KINH NGHIỆM VÀ THÁCH THỨC VỀ VIỆC SỬ DỤNG CÁC NGUỒN DỮ LIỆU MỚI TRONG CƠ QUAN THỐNG KÊ HÀN QUỐC

Tóm tắt:

Bài viết này trình bày chiến lược và dự án về các nguồn dữ liệu mới trong Cơ quan Thống kê Hàn Quốc (KOSTAT) kể từ khi ra mắt bộ phận mới về dữ liệu lớn vào tháng 10 năm 2015. Chiến lược này tập trung vào liên kết dữ liệu giữa dữ liệu của khu vực công (dữ liệu của KOSTAT như dữ liệu hành chính và điều tra dân số) và dữ liệu lớn của khu vực tư nhân (ví dụ dữ liệu điện thoại di động, dữ liệu truyền thông xã hội) cũng như thiết lập khung thể chế và hợp tác. KOSTAT đã thực hiện thành công các dự án như liên kết dữ liệu đánh giá tín dụng cá nhân cũng như dữ liệu điện thoại di động với dữ liệu KOSTAT, tổ chức các diễn đàn dữ liệu lớn và thiết lập hợp tác quốc tế. Tuy nhiên, KOSTAT vẫn phải đối mặt với những thách thức cần khắc phục: Hạn chế truy cập thông tin cá nhân vào dữ liệu của khu vực tư nhân do luật bảo vệ quyền riêng tư, thiếu sự hợp tác từ các nhà cung cấp dữ liệu khu vực tư nhân, thiếu chuyên gia như nhà khoa học dữ liệu và dữ liệu lớn của khu vực tư nhân có chất lượng thấp. Trước mọi thách thức, KOSTAT nên tiếp tục trao đổi nhiều hơn với các bên liên quan từ các chính trị gia, các nhà hoạch định chính sách, doanh nghiệp, học viện và tổ chức phi chính phủ để giúp họ hiểu tầm quan trọng của các nguồn dữ liệu mới cho thống kê chính thức và nâng cao năng lực nội bộ trên cơ sở hạ tầng dữ liệu lớn. Ngoài ra, điều quan trọng là hợp tác với các tổ chức quốc tế để giải quyết các vấn đề dữ liệu lớn.

Bài viết này đã được trình bày tại hội thảo Hội nghị các nhà thống kê châu Âu năm 2019 về nguồn dữ liệu mới - khả năng truy cập và sử dụng, tại phiên 1 "Truy cập các nguồn dữ liệu mới" để thảo luận.

I. Giới thiệu

1. Để thay đổi mô hình sản xuất thông tin thống kê từ điều tra thực địa truyền thống sang phương pháp mới để thu thập dữ liệu, KOSTAT đã tiếp tục nỗ lực sử dụng dữ liệu hành chính để tổng hợp số liệu thống kê chính thức. Do đó, điều tra dân số dựa trên đăng ký đã được tiến hành lần đầu tiên vào năm 2015. Mười hai biến cơ bản của tổng dân số như tên, tuổi, giới tính và đặc điểm hộ gia đình được thu thập bằng cách sử dụng hai mươi bốn nguồn dữ liệu hành chính từ mười ba cơ quan chính phủ. Năm mươi hai biến không thu được từ dữ liệu hành chính được thu thập bằng điều tra thực địa từ 20% tổng thể mẫu. Ngoài ra, KOSTAT đã thực hiện dự án cơ sở dữ liệu đăng ký thống kê toàn diện để thiết lập bốn cơ sở dữ liệu ngành sử dụng dữ liệu hành chính: Dân số/hộ gia đình, nhà ở/tòa nhà, kinh doanh/doanh nghiệp và hoạt động kinh tế.

