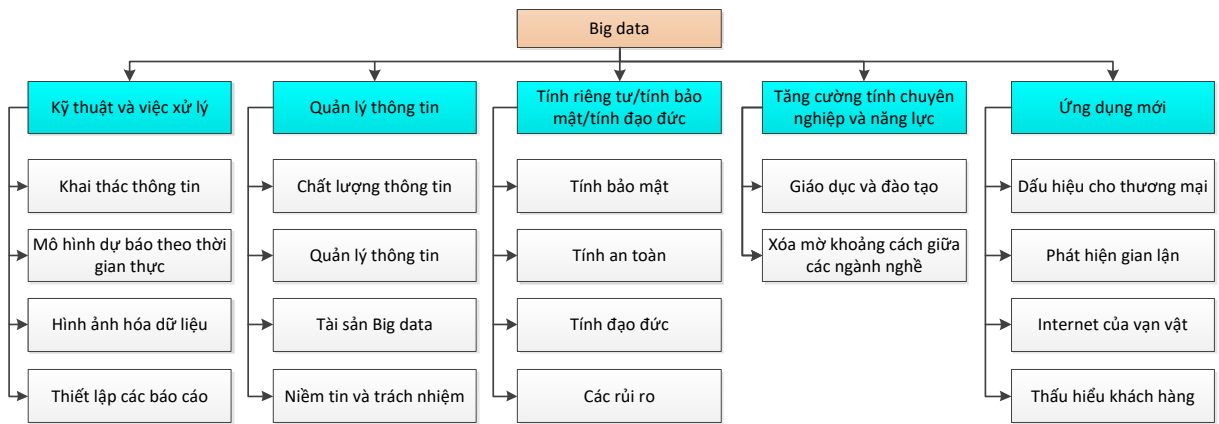
Các nguồn của Big data trong kế toán gồm cả dữ liệu nội bộ và bên ngoài. Dữ liệu doanh nghiệp được tạo ra trong nội bộ sẽ tăng lớn khi doanh nghiệp có khối lượng giao dịch ngày càng lớn. Ví dụ, hồ sơ chi tiết về các cuộc gọi của một công ty điện thoại lớn, công nghệ nhận dạng bằng tần số vô tuyến (RFID - Radio Frequency Identification) cho các sản phẩm hàng tồn kho của nhà bán lẻ hoặc nhà sản xuất lớn, hồ sơ giao dịch chứng khoán của các tổ chức tài chính lớn. Trong chuỗi cung ứng những gì có thể tạo thành dữ liệu nội bộ cho một tổ chức này có thể được xem là dữ liệu ngoài cho một tổ chức khác. Hơn nữa, sự phát triển không ngừng của RFID và công nghệ cảm biến nối mạng đang dẫn đến sự gia tăng nhanh chóng các loại giao dịch được thu thập dẫn đến số lượng doanh nghiệp tạo ra Big data ngày càng nhiều. Chính vì vậy, Oracle và SAP đã phát triển các công cụ có khả năng đặc biệt để xử lý các dữ liệu đó.

## **2 TẠI SAO BIG DATA LÀ VẤN ĐỀ CẦN ĐƯỢC QUAN TÂM TRONG LĨNH VỰC KẾ TOÁN?**

Dữ liệu doanh nghiệp, đặc biệt là các dữ liệu được sử dụng cho mục đích kế toán, đang mở rộng theo cấp số nhân và môi trường này cho thấy các đặc điểm thay đổi linh hoạt nên rất có thể bản thân môi trường sẽ yêu cầu phát triển ra một lý thuyết thông tin nâng cao (Shannon và Weaver, 1949). Sự phát triển và cải tiến cho lý thuyết này sẽ đòi hỏi sự công nhận về bản chất của việc thu thập dữ liệu theo phương thức mới (thủ công so với tự động), khối lượng dữ liệu, hiệu quả của việc tích hợp với kho dữ liệu hiện tại, hiệu quả của việc chuyển đổi dữ liệu thành thông tin, mức độ chi tiết của dữ liệu, loại hoạt động/quyết định được hỗ trợ và các biến khác.

