

# Ứng dụng dữ liệu lớn vào quản lý thuế: kinh nghiệm thế giới và định hướng phát triển của Việt Nam

Trong những năm gần đây, số lượng người nộp thuế trên phạm vi toàn cầu gia tăng nhanh chóng, quy mô hoạt động lớn và tính chất ngày càng phức tạp, tiềm ẩn nhiều rủi ro, trong khi đó nguồn lực quản lý của cơ quan thuế còn khuyết thiếu. Chính vì vậy, các nước trên thế giới cũng như Việt Nam đang đẩy mạnh cải cách thủ tục hành chính, cải cách thể chế, xây dựng hệ thống cơ sở dữ liệu lớn (Big Data) để vận hành công tác quản lý theo cơ chế đánh giá rủi ro.

Trung Kiên

## Kinh nghiệm quốc tế

Theo đại diện Ban Quản lý rủi ro, Tổng cục Thuế, việc xử lý dữ liệu với Big Data, dựa trên khái niệm năm chữ “V”, đó là Volume (khối lượng): liên quan đến lượng lớn dữ liệu được tạo ra; Variety (đa dạng): các nguồn dữ liệu đa dạng, làm tăng độ phức tạp của các phân tích; Velocity (vận tốc): với khối lượng lớn và sự đa dạng của dữ liệu, tất cả quá trình xử lý cần phải tăng tốc để tạo ra thông tin cần thiết; Veracity (tính xác thực): liên quan trực tiếp đến độ tin cậy của dữ liệu; Value (giá trị): những thông tin hữu ích có thể thu được từ dữ liệu.

Hầu hết các giải pháp Big Data đều bao gồm các thành phần: nguồn dữ liệu, kho dữ liệu, xử lý dữ liệu trong thời gian thực, xử lý luồng, kho dữ liệu phân tích, báo cáo và điều phối. Việc phân tích dữ liệu lớn, cùng với trí tuệ nhân tạo, có thể giúp các nhà quản lý tăng đáng kể việc tuân thủ pháp luật thuế. Ví dụ: dữ liệu lịch sử về phản ứng của người nộp thuế khi nhận được thông báo thuế có thể được sử dụng để dự đoán động thái tiếp theo của họ. Hay việc phân tích dự đoán cũng có thể cho thấy xu hướng, gắn cờ (đánh dấu) các vấn đề tiềm ẩn và xác định các trường hợp rủi ro cao để điều tra sâu hơn, qua đó có khả năng cắt đứt con đường gian lận ngay cả trước thời điểm diễn ra.

Tại Hoa Kỳ, việc ngăn chặn gian lận thuế được thực hiện với việc triển khai Big Data. Cụ thể, một trong những vấn đề lớn nhất



đối với các tổ chức nhà nước là phân biệt những người đóng thuế có thiện chí và những người cố tình gian lận có hệ thống để giảm các khoản thanh toán thuế phải nộp. Một vài năm trước, Sở Thuế vụ Mỹ thừa nhận đã sử dụng Big Data để ngăn chặn gian lận thuế. Một trong những chiến lược được sử dụng đó là khai thác dữ liệu truyền thông xã hội để chứng minh rằng một người sống giàu có hơn hồ sơ thuế của họ. Điều này giúp bảo toàn thành công 300 tỷ USD tiền thuế thất thoát mỗi năm. Big Data giải quyết vấn đề này bằng cách sử dụng phân loại dữ liệu và nhận dạng mẫu dựa trên dấu vết, tách biệt những kẻ lừa đảo và người nộp thuế chính đáng.