2. Gần đây, dữ liệu lớn đã nhận được sự quan tâm cao như một nguồn dữ liệu mới trong thống kê cũng như trong kinh doanh. Về khía cạnh thống kê, dữ liệu lớn có thể cung cấp dữ liệu phù hợp và kịp thời hơn cho việc ra quyết định thông qua việc liên kết các dữ liệu khác nhau và giảm chi phí sản xuất thống kê mà không cần khảo sát thực địa để thu thập dữ liệu. Về khía cạnh kinh doanh, dữ liệu lớn tạo ra động cơ tăng trưởng mới như là cốt lõi của cuộc cách mạng công

nghiệp lần thứ 4 như phân tích dữ liệu lớn cho công nghệ vạn vật kết nối (IoT) và trí tuệ nhân tạo (AI). Trong bối cảnh này, KOSTAT đã thành lập bộ phận mới về dữ liệu lớn vào tháng 10 năm 2015 và thực hiện nhiều dự án để phát triển số liệu thống kê chính thức theo chiến lược dữ liệu lớn mới. Tuy nhiên, vẫn còn nhiều hạn chế để sử dụng dữ liệu lớn cho thống kê chính thức. Do đó, bài viết này xem xét kinh nghiệm KOSTAT trong việc tạo điều kiện sử dụng dữ liệu lớn và các thách thức liên quan.

3. Cấu trúc của bài viết như sau: Phần 2 trình bày chiến lược của KOSTAT về dữ liệu lớn và dự án; Phần 3 cho thấy những thách thức phải đối mặt trong việc sử dụng dữ liệu lớn để thống kê chính thức; Phần cuối cùng trình bày tóm tắt và kết luận.

II. Trải nghiệm của KOSTAT trên nguồn dữ liệu mới

A. Chiến lược về dữ liệu lớn

4. Mặc dù không có định nghĩa khẳng định về dữ liệu lớn, nhưng nó thường đề cập đến các nguồn dữ liệu được mô tả là khối lượng lớn, tốc độ nhanh và nhiều loại dữ liệu đòi hỏi các hình thức xử lý sáng tạo, hiệu quả về chi phí để nâng cao cái nhìn sâu sắc và ra quyết định (UNECE, 2013).

5. KOSTAT đã ra mắt một bộ phận mới về dữ liệu lớn vào tháng 10 năm 2015 để tạo điều kiện sử dụng các nguồn dữ liệu mới, tức là dữ liệu lớn, để thống kê chính thức. Sau khi xác định dữ liệu lớn là thông tin thống kê, hướng tới dữ liệu hữu ích thông qua liên kết và phân tích dữ liệu, KOSTAT đã thiết lập chiến lược dữ liệu lớn bao gồm hai cách tiếp cận: Sản xuất thông tin thống kê khác nhau và thiết lập khung thể chế và hợp tác. Chiến lược này có bốn nhiệm vụ: (1) liên kết dữ liệu lớn của khu vực công và khu vực tư nhân, (2) cung cấp số liệu thống kê mới và bổ sung số liệu thống kê hiện có, (3) thiết lập khung pháp lý và thể chế và (4) tăng cường hợp tác bên ngoài. Theo chiến lược này, nhiều dự án đã được thực hiện.