Theo nghiên cứu của Hiệp Hội Kế Toán công chứng Anh Quốc – ACCA (2013), một cuộc khảo sát đã được thực hiện vào năm 2012 – 2013 với các kế toán viên và những nhà phân tích tài chính chuyên nghiệp về sự mở rộng của việc sử dụng Big Data trong lĩnh vực kế toán và tài chính. Kết quả cho thấy 78% những người tham gia trả lời rằng họ mong đợi việc sử dụng Big Data trong 2 lĩnh vực trên sẽ bùng nổ chỉ trong vòng 2 năm nữa. Một nghiên cứu tương tự cũng đã được thực hiện bởi Hiệp Hội Kế Toán công chứng Anh Quốc – ACCA (2013) cho thấy rằng xu hướng công nghệ có tầm ảnh hưởng lớn thứ hai trên khía cạnh về tính tiềm năng trong việc thay đổi mô hình kinh doanh và mô hình của kế toán chính là Big Data

Nhìn chung, vai trò của Big Data trong các khía cạnh của công việc kế toán được thể hiện trong mô hình cấu trúc theo nghiên cứu của Cockcroft và Russell (2018) – một mô hình được xây dựng và tập hợp từ 75 bài báo trên thế giới.



Hình 1: Vai trò của Big Data trong các khía cạnh của công việc kế toán

Nguồn: Cockcrof và Russell (2018)

Một trong những ứng dụng mới của Big data là hình thành dấu hiệu thương mại. Điều này có thể được giải thích thông qua việc ứng dụng Big data trong thị trường chứng khoán. Một lượng lớn các dữ liệu thị trường dưới dạng số hóa (Big data) sẽ được thêm vào trong mô hình giao dịch theo luật định dưới hệ thống giao dịch số học. Phương pháp giao dịch khi đó sẽ được lọc, xử lý, phân tích các dữ liệu thị trường và hình thành dấu hiệu thương mại. Dựa trên các dấu hiệu thương mại này, các công ty mới đưa ra quyết định cho việc thực hiện các hành động của mình (ví dụ như thời điểm nào tham gia thị trường là tốt nhất bất kể lúc đó thị trường đang đi xuống hay đi lên, thực hiện một hợp đồng mua hay bán chứng khoán, v.v) và khi các giao dịch này được thực hiện đó chính là phản ánh của thị trường với dữ liệu do Big data đưa ra (Reddy và các đồng nghiệp, 2017).

**2.1. Vai trò của Big data trong việc nâng cao tính giá trị của dữ liệu**

Theo những nghiên cứu gần đây của Hiệp Hội Kế Toán công chứng Anh Quốc – ACCA (2013) cho thấy khoảng 20% các công ty lớn xếp loại data (dữ liệu) như một loại tài sản trên bảng cân đối kế toán, và đây là những công ty thường có trên 10.000 nhân viên với tốc độ tăng trưởng khoảng 30%. Tính giá trị của dữ liệu sẽ phụ thuộc vào tính liên quan của dữ liệu cho từng đối tượng sử dụng và Big data cho thấy hiệu quả rất lớn trong việc nâng cao tính giá trị cho bộ dữ liệu.

Một trong những ứng dụng có thể nói là nổi bật nhất của Big data trong việc nâng cao tính giá trị của dữ liệu chính là cung cấp các thông tin theo tính thời gian thực cho từng đối tượng sử dụng (real-time information). Theo đánh giá của giám đốc về quản trị nguồn dữ liệu của công ty Shell, việc sửa chữa một dữ liệu sai sót sẽ có chi phí cao gấp 10 lần một nguồn dữ liệu đáng tin cậy và có tính thời gian thực ngay tại bước đầu (ACCA, 2013). Chính vì vậy, hiện nay một số các kế toán chuyên nghiệp bắt đầu xây dựng nền tảng cho việc xác định giá trị của dữ liệu bằng cách thiết lập các mục tiêu về thành quả hoạt động và các KPIs (Key performance indicators – chỉ số đo lường thành quả).