Cơ quan thuế vương quốc Anh (HMRC) cũng là một trong những đơn vị tiên phong trong phân tích Big Data để cải thiện công tác quản lý thuế. Hệ thống “HMRC Connect” đã giành được giải thưởng “Dự án dữ liệu lớn của

Vương quốc Anh 2014” và đang tiếp tục được hoàn thiện. Hệ thống có nguồn dữ liệu nội bộ từ dịch vụ công và nguồn dữ liệu thông qua các thỏa thuận với các ngân hàng, nền tảng kinh tế chia sẻ (như Airbnb, eBay), các công ty bảo hiểm, trao đổi dữ liệu với lãnh thổ hải ngoại của Anh và các nước OECD. Ngoài ra, còn sử dụng dữ liệu từ các mạng xã hội (Facebook, Twitter). Bên cạnh đó, HMRC còn xây dựng ADEPT (Analytics for Debtor Profiling and Targeting) là một hệ thống phân tích Big Data có dung lượng 13,5 terabyte (kích thước năm 2017), với các ứng dụng quản lý nợ và công cụ hỗ trợ ra quyết định. Hệ thống này tự động kết hợp dữ liệu từ 20 hệ thống bên trong và bên ngoài để kiểm soát những thay đổi, liên quan đến một triệu khoản nợ mỗi ngày. Từ đó, ADEPT có khả năng đề xuất chuỗi các biện pháp can thiệp thu nợ tùy chỉnh cho từng khoản nợ.

Tại Australia, Văn phòng Thuế (ATO) đang xây dựng một giải pháp phân tích mạng gọi là “ANGIE-Automated Network & Grouping Identification Engine” để giúp lực lượng đặc nhiệm phân biệt các mối quan hệ phức tạp, nhiều lớp giữa các khách hàng. Lực lượng đặc nhiệm được thành lập vào năm 2016 để trấn áp việc trốn thuế của các công ty đa quốc gia, các DNNN, DN tư nhân lớn và các cá nhân liên quan. ANGIE sẽ tự động xác định và xếp nhóm các khách hàng để hiểu mối quan hệ của họ với nhau, từ đó cho phép lực lượng đặc nhiệm phát hiện các mô hình quan tâm và hình dung các liên kết mới giữa các khách hàng để phân tích, kịp thời phát hiện và xử lý các trường hợp vi phạm.

#### **Và một số định hướng phát triển ở Việt Nam**

Tại Việt Nam, cơ sở dữ liệu chuyên ngành quản lý thuế đã sớm được xây dựng và đưa vào triển khai sử dụng. Cơ sở dữ liệu chuyên ngành quản lý thuế được xây dựng theo công nghệ kho dữ liệu nguồn - Data Warehouse hỗ trợ công tác nghiệp vụ, công tác quản lý thuế, công tác chỉ đạo, điều hành và ra quyết định. Hệ thống lưu trữ quản lý các nguồn dữ liệu từ hệ thống: Quản lý thuế tập trung (TMS), Báo cáo tài chính (BCTC), Quản lý ấn chỉ (QLAC), Thanh tra kiểm tra (TTR). Việc sử dụng cơ sở dữ liệu chuyên ngành quản lý thuế đã góp phần hỗ trợ công tác khai thác số liệu phục vụ quản lý thuế cũng như giám sát áp lực lên hệ thống tác nghiệp trong việc lấy báo cáo, dữ liệu phục vụ yêu cầu nghiệp vụ, cung cấp thông tin cho các đơn vị bên ngoài về đăng ký thuế, nghĩa vụ kê khai, số thuế phải nộp qua công Thông tin ngành Thuế.

Không chỉ có vậy, từ ngày 21/4/2022, Tổng cục Thuế đã triển khai hệ thống hóa đơn điện tử trên toàn quốc. Với hàng tỷ hóa đơn/năm, đây sẽ là nguồn dữ liệu quý trong việc xây dựng cơ sở dữ liệu phục vụ áp dụng quản lý rủi ro trong quản lý thuế.

Đại diện Tổng cục Thuế cho biết, từ các nghiên cứu kinh nghiệm quốc tế triển khai ứng dụng Big Data trong quản lý thuế và trên cơ sở đánh giá hiện trạng cơ sở dữ liệu, hạ tầng, hệ thống ứng dụng, nguồn nhân lực của ngành Thuế Việt Nam, cơ quan thuế đã đặt mục tiêu xây dựng Hệ thống ứng dụng dữ liệu phân tích và quản lý rủi ro tổng thể, toàn diện trên nền tảng Big Data với các công cụ thu thập, xử lý, phân tích dữ liệu hiện đại, phù hợp với thông lệ quốc tế và xu thế phát triển của công nghệ, đáp ứng yêu cầu quản lý thuế toàn diện trên cơ sở phân tích rủi ro và yêu cầu chia sẻ dữ liệu giữa các cơ quan quản lý nhà nước, tạo thuận lợi cho người dân và nâng cao hiệu quả quản lý.

