

Áp dụng các nguồn dữ liệu thay thế trong thống kê nhà nước - trường hợp của Hungary

Tóm tắt:

Những thách thức thống kê nhà nước phải đối mặt trong thế kỷ XXI là rất đa dạng. Bài viết này trình bày một số phát triển chính sử dụng các nguồn dữ liệu mới tại Cơ quan Thống kê quốc gia Hungary, cụ thể là máy tính tiền trực tuyến - cho doanh thu bán lẻ, dữ liệu camera giám sát giao thông - cho thống kê du lịch, cơ sở dữ liệu của Cơ quan Thuế quốc gia - thống kê lao động, dữ liệu máy quét - cho giá tiêu dùng,... Bằng cách sử dụng hiệu quả các bộ dữ liệu thay thế này, có thể giảm gánh nặng hành chính và tăng tiềm năng thống kê, đồng thời cung cấp thông tin điều chỉnh cho người ra quyết định và cải thiện chất lượng thống kê nhà nước.

I. Giới thiệu

1. Kết quả của cuộc cách mạng dữ liệu là ngày nay có rất nhiều nguồn dữ liệu khác nhau: Dữ liệu thống kê dựa trên các bảng hỏi trong các cuộc điều tra, tổng điều tra, dữ liệu lớn, dữ liệu thông minh, dữ liệu máy, dữ liệu hành chính, dữ liệu riêng tư, v.v... Một mặt, có nhiều trường hợp trong đó các lĩnh vực thống kê dựa trên các nguồn dữ liệu truyền thống, đã đạt đến giới hạn về tính kịp thời, sự phù hợp và tuân thủ yêu cầu giảm gánh nặng cho người trả lời, điều này làm giảm tỷ lệ trả lời. Mặt khác, những dữ liệu mới thường không được tạo ra cho mục đích thống kê nên chất lượng của các nguồn này đôi khi thấp hơn so với các nguồn truyền thống, hoặc việc đo lường chất lượng của chúng cũng đòi hỏi các giải pháp phi truyền thống và việc sử dụng chúng gây ra nhiều lỗ hổng bên ngoài hơn. Do đó, các nhà thống kê có cơ hội lớn hơn bao giờ hết để sử dụng các nguồn dữ liệu này, nhưng các Viện Thống kê quốc gia (NSI) cần phải hài hòa chúng với các tiêu chuẩn chất lượng cao và tích hợp chúng vào hệ thống sản xuất thống kê của họ.

2. Cho đến nay, các cơ quan thống kê có một lợi thế cạnh tranh lớn trong việc truy cập gần như độc quyền vào dữ liệu đầu vào từ cả điều tra thống kê và hồ sơ hành chính. Tình hình này đã thay đổi mạnh mẽ. Trong thời đại toàn cầu hóa, dữ liệu có giá trị thị trường ngày càng tăng. Hầu hết mọi người đều có thể truy cập các nguồn dữ liệu mới, cung cấp các cơ hội mới và đồng thời dẫn đến cạnh tranh với các cơ quan thống kê.

3. Hơn nữa, các xã hội và nền kinh tế đang trải qua những thay đổi quan trọng trong đó yêu cầu các dịch vụ thống kê phức tạp. Nhu cầu người dùng truyền thống cũng đã thay đổi. Nhu cầu liên quan đến chính sách mới, ưu tiên khoa học mới xuất hiện ngày càng nhiều - thúc đẩy nhu cầu mới về thống kê (số hóa, toàn cầu hóa, Mục tiêu phát triển bền vững, v.v...).

4. Quan điểm của các nhóm người dùng chính hiện tại của chúng tôi (những người ra quyết định trong khu vực công và tư nhân, học viện, truyền thông và công chúng nói chung) đối với chất lượng dữ liệu dường như cũng đã thay đổi. Ngày càng có nhiều

người dùng dường như coi trọng tính kịp thời hơn tính chính xác. Họ hy vọng số liệu thống kê sẽ được cung cấp nhanh hơn, vì họ có nhiều nguồn thay thế không chính thức trong tầm tay.