Ngoài ra, để cho việc kết hợp Big data vào trong hoạt động kinh doanh đạt hiệu quả cần phải thiết lập được các hệ thống kế toán phù hợp với dữ liệu được tạo ra. Ví dụ, trong kế toán tài chính, mục tiêu đó là cải thiện độ tin cậy và mức độ phù hợp của dữ liệu kế toán, và tuân thủ tốt hơn các Chuẩn mực kế toán Việt Nam (VAS). Hơn nữa, các nguồn tài liệu truyền thống (hóa đơn, biên lai, chứng từ vận chuyển, v.v) vẫn là bằng chứng chính trong việc thực hiện sổ sách kế toán, nên vai trò của Big data chủ yếu được xem là thông tin bổ sung (ví dụ, các dữ liệu phi tài chính như hình ảnh, cuộc gọi điện thoại hoặc video có thể được kết hợp với dữ liệu tài chính để cung cấp tính xác thực của số tiền trong các tài khoản chi phí. Chẳng hạn, thông tin liên quan đến uy tín của khách hàng, bao gồm danh tiếng và lịch sử giao dịch của họ có thể được thu thập thông qua các báo cáo tin tức và đánh giá trực tuyến, từ đó cho phép công ty ước tính chính xác hơn khoản chi phí nợ xấu liên quan). Nói chung, Big data giúp tạo ra các số liệu kế toán chính xác, kịp thời và được hỗ trợ bởi các bằng chứng đa dạng.

Đồng thời Big data cũng có giá trị trong kế toán quản trị bằng cách cải tiến các mô hình chi phí của công ty. Khả năng xử lý một lượng dữ liệu khổng lồ của Big data là cơ hội để tập hợp dữ liệu có sẵn và hình thành các chỉ số tiêu chuẩn cho các yếu tố về chi phí và giá cả. Ưu điểm này của Big data đặc biệt phù hợp với các công ty toàn cầu sử dụng các phương pháp chi phí khác nhau trong mỗi công ty con của họ. Một hệ

thống kế toán có thể lưu trữ dữ liệu một đơn vị chi phí bằng các loại tiền tệ khác nhau trên cơ sở thời gian thực sẽ cho phép công ty kiểm soát chi phí và lập ngân sách hiệu quả hơn.

Một khả năng hữu ích khác của Big data được thể hiện qua việc phân bổ chi phí chung. Mặc dù hệ thống xác lập phương pháp tính giá thành dựa trên hoạt động (ABC - activity-based costing) được xem là một công cụ tương đối chính xác và hợp lý trong việc phân bổ chi phí chung, việc công nhận các kích tố chi phí thực (real cost drivers) vẫn còn khoảng trống cho những sai sót và các yếu tố bị bỏ sót. Big data có thể cung cấp sự hỗ trợ trong công việc này bằng cách thu thập thông tin liên quan thông qua các cuộc phỏng vấn với nhân viên, người lao động và video giám sát để xác định rõ hơn các hoạt động đang tác động đến chi phí chung. Các nhà quản lý cũng có thể thực hiện phân tích xu hướng dựa trên các mối quan hệ trong quá khứ giữa các hoạt động và chi phí hoặc trung bình ngành để dự đoán các kích tố chi phí trong tương lai.