B. Dự án

1. Liên kết dữ liệu lớn của khu vực công và khu vực tư nhân

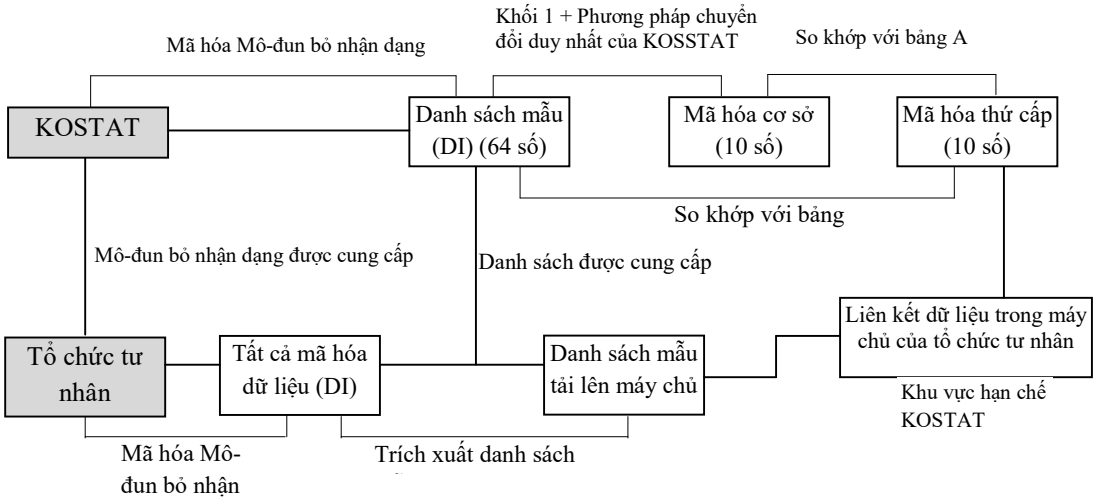
6. KOSTAT có rất nhiều dữ liệu hành chính (khoảng 89 loại) được thu thập từ các cơ quan chính phủ khác cũng như dữ liệu điều tra (khoảng 42 loại) bao gồm cả điều tra dân số và điều tra kinh doanh. Để sử dụng và liên kết dữ liệu hành chính với dữ liệu khác, số đăng ký thường trú (RRN) trong dữ liệu hành chính được chuyển đổi thành Số nhận dạng thống kê (SIN). Số này bị xóa khỏi cơ sở dữ liệu đăng ký KOSTAT để bảo vệ quyền riêng tư. Vì mỗi người có SIN riêng, do đó, dữ liệu trong cơ sở dữ liệu đăng ký có thể được sử dụng để tạo dữ liệu mới hoặc cải thiện số liệu thống kê chính thức thông qua liên kết với dữ liệu lớn của khu vực tư nhân như dữ liệu điện thoại di động, dữ liệu thẻ tín dụng, dữ liệu nợ cá nhân, v.v…

7. Tuy nhiên, không dễ để có được dữ liệu của khu vực tư nhân vì các công ty tư nhân không bắt buộc phải cung cấp dữ liệu cho KOSTAT cho các mục đích khác ngoài việc tạo ra số liệu thống kê chính thức. Trong trường hợp dùng cho sản xuất thông tin thống kê chính thức, KOSTAT có thể lấy dữ liệu của khu vực tư nhân theo Luật Thống kê. Vì các nghiên cứu thí điểm trong các dự án dữ liệu lớn không có nghĩa là đưa ra số liệu thống kê chính thức được phê duyệt theo quy trình chính thức, rất khó để có được dữ liệu của khu vực tư nhân.

8. Về vấn đề này, KOSTAT đã thiết lập một khung hợp tác thông qua việc ký một

biên bản ghi nhớ (MOU) với các nhà cung cấp dữ liệu khu vực tư nhân và thực hiện các dự án liên kết dữ liệu cùng với họ. Cách tiếp cận này có thể có lợi cho cả hai bên vì khu vực tư nhân có thể đóng góp cho lợi ích công cộng và cải thiện phương pháp biên soạn dữ liệu lớn của họ. Để bảo vệ quyền riêng tư trong quá trình liên kết dữ liệu, một phương pháp bỏ thông tin cá nhân đã được phát triển. Mô-đun bỏ nhận dạng (DI) tương tự được sử dụng trong KOSTAT để sản xuất SIN được áp dụng cho dữ liệu lớn của khu vực tư nhân để tạo cùng khóa liên kết (Hình 1).