5. Và tất cả, chúng ta đang nắm giữ nhiều lựa chọn. Do đó, cần quản lý dữ liệu tốt hơn, chiến lược, phương pháp tiếp cận dựa trên quy trình, kỹ năng và chiến lược tốt hơn để kiểm soát chi phí, cải thiện quy trình và tính kịp thời trong việc nắm bắt quyền, phù hợp với mục đích thông tin.

6. Là một phần của chiến lược này, chúng ta cần:

- Phương pháp tiếp cận định hướng quy trình, trong đó cần tăng cường kết hợp các nguồn dữ liệu có sẵn và các đăng ký thống kê hoặc hạn chế tiếp cận nguồn dữ liệu đơn lẻ, quản lý rủi ro của quy trình sản xuất thông tin thống kê và giảm bớt thời gian cần thiết để phát hành, mà không làm giảm các khía cạnh chất lượng khác, cụ thể là tính chính xác.

- Thiết lập/khôi phục kho dữ liệu theo thời gian và kịp thời - khối tài sản lớn nhất của NSI - để xử lý hiệu quả nhu cầu dữ liệu ngày càng tăng. Các nguyên tắc về kho dữ liệu, phạm vi dữ liệu, quy tắc cập nhật và lưu trữ cần phải được xác định. Công việc này cũng nhấn mạnh sự cần thiết của các hướng dẫn phương pháp được thiết lập tốt, các tiêu chuẩn siêu dữ liệu và tài liệu chi tiết. Chúng ta cần nhận ra rằng thông tin thu thập không chỉ cho hiện tại mà còn cho tương lai.

- Sắp xếp và kết hợp hiệu quả hơn các nguồn dữ liệu hành chính vào hệ thống thống kê. Cần phải đưa ra khả năng rằng việc sử dụng dữ liệu hành chính có thể không chỉ được thực hiện trong một lĩnh vực thống kê mà phải rộng rãi hơn. Mục tiêu

chính là sử dụng hiệu quả các nguồn dữ liệu tiềm năng. Các nguyên tắc cơ bản của thống kê nhà nước cần được xem xét khi thực hiện nhiệm vụ này.

7. Là một phần của chiến lược dữ liệu, Cơ quan Thống kê quốc gia Hungary (HCSO) đang cố gắng duy trì chất lượng của các bộ sưu tập dữ liệu truyền thống, cũng như kết hợp các nguồn dữ liệu hành chính và thay thế vào các quy trình sản xuất thông tin thống kê nhà nước, do đó được hưởng lợi từ các cơ hội của môi trường bên ngoài thay đổi. Khi các yêu cầu chất lượng cho phép thực hiện và khả thi về mặt phương pháp, HCSO cũng nhằm mục đích thay thế các bảng hỏi truyền thống bằng các nguồn dữ liệu mới. HCSO cam kết tạo điều kiện chuyển hướng sang danh mục dữ liệu đa dạng hơn.

8. Chúng tôi đưa ra một phác thảo ngắn gọn về khung chiến lược ở trên, sau đó chúng tôi triển khai các nguồn dữ liệu thay thế vào hệ thống thống kê của mình. Là một phần của chiến lược dữ liệu của chúng tôi, chúng tôi đổi mới, trong số các hệ thống khác, hệ thống thu thập dữ liệu giao thông biên giới và bán lẻ, sử dụng dữ liệu trực tuyến của máy tính tiền và dữ liệu giao thông đường bộ.

II. Phát triển cụ thể để truy cập danh mục dữ liệu đa dạng hơn

9. Một trong những mục tiêu chiến lược chính của HCSO là giảm gánh nặng hành chính đối với các nhà cung cấp dữ liệu bằng cách đơn giản hóa các câu hỏi và sử dụng dữ liệu hành chính đã có sẵn của các tổ chức khác. Bước đầu tiên, chúng tôi thực hiện kiểm kê tất cả các nguồn dữ liệu hiện có tiềm năng.