### **2.2 Vai trò của Big data trong việc ra quyết định**

Đối với hoạt động kế toán, Big data gắn liền với quá trình phân tích dữ liệu và/hoặc dự báo, và nó được xem như là một công cụ phân tích quan trọng để tìm ra câu trả lời cho các câu hỏi: giảm chi phí, giảm thời gian, ra quyết định thông minh... Khi việc phân tích nguồn dữ liệu lớn được hỗ trợ tối đa không chỉ cho phép tổ chức có thể hoàn thành tốt một số công việc như: xác định nguyên nhân của những thất bại, phân tích thị trường, tìm hiểu hành vi khách hàng, tính toán được những rủi ro gặp phải, phát hiện hành vi gian lận trước khi nó ảnh hưởng đến tổ chức..., mà còn giúp cho kế toán kiểm tra được các khối lượng giao dịch lớn cũng như dễ dàng phát hiện ra các sai sót trên các thiết bị công nghệ đang sử dụng hiện hành, từ đó tạo động lực cho các đơn vị nâng cao chất lượng các thiết bị công nghệ cho hệ thống thông tin kế toán. Chính các ưu điểm của Big data, để có được lợi thế cạnh tranh, các tổ chức đang dần mở rộng phạm vi hệ thống thông tin của họ từ xử lý dữ liệu truyền thống sang thu thập dữ liệu tự động trong đó họ dựa vào các thiết bị cảm biến tự động để tự động hóa hoàn toàn hệ thống hỗ trợ quản lý và sản xuất của chính họ. Trong đó, phương pháp đo lường và đảm bảo hiệu quả kinh doanh vẫn là điều đáng chú ý nhất khi sử dụng Big data (Romero, Gal, Mock, và Vasarhelyi, 2012).

### **2.3 Vai trò của Big data trong quản trị rủi ro**

Theo nghiên cứu của ACCA (2013), việc mở rộng nguồn dữ liệu dùng trong việc ước đoán các rủi ro có thể xảy ra chính là một trong những ứng dụng của Big data trong việc quản trị rủi ro. Trong nghiên cứu này chỉ ra một ví dụ điển hình đã xảy ra tại một công ty tên Canadian Tires chuyên về bán lẻ các mặt hàng bao gồm điện tử, đồ dùng nhà bếp, dụng cụ thể thao và phụ kiện xe hơi đã làm một nghiên cứu đột phá về sự liên kết giữa hành vi tiêu dùng của khách hàng và rủi ro tín dụng của họ. Nhờ vào việc phân tích chi tiết các giao dịch được thực hiện với thẻ tín dụng của khách hàng và được chấp thuận bởi các cửa hàng, nghiên cứu cho thấy rằng hành vi thanh toán trễ và thậm chí là không thanh toán của khách hàng có thể dự đoán được theo loại hàng hóa và nhãn hiệu mà họ mua cũng như các loại quán bar họ ghé thăm. Dữ liệu Big data của công ty Canadian Tire chỉ ra rằng bất cứ khách hàng nào đã mua phụ kiện xe dạng nhôm có hình hợp sọ hoặc hệ thống xả Thruster có khả năng trễ hạn thanh toán hóa đơn, và những khách hàng đi vào quán bar có tên Sharx Pool Bar ở Montreal thường sẽ trễ hạn thanh toán khoảng 4 lần với việc trễ hạn lên tới 12 tháng.

Một trong những ứng dụng khác của Big data trong việc quản trị rủi ro chính là phát hiện các rủi ro theo thời gian thực. Theo nghiên cứu của ACCA (2013), các nguồn dữ liệu “sống” (dữ liệu theo thời gian thực) là cực kỳ cần thiết cho việc phát hiện gian lận và là công cụ cho kế toán điều tra. Nhiều công ty kế toán tiên phong trong việc ứng dụng Big data thậm chí đã sử dụng việc phát hiện khuôn mẫu một cách tự động thành những ứng dụng có thể kiếm được tiền. Chẳng hạn một số các công ty kế toán ở Mỹ đã phát triển một hệ thống phát hiện gian lận cho các khách hàng của họ về các hành vi trộm cắp của nhân viên. Một phần mềm tự động phân tích xu hướng được sử dụng và sẽ gửi thông báo về những sự kiện vượt khỏi các thông số được cho trước. Khi phần mềm phát hiện những hành vi bất thường phần mềm sẽ tự động gửi thông báo tới khách hàng.