Để thực hiện được mục tiêu này, ngành Thuế đang tích cực nâng cấp hệ thống hạ tầng kỹ thuật, máy chủ lưu trữ, phân tích dữ liệu, hạ tầng kết nối và phần mềm ứng dụng các yêu cầu xây dựng kho cơ sở dữ liệu quản lý rủi ro; đào tạo quản trị hệ thống hạ tầng kỹ thuật đáp ứng yêu cầu vận hành Hệ thống ứng dụng phân tích dữ liệu và quản lý rủi ro. Khi Hệ thống ứng dụng phân tích dữ liệu và quản lý rủi ro được đưa vào sử dụng, sẽ giúp ngành Thuế thay đổi phương thức quản lý, tăng cường tính công khai, minh bạch, giảm thiểu sự can thiệp trực tiếp của cán bộ thuế, cơ quan thuế vào việc ra quyết định đối với người nộp thuế. Việc áp dụng quản lý rủi ro cũng đảm bảo hiệu lực, hiệu quả của công tác quản lý thuế; khuyến khích, tạo điều kiện thuận lợi để người nộp thuế tự nguyện tuân thủ tốt các quy định của pháp luật; đồng thời góp phần phòng chống, phát hiện, xử lý kịp thời và hiệu quả các hành vi vi phạm pháp luật về thuế. Đặc biệt đối với người nộp thuế, Hệ thống phân tích dữ liệu và quản lý rủi ro sẽ đảm bảo tính công bằng, minh bạch trong việc thực hiện trách nhiệm đối với NSNN ■

## **Khẩn trương...**

▶▶ Tiếp theo trang 19

của thẩm quyền quản lý, nhất là các dữ liệu liên quan đến bí mật kinh doanh của DN, hoặc dữ liệu khách hàng được bảo vệ và chỉ được chia sẻ theo quy định của pháp luật.

Bổ sung thêm, ông Ngô Vĩnh Bạch Dương, Viện Nhà nước và Pháp luật đánh giá, nhu cầu và yêu cầu chia sẻ dữ liệu giữa DN và cơ quan nhà nước đang ngày càng tăng và có thể thực hiện theo hình thức tự nguyện hay bắt buộc. Tuy nhiên, dữ liệu là tài sản được pháp luật bảo vệ theo các quy tắc khác nhau, như quyền riêng tư, quyền tài sản, quyền sở hữu trí tuệ... Do đó, tới đây hành lang pháp lý quy định việc chia sẻ dữ liệu cần thiết lập trên nguyên tắc: về mục đích - chỉ áp dụng vì lợi ích cộng đồng; về chi phí - DN được bù đắp chi phí cho việc xây dựng và bảo vệ dữ liệu cùng các chi phí phục vụ kết nối, chia sẻ dữ liệu.

Trên cơ sở những vấn đề thực tiễn phát sinh, CIEM đưa ra khuyến nghị, cần khẩn trương xem xét xây dựng quy định về kết nối, chia sẻ dữ liệu giữa DN với cơ quan nhà nước theo hướng làm rõ căn cứ pháp lý, hợp lý để cơ quan nhà nước có thẩm quyền và “đu quyền” yêu cầu DN chia sẻ. Trong đó, quy định rõ hơn về các loại dữ liệu có liên quan mà DN có nghĩa vụ chia sẻ với cơ quan nhà nước trong trường hợp cụ thể; DN có quyền phản hồi về yêu cầu chia sẻ dữ liệu, cả về tính hợp pháp, hợp lý, mục đích, thời gian chia sẻ dữ liệu. Ngoài ra, Chính phủ cũng cần sớm ban hành quy định về dữ liệu mở của cơ quan nhà nước theo thẩm quyền được giao trong Luật Giao dịch điện tử; đồng thời nâng cao năng lực điều tra, xử lý vi phạm của cơ quan nhà nước theo trình tự thủ tục theo luật định ■