10. Bằng cách tập hợp các bộ dữ liệu đã được lấy từ các nguồn hành chính, xem xét

➤➤➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

xem dữ liệu nào có thể được sử dụng thay cho việc thu thập dữ liệu sơ cấp hay không. Chúng tôi phân tích các vấn đề trong việc sử dụng dữ liệu hành chính, cho dù dữ liệu có thể:

(a) Được sử dụng để thay thế cho việc thu thập dữ liệu hiện có;

(b) Thay thế một phần của bộ sưu tập dữ liệu thống kê;

(c) Được sử dụng để xác nhận; hoặc là

(d) Được xuất bản mà không xây dựng một quy trình thống kê phức tạp (ví dụ: Xuất bản thông tin ngay sau khi kiểm tra chất lượng dữ liệu).

11. Bản kiểm kê cho phép chúng tôi phân tích các cách sử dụng lại dữ liệu hiện có để tạo hoặc bổ sung số liệu thống kê mới. Từ 370 bộ sưu tập dữ liệu ở giai đoạn này, chúng tôi thấy có sự thay đổi ở 21 trường hợp năm 2018. Năm 2019, có 20 sửa đổi, trong đó có 14 lần tiếp quản dữ liệu hành chính mới. Phần lớn các tiếp quản dữ liệu mới chủ yếu liên quan đến Tổng điều tra dân số và nhà ở năm 2021. Tuy nhiên, một số trong đó có thể thay thế các bộ sưu tập dữ liệu sơ cấp hoặc được ban hành để đáp ứng các yêu cầu của quy định trong Liên minh châu Âu.

12. Việc đưa các nguồn dữ liệu mới vào hệ thống thống kê chỉ có thể được thực hiện dần dần như là một chức năng đáp ứng các yêu cầu chất lượng. Trong trường hợp này, số liệu thống kê thử nghiệm đóng một vai trò quan trọng trong việc phổ biến kết quả chính của chúng tôi, và sau đó, khi sản phẩm hoàn thiện, nó sẽ trở thành một phần của thống kê nhà nước. Chúng tôi cam kết theo cách tiếp cận từng bước này và thúc đẩy các số liệu thống kê thử nghiệm trong HCSO.

A. Thống kê thu nhập hàng tháng từ nguồn dữ liệu hành chính mới

13. Từ năm 2019 trở đi, việc thu thập dữ liệu thống kê tiền lương lao động hàng tháng có thể được thay thế hoàn toàn bằng các nguồn dữ liệu hành chính có sẵn. Thông tin về thu nhập (theo khái niệm ở Hungary) được tạo ra dựa trên các tờ khai về các khoản đóng góp nhận được từ Cơ quan Thuế và Hải quan quốc gia. Đối với các tổ chức ngân sách, dữ liệu nhận được từ Kho bạc Nhà nước Hungary được sản xuất và công bố, song song với việc hủy bỏ việc thu thập dữ liệu hàng tháng trước đó về lao động. Việc bỏ thu thập dữ liệu hàng tháng đã giảm đáng kể gánh nặng cung cấp dữ liệu xuống còn một nửa (liên quan đến việc thu thập dữ liệu hàng năm về dữ liệu lao động). Dữ liệu hành chính chứa dữ liệu chi tiết hơn đáng kể - ở cấp độ cá nhân/công việc - liên quan đến mọi vị trí và thu nhập. Việc tính toán thông tin cơ bản như vậy, trước đây không thể truy cập được trong bộ sưu tập dữ liệu HCSO, được thực hiện theo cách này. Giá trị thu nhập trung bình, thu nhập ròng dựa trên thuế và các khoản phụ cấp đóng góp thích hợp ở đây. Cơ sở dữ liệu tăng lên cùng với sự sẵn có của thông tin cơ bản - không có sẵn trong các bộ sưu tập dữ liệu trước đó - chẳng hạn như nhân viên, giới tính, tuổi và đặc điểm nghề nghiệp. Các thông tin mới sẽ được giới thiệu dần dần. Thông tin không có trong các tờ khai - như tiền công và tiền lương theo khái niệm SNA, số giờ làm việc, dữ liệu thu nhập thường xuyên chi tiết - có sẵn hàng quý kể từ năm 2019 từ các bộ sưu tập dữ liệu thống kê lao động cho các doanh nghiệp sử dụng ít nhất 5 lao động, tổ chức ngân sách và tổ chức phi lợi nhuận là quan trọng đối với việc làm.