Big data cũng mở ra một phương pháp phát hiện gian lận tức thì và kiểm toán “sống” (kiểm toán theo số liệu theo thời gian thực). Dưới đây là ví dụ điển hình cho 3 chữ V trong số các đặc tính hữu dụng của Big data trong việc phát hiện gian lận, bao gồm tính đa dạng – “Variety”, dung lượng – “Volume”, tốc độ – “Velocity” mà các kiểm toán viên tại PwC đã thực hiện. Khi được hỏi về việc phát hiện gian lận, nhóm kế toán điều tra của PwC sẽ “minh họa” thông tin từ nhiều nguồn dữ liệu khác nhau, bao gồm các tài liệu, các thông tin trên mạng xã hội, các email và thậm chí là các cuộc hội thoại. Việc phân tích và xử lý thông tin của hàng triệu các dữ liệu nếu như trước kia phải mất ít nhất là hai tuần hoặc vài tháng thì nay chỉ với việc

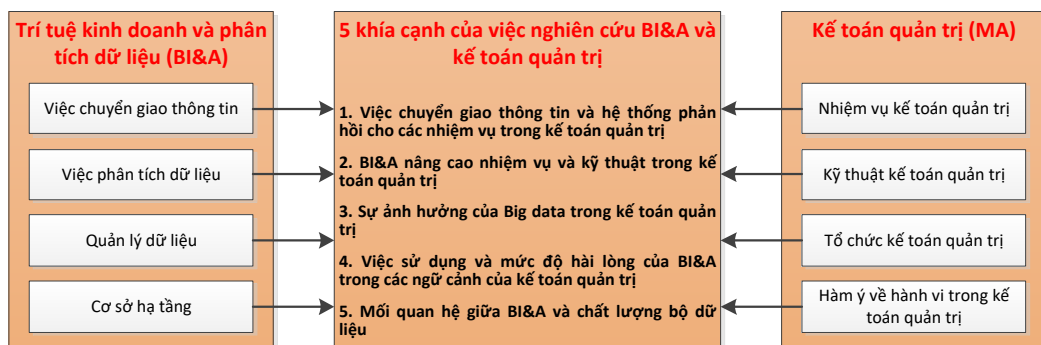
sử dụng hiệu quả Big data chỉ cần vài ngày.

**2.4 Vai trò của Big data trong việc triển khai trí tuệ kinh doanh (BI - Business intelligence) và phân tích dữ liệu**

Có rất nhiều cách nhìn nhận về Big data, trí tuệ kinh doanh và phân tích dữ liệu. Nói một cách khác, thuật ngữ được phát triển tại các thời điểm khác nhau và diễn giải trên các lĩnh vực khác nhau sẽ mang ý nghĩa rất khác nhau.

Wikipedia định nghĩa trí tuệ kinh doanh (BI) là “tập hợp các kỹ thuật và công cụ để chuyển đổi dữ liệu thô thành thông tin có ý nghĩa và hữu ích cho mục đích phân tích kinh doanh. Các chức năng phổ biến của công nghệ BI gồm báo cáo, xử lý phân tích trực tuyến, phân tích, khai thác dữ liệu, xử lý sự kiện phức tạp, quản lý hiệu quả kinh doanh, khai thác văn bản, phân tích dự đoán, v.v. Công nghệ BI có khả năng xử lý một lượng lớn dữ liệu có cấu trúc và đôi khi phi cấu trúc để giúp xác định, phát triển, và tạo ra các cơ hội và chiến lược kinh doanh mới, v.v. BI hiệu quả nhất khi kết hợp dữ liệu xuất phát từ thị trường (dữ liệu ngoài) với dữ liệu từ nguồn bên trong doanh nghiệp như dữ liệu tài chính và dữ liệu hoạt động (dữ liệu nội bộ). Khi dữ liệu bên ngoài và bên trong được kết hợp có thể cung cấp một bức tranh toàn diện về hiệu quả. BI nhằm mục đích cho phép giải thích dễ dàng những dữ liệu lớn. Xác định các cơ hội mới và thực hiện chiến lược hiệu quả dựa trên những hiểu biết của công nghệ BI có thể cung cấp cho doanh nghiệp một lợi thế thị trường cạnh tranh và sự ổn định lâu dài”.