Hình 1: Quy trình bỏ nhận dạng



9. Các dự án liên kết bao gồm xây dựng số liệu thống kê nợ của hộ gia đình sử dụng dữ liệu đánh giá tín dụng tư nhân từ Cục tín dụng Hàn Quốc (KCB) và đo lường mới về thời gian dành cho giải trí và làm việc thông qua sử dụng dữ liệu định vị điện thoại di động từ Korea Telecom (KT).

10. Mục tiêu của dự án về nợ hộ gia đình là cung cấp số liệu thống kê nợ chính xác theo đặc điểm của hộ gia đình (ví dụ: Hộ gia đình độc thân, tự làm chủ, v.v...) cho các nhà hoạch định chính sách do nợ hộ gia đình ở Hàn Quốc gia tăng. Có số liệu thống kê nợ hộ gia đình vĩ mô và vi mô ở Hàn Quốc. Thống kê vĩ mô được thu thập từ ngành tài chính phản ánh toàn bộ khối lượng nợ của hộ gia đình nhưng họ không cung cấp thông tin về các loại hộ gia đình khác nhau. Mặt khác, thống kê vi mô từ khảo sát hộ gia đình cung cấp tình hình nợ theo đặc điểm của hộ gia đình nhưng đánh giá thấp tổng số nợ. Do đó, rất hữu ích khi kết hợp dữ liệu nợ hộ gia đình vĩ mô với dữ liệu KOSTAT, chẳng hạn như điều tra dân số để lấy thông tin hộ gia đình. Bước đầu tiên, 5.000 cơ sở dữ liệu nợ mới lập gia đình (được xác định là chưa đầy 5 năm sau khi kết hôn) từ tháng 10 năm 2010 đến tháng 11 năm 2014 đã được thiết lập và phân tích thông qua liên kết dữ liệu KCB[1] với dữ liệu KOSTAT[2].

[1] *Biến từ dữ liệu KCB: Thu nhập, xếp hạng tín dụng, số dư cho vay, số tiền quá hạn, trả nợ, sử dụng thẻ.*

[2] *Biến từ dữ liệu KOSTAT: Ngày kết hôn, tuổi, nghề nghiệp, trình độ học vấn, số con, loại nhà ở, loại hộ gia đình, số lượng nhà sở hữu, thu nhập, loại công việc.*

11. Trong số các nguồn dữ liệu lớn khác, dữ liệu điện thoại di động có sự quan tâm cao từ cộng đồng thống kê vì tỷ lệ thâm nhập cao và tính kịp thời của chúng. Sự sẵn có của chúng cho các khu vực địa lý nhỏ với tính kịp thời cung cấp cơ hội tạo ra các số liệu thống kê phân tách về di cư, du lịch, quản lý thảm họa, v.v... Trong bối cảnh này, KOSTAT đã triển khai một dự án điện thoại di động để kiểm tra khả năng và tính hữu ích của việc sử dụng dữ liệu điện thoại di động để sản xuất mới số liệu thống kê đo lường chất lượng cuộc sống như thời gian cho giải trí, thời gian đi lại, nghèo thời gian thông qua liên kết dữ liệu KOSTAT và dữ liệu điện thoại di động. Có ba nhà khai thác mạng di động (MNO) tại Hàn Quốc, gồm có SKT, KT và LGU+. Trong số đó, KT có thị phần khoảng 31% tham gia dự án KOSTAT. Trong dự án này, chỉ có hai quận ở Seoul (tức là Gangnam-gu và Dobong-gu) theo Tổng sản phẩm quốc nội khu vực được chọn để so sánh mô hình hạnh phúc giữa khu vực giàu và nghèo. Do một khối lượng lớn dữ liệu điện thoại di động, dữ liệu KOSTAT đã được lưu trữ trong hệ thống phân tích dữ liệu lớn KT sau khi được xác định lại và liên kết với dữ liệu điện thoại di động. Các bộ dữ liệu được liên kết đã được nhân viên KOSTAT truy cập và phân tích tại một địa điểm được chỉ định trong văn phòng KT. Các kết quả ước tính không đại diện cho toàn bộ dân số ở hai quận vì dữ liệu KT chỉ chiếm khoảng 30% tổng dân số. Do đó, các bảng tổng hợp được tổng hợp bằng phương pháp "Trọng số xếp hạng" của thang điểm xem xét bốn biến số (vùng, giới tính, tuổi tác, tình trạng hôn nhân, loại nhà) thông qua ánh xạ dữ liệu KT để tính toán dân số dựa trên đăng ký.