B. Dữ liệu hành chính để tính chỉ số chi phí lao động

14. Dự án này nhằm mục đích nâng cao chất lượng dữ liệu theo tiềm năng sử dụng dữ liệu hành chính. HCSO đang có kế hoạch đổi mới phương pháp ước tính chỉ số chi phí lao động hàng quý (trong phạm vi này là chi phí xã hội và thuế) bằng cách sử dụng dữ liệu hành chính. Tờ khai thuế về các khoản đóng góp cũng chứa thông tin về các khoản thuế mà người sử dụng lao động phải trả (ví dụ: Thuế đóng góp xã hội, các loại thuế khác đã trả sau khi cung cấp các lợi ích cho người lao động).

C. Sử dụng dữ liệu hành chính thông qua dự án của chính phủ ASP2

15. Mô hình Nhà cung cấp dịch vụ ứng dụng (ASP) đã được sử dụng chủ yếu trên toàn thế giới trong lĩnh vực kinh doanh và hành chính công như một giải pháp hiệu quả về chi phí. Ưu điểm của mô hình là người dùng có thể truy cập phần mềm dưới dạng dịch vụ được cung cấp bởi nhà cung cấp dịch vụ trực tuyến từ xa thông qua trình duyệt web đơn giản. Mô hình này cũng có lợi thế về mặt công nghệ và kinh tế cho các thành phố để hỗ trợ nhiều chức năng của họ. Phạm vi của hệ thống ASP của Hungary là đơn giản hóa các thủ tục hành chính của chính quyền địa phương và làm cho minh bạch ngân sách của các hệ thống con của chính quyền địa phương. Mục tiêu xa hơn của đô thị tự trị ASP là cung cấp các giải pháp CNTT hiện đại, tích hợp và hiệu quả về chi phí cho chính quyền địa phương, thúc đẩy hoạt động nội bộ được tiêu chuẩn hóa và cung cấp dịch vụ Chính phủ điện tử dựa trên nền tảng chung cho người dân và doanh nghiệp. Vào năm 2015, Hungary đã khởi động một dự án thí điểm mang tên Thiết lập một trung tâm đô thị ASP, với 55 đô thị tự nguyện tham gia và

hơn 39 đô thị tham gia từ tháng 1 năm 2016. Sau khi thí điểm thành công, Chính phủ đã quyết định khởi động Hệ thống đô thị ASP 2.0 vào năm 2016 để mở rộng hệ thống lên cấp quốc gia trên cơ sở bắt buộc (theo Nghị định của Chính phủ số 257/2016. (VIII. 31.) trên Đô thị tự trị ASP). Tính đến tháng 1 năm 2019, hơn 90% chính quyền địa phương Hungary đang sử dụng hệ thống ASP trung tâm. Lãnh đạo của dự án là Cơ quan Phát triển Công nghệ thông tin Chính phủ - liên kết với Bộ Nội vụ và Kho bạc Nhà nước Hungary. Nhiều cơ quan chính phủ khác có liên quan đến các khía cạnh khác nhau trong sự phát triển hệ thống, bao gồm cả HCSO. Mục tiêu của các cuộc tham vấn chuyên nghiệp liên tục giữa HCSO và các thành viên của dự án ASP là thay thế bảng hỏi thống kê do chính quyền địa phương gửi cho HCSO. Việc truyền dữ liệu tự động từ các hệ thống tích hợp của ASP, giảm gánh nặng hành chính bằng cách loại bỏ nghĩa vụ của họ để thực hiện các cuộc điều tra thống kê cụ thể. Kể từ tháng 2 năm 2019, HCSO có kế hoạch truyền dữ liệu tự động trong trường hợp có hai ấn phẩm thống kê quan trọng từ ASP:

- Báo cáo về các doanh nghiệp có hoạt động thương mại phân phối;
- Một cuộc điều tra đầu tư các đô thị.

D. Phát triển cuộc Khảo sát điều kiện sống và ngân sách hộ gia đình

16. HCSO có các dự án nhằm giảm gánh nặng đáp ứng và cải thiện chất lượng dữ liệu thu nhập. Điều này bao gồm sự phát triển phương pháp của Khảo sát điều kiện sống và ngân sách hộ gia đình (HBS + SILC) về việc sử dụng các công cụ cải tiến trong quá trình thu thập dữ liệu các mặt hàng đã mua, sử dụng máy quét làm hệ thống nhận dạng ký tự trực tuyến để ghi hóa đơn mua hàng. Đối với việc thu thập dữ liệu về dữ liệu thu nhập

➤➤➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

- trong khuôn khổ một dự án tài trợ - việc sử dụng thu nhập của người lao động cho dữ liệu đăng ký thuế đang được thử nghiệm. Các nhà tài trợ từ chối một số nguồn dữ liệu bên ngoài đã được thực hiện do không có phản hồi cao và thường đánh giá thấp thu nhập của người lao động trong Điều tra lực lượng lao động (LFS) trong khuôn khổ dự án tài trợ. Theo báo cáo cuối cùng của dự án, các nguồn dữ liệu đáng tin cậy nhất đã được chứng minh là đăng ký các khoản đóng góp hoàn trả của Cơ quan Thuế. Sau khi tích hợp Thống kê xã hội châu Âu có hiệu lực, thu nhập của người lao động của LFS sẽ được xây dựng theo phương pháp được phát triển trong dự án này.

E. Cải thiện việc sử dụng các nguồn dữ liệu hành chính trong Tổng điều tra năm 2021

17. Tổng điều tra năm 2021 sẽ thể hiện sự cải thiện đáng kể trong việc sử dụng dữ liệu hành chính. Đạo luật tổng điều tra Hungary CI/2018 quy định tên của người trả lời là một phần của dữ liệu bắt buộc phải được thu thập trong cuộc điều tra dân số năm 2021, đây là một thay đổi quan trọng từ các cuộc điều tra ẩn danh năm 2011 và 2001. Tên này, cùng với dữ liệu về địa chỉ, giới tính và ngày sinh, có thể kết nối trực tiếp dữ liệu câu hỏi điều tra dân số với bộ dữ liệu hành chính. Dựa trên kết quả nghiên cứu được hỗ trợ bởi các khoản tài trợ của Eurostat, 13 nguồn dữ liệu hành chính cần thiết cho điều tra dân số đã được xác định và liệt kê trong đạo luật Tổng điều tra. Các tập hợp dữ liệu hành chính sẽ đóng vai trò là cơ sở của một đăng ký thống kê dân số sẽ được sử dụng trong tổng điều tra cho các nhiệm vụ khác nhau. Nó sẽ cung cấp khung tổng điều tra và sẽ giúp kiểm tra phạm vi trong giai đoạn thu thập dữ liệu. Nó cũng sẽ đóng vai trò là bộ dữ liệu được sử dụng để xác

nhận, gán và làm sạch trong quá trình xử lý dữ liệu. Sau tổng điều tra, đăng ký dân số hành chính được bổ sung cùng với dữ liệu bảng hỏi tổng điều tra sẽ là nguồn dữ liệu chính để cung cấp dữ liệu loại tổng điều tra hàng năm và là một phần quan trọng trong việc phát triển số liệu thống kê dân số và xã hội khác. Nó cũng sẽ có thể xây dựng các mẫu dựa trên cá nhân bằng cách sử dụng đăng ký nội bộ.