Cao và cộng sự (2015) cho các ví dụ và thảo luận về các đặc điểm của phương pháp phân tích Big data, đặc biệt khi áp dụng cho việc kết hợp nhiều nguồn Big data khác nhau. Các tác giả đã xác định được những thách thức của phân tích Big data trong kiểm toán. BI có thể là một mục tiêu của phân tích dữ liệu. Phân tích Big data là một phương tiện có thể đạt được mục tiêu đó. Người ta có thể có được BI mà không cần thực hiện phân tích Big data. Mặt khác, phân tích Big data có thể được sử dụng để đạt được các mục tiêu khác nhau không chỉ BI. Hơn nữa, Big data là một phần mở rộng của bộ cơ sở dữ liệu phân tích hiện tại. Do khối lượng và đặc tính động của Big data nên nó tạo ra các thách thức kỹ thuật mới. Những thách thức này đang nổi lên hàng ngày và phải đối mặt về mặt khái niệm và thực tế. Nghiên cứu này chỉ đề cập đến một tập hợp nhỏ của những vấn đề này. Để minh họa cho vấn đề này nghiên cứu của Rikhardsson và Yigitbasioglu (2018) cho thấy mối liên hệ giữa BI và phân tích dữ liệu (gọi tắt là BI&A) với kế toán quản trị thông qua 5 khía cạnh chủ yếu, bao gồm: việc chuyển giao thông tin và hệ thống phản hồi cho các nhiệm vụ trong kế toán quản trị, nâng cao nhiệm vụ và kỹ thuật trong kế toán quản trị, sự ảnh hưởng của Big data trong kế toán quản trị, việc sử dụng và mức độ hài lòng của BI&A trong các ngữ cảnh của kế toán quản trị và cuối cùng là mối quan hệ giữa BI&A và chất lượng bộ dữ liệu.



Hình 2: Năm khía cạnh của việc nghiên cứu BI&A và kế toán quản trị

Nguồn: Rikhardsson và Yigitbasioglu (2018)

**3 HẠN CHẾ, RỦI RO VÀ KỲ VỌNG CỦA BIG DATA**

**3.1. Hạn chế và rủi ro khi sử dụng Big data**

Các thiết bị cảm biến chất lượng cao với dung lượng lưu trữ tăng theo cấp số nhân, và các thuật toán tính vi đã tạo ra Big data cùng với các ứng dụng ngày càng rộng của nó. Tuy nhiên, đối với mỗi công ty, tính hữu ích của dữ liệu bị giới hạn bởi số lượng, chất lượng và khả năng tiếp cận. Nhiều tổ chức không thể áp dụng các kỹ thuật Big data đơn giản vì các tổ chức không thể khắc phục được các yếu tố giới hạn chẳng hạn như thiếu dữ liệu (số lượng), dữ liệu không liên quan hoặc dữ liệu đến từ các nguồn không tin cậy (chất lượng), hoặc các kỹ năng cần thiết trong việc trích xuất thông tin (khả năng về công nghệ). Do đó, một trong những động thái đầu tiên của các doanh nghiệp nếu muốn đưa Big data vào hoạt động kế toán của họ

là phải xác định được khối lượng dữ liệu cần xử lý, đánh giá mức độ phù hợp của dữ liệu đối với công việc kế toán đang làm tại đơn vị, và quyết định liệu các công việc phân tích có nên thuê ngoài hay không. Nếu không thực hiện đúng quy trình này có thể làm giảm chất lượng của hồ sơ kế toán và ảnh hưởng cơ bản lên kết quả tài chính.