2. Cung cấp số liệu thống kê mới và bổ sung số liệu thống kê hiện có

12. Ngày càng có nhu cầu cao hơn từ các nhà hoạch định chính sách về dữ liệu kinh tế kịp thời vì hầu hết dữ liệu kinh tế được phát hành hàng tháng hoặc hàng quý. Để đáp ứng nhu cầu, KOSTAT đã phát triển 14 "chỉ số kinh tế kịp thời", sử dụng nhiều nguồn dữ liệu khác nhau: Chỉ số giá giỏ hàng hóa, phí điện quá hạn thanh toán, v.v... Các chỉ số được phát hành mỗi tuần.

13. Để bổ sung số liệu thống kê hiện có, các chỉ số giá trực tuyến hàng ngày và hàng tháng dựa trên 284 mặt hàng sản phẩm được tính bằng dữ liệu giá từ 6 trang web trung tâm mua sắm trực tuyến không bao gồm giá dịch vụ. Tuy nhiên, có một số hạn chế: (i) không thể thu thập dữ liệu khi thay đổi liên kết web bằng cách sửa đổi trang web hoặc danh mục thay đổi mà không cần thông báo, (ii) sản phẩm theo mùa vụ không thu thập được đầy đủ, (iii) không thực hiện điều chỉnh chất lượng như đối với CPI do đó giá điện tử, quần áo, v.v... giảm.

14. Ngoài ra, một chỉ số kinh tế truyền thông xã hội được tính bằng dữ liệu truyền thông xã hội (ví dụ: Tin tức, blog, bảng thông báo và twitter) liên quan đến tình hình kinh tế trong bốn lĩnh vực: Điều kiện sống, tình hình kinh tế, thu nhập hộ gia đình và chi tiêu tiêu dùng. Sau khi thu thập tài liệu có chứa từ khóa (138) từ Blog, quán cà phê internet, tin tức và Twitter bằng cách thu thập dữ liệu trên web hàng ngày, các tài liệu tích cực và tiêu cực được tính và các chỉ số được tiêu chuẩn hóa cho bốn tên miền được tính toán. Cuối cùng, một chỉ số tổng được chuyển hóa.