F. Việc sử dụng Thuế giá trị gia tăng trực tuyến (hóa đơn điện tử)

18. Hệ thống báo cáo Thuế giá trị gia tăng (VAT) đã được thay đổi hoàn toàn kể từ ngày 1 tháng 7 năm 2018. Mục đích của việc giới thiệu hệ thống báo cáo trực tuyến và phát triển hệ thống xử lý dữ liệu là nhằm thúc đẩy nền kinh tế bằng cách hạn chế hoạt động gian lận thuế. Các thay đổi chỉ áp dụng cho các hóa đơn có các thuộc tính sau:

- Số tiền thuế đầu ra trên hóa đơn từ 100.000 HUF trở lên;
- Hóa đơn được phát hành cho một thực thể kinh tế có mã số thuế nội địa và không phải chịu thuế (Đơn giản hóa thuế của doanh nhân).

19. Tóm tắt là một nguồn dữ liệu quan trọng cho HCSO, chủ yếu cho thống kê ngoại thương, bởi vì đây là cơ sở để xác định các đối tác thương mại. Phần Đơn vị lớn (LCU) cũng đang có kế hoạch sử dụng chúng trong khi những dữ liệu này cung cấp cái nhìn sâu sắc về các hoạt động thương mại giữa các doanh nghiệp và chúng có thể được sử dụng để phân tích các số liệu thu được. Do đó, chúng tôi cần tiếp tục nhận thông tin về các hóa đơn được truyền tới hệ thống trực tuyến. Việc tiếp nhận dữ liệu thử nghiệm đang được tiến hành nhằm mục đích đặc tả các biến đó và các ngưỡng đó có được lượng dữ liệu đáng tin cậy có thể quản lý được bằng HCSO.

G. Ước tính dữ liệu đăng ký tiền mặt trực tuyến trong thống kê thương mại bán lẻ

20. Ước tính dữ liệu bán lẻ và phục vụ trong HCSO dựa trên một phương pháp mới, trong đó việc thu thập dữ liệu được bổ sung dữ liệu từ các nguồn hành chính. Dựa trên phương pháp mới, chúng tôi có thể tạo các dữ liệu này dựa trên dữ liệu dân số toàn diện hơn so với thực tiễn trước đó, dẫn đến giảm sai số chọn mẫu. Trên cơ sở này, kết quả của phương pháp mới sử dụng dữ liệu đăng ký tiền mặt trực tuyến, cung cấp một bức tranh đáng tin cậy hơn về doanh số bán lẻ và phục vụ ăn uống so với thực tế trước đây, dựa trên việc chọn mẫu. Sử dụng những dữ liệu này, HCSO đã tăng mức độ bao phủ của các nhà cung cấp dữ liệu trong khi giảm gánh nặng của người trả lời khoảng 80%.

H. Sử dụng dữ liệu camera bên đường để ước tính du lịch cơ bản

21. Mục tiêu chính là cải thiện chất lượng dữ liệu thống kê du lịch. Ngoài các phương pháp hiện có (phương pháp đếm phương tiện giao thông theo thủ công), ngoài Dịch vụ thanh toán phí quốc gia Plc, dữ liệu giao thông bên đường, chúng tôi cũng kiểm tra dữ liệu biến số phương tiện và dữ liệu giao thông phương tiện của Trụ sở Cảnh sát quốc gia Hungary. Sử dụng các nguồn dữ liệu mới này, chúng tôi tạo ra các ước tính đáng tin cậy cho các nhóm nhỏ điển hình (ví dụ: vận chuyển, đi lại hàng ngày), có thể được sử dụng ngoài mục tiêu chính (ước tính chính xác hơn về phân bố quốc tịch của khách du lịch) để xác thực kết quả của việc thu thập dữ liệu về du lịch và các khoản chi tiêu khác của người nước ngoài tại Hungary. Chất lượng của số liệu thống kê du lịch được cải thiện bởi thực tế là các kết quả có thể được sử dụng trong các điều tra thống kê về