Rủi ro về bảo mật thông tin của người sử dụng ngày càng tăng cùng với sự phát triển ngày càng mạnh mẽ, tinh vi của công nghệ số và điện toán đám mây. Do đó, các tổ chức phải có trách nhiệm hơn trong quá trình sử dụng Big data vào hoạt động kế toán, quan tâm nhiều hơn đến tính bảo mật thông tin, và thiết lập các biện pháp phòng ngừa mới nhằm bảo đảm an toàn bảo mật mạng nếu không muốn các cá nhân và các tổ chức khác xâm nhập vào hệ thống để thu thập thông tin bảo mật của chính đơn vị mình.

Cuối cùng, sự phức tạp của việc sử dụng Big Data có thể tạo ra sự nhầm lẫn trong việc tìm hiểu các giả định cơ bản để phân tích, đánh giá, và xác định các khoản lợi ích cho tổ chức. Vì vậy, với bất kỳ thay đổi nào trong các mô hình giả định của mình, tổ chức phải xem chúng là những mối quan tâm quan trọng, và cần có những biện pháp xử lý phù hợp và triệt để nhằm khắc phục những sai sót của việc sử dụng Big data trước sự phát triển liên tục của nó.

### 3.2. Kỳ vọng về tương lai Big data

Tương lai của Big data chắc chắn là một điều tươi sáng. Tiềm năng của Big data vừa lớn mạnh vừa có ý nghĩa. Tiềm năng này đang được chứng minh dần trong thời gian qua đó là ngày càng có nhiều các tổ chức thuộc các lĩnh vực khác nhau đang và sẽ dần ứng dụng Big data để rút ngắn khoảng cách với khách hàng của họ và khai thác lợi ích tài chính trong tương lai. Trong kế toán, Big data có thể được sử dụng nhiều trong các lĩnh vực hiện đang dựa vào các giả định (ví dụ: nhập trước, xuất trước; nhập sau, xuất trước và chi phí nợ xấu), các nghiệp vụ kế toán phức tạp (định giá, lập dự phòng), giúp thực hiện các quy trình phân tích dữ liệu toàn diện hơn cũng như hỗ trợ các giao dịch ngày càng lớn và phức tạp.

## 4 KẾT LUẬN

Nghiên cứu này trình bày ý nghĩa của Big data, tại sao nó quan trọng và những thay đổi nào sẽ thúc đẩy trong lĩnh vực kế toán. Nghiên cứu thấy rằng, với những tiện ích hiện có của Big data đã hỗ trợ cho hoạt động phân tích các dữ liệu kế toán ngày càng được cải thiện rõ rệt, vì vậy nếu Big data được sự quan tâm chính đáng và được đưa vào áp dụng hợp lý bởi các tổ chức thuộc các lĩnh vực khác nhau sẽ góp phần tạo ra những thay đổi lớn không chỉ trong kinh doanh, mà còn cho cả giáo dục, nghiên cứu và thực hành kế toán. Ví dụ, trong giáo dục, cụ thể là kế toán, việc sử dụng Big data sẽ làm tăng nội dung thống kê và công nghệ thông tin trong chương trình giảng dạy, từ đó góp phần không nhỏ vào việc tạo ra một lực lượng lao động kế toán tương lai có kỹ năng, trình độ chuyên môn, sức sáng tạo và khả năng thích ứng với sự phát triển liên tục của khoa học công nghệ. Và những thay đổi trong giáo dục là cần thiết để sử dụng thành công Big data trong thực tiễn kế toán doanh nghiệp và kế toán công (Cao và cộng sự, 2015).

Mặt khác, khi xem xét đưa Big data vào nghiên cứu trong các lĩnh vực kế toán truyền thống, ví dụ như nghiên cứu thị trường vốn thấy rằng việc sử dụng Big data để xử lý sẽ mang lại lợi ích từ sự gia tăng dữ liệu, đồng thời góp phần tạo điều kiện để cải thiện và nâng cao các kỹ năng của nhà nghiên cứu trong các lĩnh vực như mô hình hóa, thống kê và khai thác văn bản. Sự phát triển của nghiên cứu thị trường vốn sẽ trở nên đặc biệt quan trọng nếu các nguồn Big data phi truyền thống khác nhau phụ thuộc vào những người tham gia thị trường tài chính khác nhau (Warren và cộng sự, 2015).