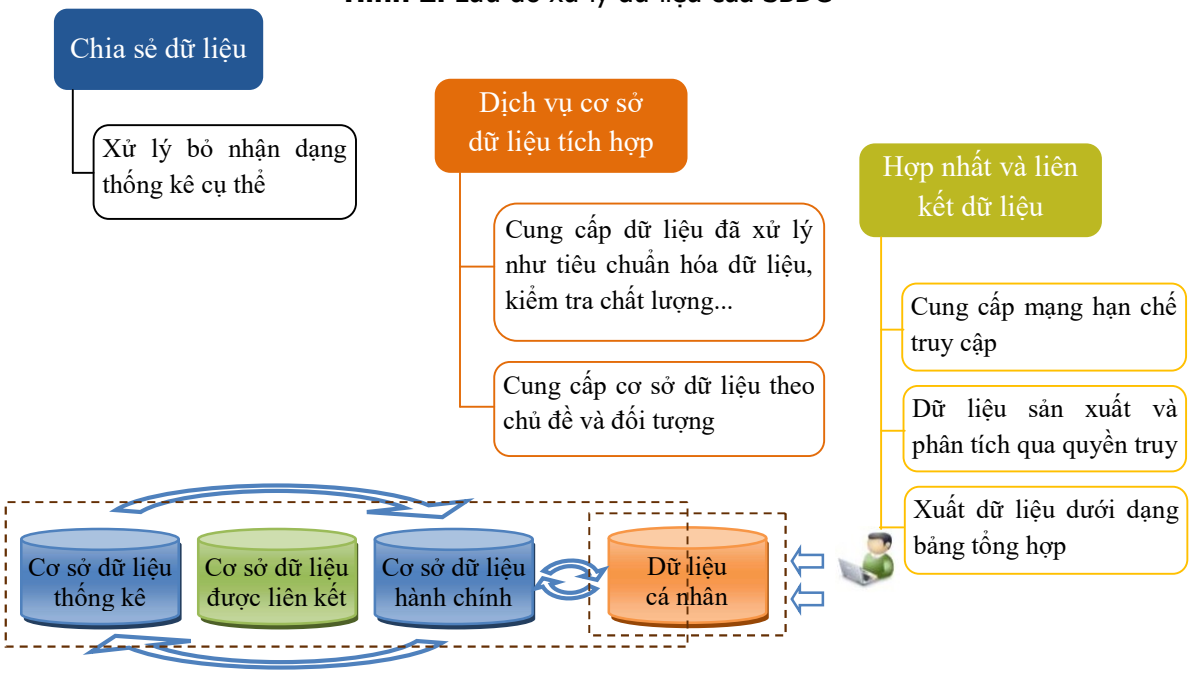
3. Thiết lập khung pháp lý và thể chế

15. KOSTAT liên tục cố gắng sửa đổi "Luật Thống kê" để có cơ sở pháp lý truy cập dữ liệu lớn của khu vực tư nhân. Luật hiện hành cho phép cơ quan thống kê chỉ thu thập

dữ liệu của khu vực tư nhân để sản xuất số liệu thống kê chính thức. Do đó, sửa đổi bao gồm quyền hợp pháp để thu thập dữ liệu từ khu vực tư nhân trong trường hợp thí điểm các dự án dữ liệu lớn kiểm tra khả năng biên soạn số liệu thống kê chính thức.

KOSTAT đã thiết lập cơ sở hạ tầng dữ liệu mở và chia sẻ được gọi là "Trung tâm dữ liệu lớn thống kê (SBDC)" với mục đích hỗ trợ liên kết dữ liệu lớn của khu vực công và khu vực tư nhân và cung cấp dịch vụ bỏ nhận dạng. Chức năng chính của nó là thực hiện kiểm tra chất lượng dữ liệu hành chính; để cung cấp cơ sở dữ liệu đăng ký theo đối tượng (dân số, nhà ở, hoạt động kinh tế, v.v…) và cơ sở dữ liệu điều tra thống kê; và để cung cấp các dịch vụ liên kết theo yêu cầu như nhận dạng. Khách hàng có thể liên kết dữ liệu của họ với dữ liệu KOSTAT ở một nơi được chỉ định và xuất dữ liệu dưới dạng bảng tổng hợp. Hiện tại Trung tâm được đặt tại ba thành phố: Seoul, Busan và Daejeon. Quá trình xử lý dữ liệu được trình bày trong Hình 2.

Hình 2: Lưu đồ xử lý dữ liệu của SBDC



4. Tăng cường hợp tác bên ngoài

16. Để giao tiếp và thảo luận với các bên liên quan từ các viện nghiên cứu, doanh nghiệp và chính phủ là rất quan trọng để giải quyết các vấn đề liên quan đến việc tạo điều kiện sử dụng dữ liệu lớn. Trong bối cảnh này, KOSTAT tổ chức "Diễn đàn Chiến lược Thống kê" diễn ra hàng quý kể từ năm 2015. Ngoài ra, KOSTAT đã đồng tổ chức một "Diễn đàn dữ liệu lớn" với hai bộ liên quan đến dữ liệu lớn: Bộ An Ninh & Nội vụ và Bộ Khoa học & CNTT.

17. KOSTAT cũng đang tham gia hợp tác quốc tế liên quan đến dữ liệu lớn: Nhóm làm việc toàn cầu của Liên hợp quốc về dữ liệu lớn và với Cơ quan Thống kê Hà Lan (CBS). CBS và KOSTAT đã xây dựng hợp tác song phương kể từ khi thỏa thuận về dữ liệu lớn được ký kết trong Hội nghị thượng đỉnh thương mại Hà Lan - Hàn Quốc vào tháng 9

năm 2016. Trong thỏa thuận đó, các hoạt động chung trong bảy lĩnh vực được xác nhận: (i) có được nguồn dữ liệu lớn; (ii) phát triển các kỹ thuật để khám phá dữ liệu lớn, chẳng hạn như các kỹ thuật dựa trên trí tuệ nhân tạo hoặc kỹ thuật khai thác dữ liệu và văn bản; (iii) chuyên sâu về phương pháp/ phân tích trong việc chọn lọc dữ liệu lớn và xử lý dữ liệu; (iv) e-learning; (v) trao đổi nhân lực; (vi) chia sẻ kinh nghiệm trong liên kết dữ liệu lớn công-tư; (vii) dữ liệu lớn và quyền riêng tư.

III. Những thách thức

18. Bất kể nhiều câu chuyện thành công trong việc tạo điều kiện sử dụng các nguồn dữ liệu mới, KOSTAT vẫn phải đối mặt với nhiều thách thức, cả bên trong và bên ngoài.

19. Thứ nhất, vẫn khó tiếp cận thông tin cá nhân trong khu vực tư nhân do luật bảo vệ quyền riêng tư mạnh mẽ ở Hàn Quốc. Đạo luật bảo vệ thông tin cá nhân (PPIA) là luật bảo vệ dữ liệu chung chi phối việc thu thập và xử lý dữ liệu cá nhân. Có các luật cụ thể theo ngành: Đạo luật Mạng, Đạo luật Thông tin tín dụng và Đạo luật Thông tin vị trí. Trong PPIA, định nghĩa về dữ liệu cá nhân quá rộng. Việc sử dụng dữ liệu cá nhân cần có sự đồng ý trước, tức là phương pháp chọn tham gia. Dữ liệu cá nhân[3] được định nghĩa là dữ liệu về một người sống có thể xác định cá nhân cũng như dữ liệu có thể xác định bằng cách dễ dàng kết hợp với các thông tin khác. Do đó, luật này gây khó khăn cho việc sử dụng dữ liệu lớn để liên kết sử dụng thông tin cá nhân.

20. Thứ hai, các nhà cung cấp dữ liệu từ khu vực tư nhân có nhận thức thấp về hợp tác về dữ liệu. Họ miễn cưỡng chia sẻ dữ liệu do luật bảo vệ quyền riêng tư mạnh mẽ cũng như cách tiếp cận thụ động của họ về chia sẻ dữ liệu.

21. Thứ ba, dữ liệu lớn có thể có chất lượng thấp do chúng không được thu thập bằng các phương pháp điều tra truyền thống theo hướng dẫn thống kê chính thức hoặc khung chất lượng mà bằng các phương pháp dựa trên CNTT-TT như kiểm duyệt từ điện thoại di động, nhà cung cấp dữ liệu, v.v... thiếu các chiều chất lượng như tính đại diện, tính nhất quán và tính đầy đủ.