## TÀI LIỆU THAM KHẢO

- [1] ACCA & IMA. (2013). Big data: its power and perils. Retrieved from <http://www.accaglobal.com/bigdata>
- [2] Accounting Dictionary. (2014). Accounting Records: Definitions. Available at: <http://accountingdictionary.org/dictionary/accounting-records>.
- [3] Alles, M., & Gray, G. (2016). Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors. *International Journal of Accounting Information Systems*, 22, 44-59.
- [4] Arora, S., Barak, B., Brunnermeier, M., & Ge, R. (2011). Computational complexity and information asymmetry in financial products. *Communications of the ACM*, 54 (5), 101–107.
- [5] Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of “Big Data?”. *McKinsey Quarterly*, 24–35.

- [6] Brynjolfsson, E., Hammerbacher, J., & Stevens, B. (2011). Competing through data: Three experts offer their game plans. *McKinsey Quarterly*, 36–47.
- [7] Bughin, J., Livingston, J., & Marwaha, S. (2011). Seizing the potential of “Big Data”. *McKinsey Quarterly*, 103–109.
- [8] Cao, M., Chychyla, R., & Stewart, T. (2015). Big data analytics in financial statement audits. *Accounting Horizons*, 29(2), 423–429
- [9] Cockcroft, S., & Russell, M. (2018). Big Data Opportunities for Accounting and Finance Practice and Research: Big Data in Accounting and Finance. *Australian Accounting Review*, 28 (3), 323–333. doi: 10.1111/auar.12218.
- [10] Cukier, K., & Mayer-Schonberger, V. (2013). The rise of Big Data. *Foreign Affairs*, 28–40.
- [11] IBM big data and analytics hub. Extracting Business Value from the 4 V’s of Big Data. (2016). Retrieved from <http://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data>.
- [12] Kozlovski, S., & Vasarhelyi, M.A. (2014). An Audit Ecosystem: A Starting Point with Definitions, Attributes, and Agents. Working paper, Rutgers. The State University of New Jersey.
- [13] Moffitt, K. C., & Vasarhelyi, M.A. (2013). AIS in an age of Big Data. *Journal of Information Systems*, 27(2), 1–19.
- [14] Reddy, P.K., Sureka, A., Chakravarthy, S., Bhalla, S. (2017). *Big data analysis*. (1<sup>st</sup> ed.). Springer International Publishing
- [15] Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems, Elsevier*, 29(C), 37–58.
- [16] Romero, S., Gal, G., Mock, T.J., & Vasarhelyi, M.A. (2012). A measurement theory perspective on business measurement. *Journal of Emerging Technologies in Accounting*, 9(1), 1–24.
- [17] Shannon, C.W., & Weaver, W.W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
- [18] Syed, A., Gillela, K., & Venugopal, C. (2013). The future revolution on Big Data. *International Journal of Advanced Research in Computer and Communication Engineering*, 2 (6), 2446–2451.
- [19] Vasarhelyi, M.A, Kogan, A., & Tuttle, B.M. (2015). Big Data in Accounting: An Overview. *Accounting Horizons*, 29 (2), 381–396.
- [20] Vasarhelyi, M.A., Alles, M.G., & Williams, K.T. (2010). *Continuous Assurance for the Now Economy*. Sydney, Australia: Institute of Chartered Accountants in Australia.
- [21] Warren, J.D., Jr., K. C. Moffitt, and P. Byrnes. (2015). How accounting records will change with Big Data. *Accounting Horizons*, 29 (2), 397–407.

Ngày nhận bài: 20/12/2019

Ngày chấp nhận đăng: 13/03/2020