22. Cuối cùng, thiếu các chuyên gia như các nhà khoa học dữ liệu và cơ sở hạ tầng CNTT để xử lý dữ liệu lớn trong KOSTAT. Để phân tích dữ liệu lớn đòi hỏi các kỹ năng và cơ sở hạ tầng CNTT khác nhau so với phân tích thống kê và xử lý dữ liệu truyền thống. Các nhà khoa học dữ liệu cần có kiến thức về nhiều lĩnh vực như: Hadoop, NoSQL, trực quan hóa dữ liệu, học máy và khai thác văn bản, v.v… KOSTAT đã hạn chế khả năng tuyển dụng nhân viên mới có kỹ năng phân tích cao theo hệ thống tuyển dụng của chính phủ hiện tại, vì hạn chế về ngân sách và quy trình tuyển dụng không linh hoạt. Để đào tạo nhân viên hiện tại để phát triển kỹ năng của họ mất nhiều thời gian. Về cơ sở hạ tầng CNTT, các khoản đầu tư lớn vào kho dữ liệu và phần mềm để thu thập dữ liệu, lưu trữ dữ liệu, phân tích dữ liệu và trực quan hóa dữ liệu được yêu cầu. Do đó, KOSTAT không thể xây dựng hệ thống phân tích dữ liệu lớn của riêng mình do ngân sách hạn chế.

IV. Kết luận

(Xem tiếp trang 34)

[3] *Bất kỳ dữ liệu nào liên quan đến một người đang sống mà cá nhân có thể được xác định thông qua tên, số đăng ký cư trú, hình ảnh trực quan, v.v… (bao gồm thông tin có thể dễ dàng kết hợp với thông tin khác để xác định một cá nhân cụ thể).*

Tiếp theo trang 29

23. Theo chiến lược dữ liệu lớn được thành lập vào năm 2016, KOSTAT đã thực hiện các dự án để kiểm tra khả năng tạo điều kiện sử dụng dữ liệu lớn cho thống kê chính thức tập trung vào liên kết dữ liệu của khu vực công (như dữ liệu KOSTAT như dữ liệu hành chính và dữ liệu tổng điều tra) và dữ liệu lớn của khu vực tư nhân (ví dụ dữ liệu điện thoại di động, dữ liệu truyền thông xã hội). Ngoài ra, KOSTAT đã nỗ lực trong việc thiết lập khuôn khổ pháp lý và thể chế và hợp tác với các bên liên quan trong nước và quốc tế.

24. KOSTAT đã thực hiện thành công các dự án như liên kết dữ liệu đánh giá tín dụng cá nhân cũng như dữ liệu điện thoại di động với dữ liệu KOSTAT, phát triển khung hợp tác như tổ chức các diễn đàn dữ liệu lớn để tăng cường liên lạc với các bên liên quan và thiết lập hợp tác quốc tế với Liên hợp quốc và Hà Lan.

25. Tuy nhiên, KOSTAT vẫn phải đối mặt với những thách thức sau: Hạn chế truy cập thông tin cá nhân trong dữ liệu của khu vực tư nhân do luật bảo vệ quyền riêng tư mạnh mẽ; thiếu sự hợp tác từ các nhà cung cấp dữ liệu khu vực tư nhân; thiếu các chuyên gia như các nhà khoa học dữ liệu và dữ liệu lớn của khu vực tư nhân có chất lượng thấp. Trước tất cả những thách thức này, KOSTAT có kế hoạch tiếp tục trao đổi nhiều hơn với các bên liên quan từ các chính trị gia, các nhà hoạch định chính sách, doanh nghiệp, viện nghiên cứu và tổ chức phi chính phủ để giải thích tầm quan trọng của các nguồn dữ liệu mới cho thống kê chính thức và nâng cao năng lực nội bộ trên cơ sở hạ tầng dữ liệu lớn. Ngoài ra, điều quan trọng là liên lạc với các tổ chức quốc tế để giải quyết các vấn đề dữ liệu lớn.

Anh Tuấn (dịch)

Nguồn: Hội nghị các nhà thống kê châu Âu năm 2019 về nguồn dữ liệu mới - khả năng truy cập và sử dụng, https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2019/CES_30_Sem1_Ses1_KoreaE.pdf