du lịch Hungary, làm cơ sở cho định nghĩa về chi tiêu du lịch và được sử dụng để ước tính xuất khẩu và nhập khẩu du lịch để tính toán của tổng sản phẩm trong nước.

I. Sử dụng các nguồn dữ liệu mới cho các cơ sở thương mại và kinh doanh khác

22. Việc thành lập Trung tâm cung cấp dữ liệu du lịch quốc gia thay đổi việc cung cấp dữ liệu thống kê của các cơ sở lưu trú thương mại và kinh doanh khác theo cách cơ bản, vì dữ liệu của họ được chuyển bằng điện tử đến Trung tâm hàng ngày thay vì HCSO. Hiện tại, ứng dụng nhận dữ liệu điện tử (KARÁT) của HCSO đang được lên kế hoạch và thử nghiệm; dữ liệu đầu tiên cho cơ sở thương mại dự kiến vào tháng 8 năm 2019. Dữ liệu cho các cơ sở lưu trú kinh doanh khác sẽ chi tiết hơn với việc tích hợp loại này vào Trung tâm so với việc thu thập dữ liệu hiện tại của HCSO, phù hợp với nhu cầu ngày càng tăng đối với loại thông tin này.

J. Cơ sở hạ tầng y tế quốc gia

23. Cơ sở hạ tầng y tế quốc gia (EESZT) đã được phát triển để tận dụng các cơ hội của y tế bằng cách liên kết các nhà cung cấp dịch vụ y tế và một không gian truyền thông duy nhất cho họ. EESZT giúp cho dữ liệu và tài liệu y tế trở nên dễ dàng hơn, tạo điều kiện quản lý và phân tích dữ liệu hiệu quả hơn bởi các thành viên của ngành chăm sóc sức khỏe. Phối hợp với các chuyên gia của Trung tâm dịch vụ chăm sóc sức khỏe quốc gia, HCSO đang tìm kiếm cơ hội trong tương lai để sử dụng thống kê dữ liệu được lưu trữ trong hệ thống.

K. Phát triển thỏa thuận hợp tác với các đối tác kinh doanh để tiếp quản và sử dụng thông tin quảng cáo bất động sản

➤ ➤ ➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

24. HCSO có kế hoạch xây dựng mô hình hợp tác với các đối tác kinh doanh để sử dụng cơ sở dữ liệu của họ. Chúng tôi mong muốn đạt được sự hợp tác cùng có lợi trong đó HCSO hỗ trợ kiến thức chuyên môn/nghiệp vụ và cung cấp một số loại cơ sở dữ liệu kinh doanh. Đối tác có thể là một trang web bất động sản đã phát triển các kỹ thuật xử lý dữ liệu để tạo thuận lợi cho việc sử dụng dữ liệu thống kê. Để hiểu hơn về cách các trang web bất động sản bao trùm toàn bộ thị trường nhà ở cho thuê, HCSO đã hoàn thành một cuộc điều tra tiêu biểu về nhà cho thuê vào cuối năm 2018. Theo cách này, thông tin phù hợp của cuộc điều tra với dữ liệu kinh doanh có thể có nhiều cách sử dụng cho cả hai bên. Mặc dù Luật Thống kê ở Hungary cho phép HCSO sử dụng cơ sở dữ liệu kinh doanh, nhưng HCSO mong muốn đạt được thỏa thuận có lợi cho cả hai bên. Đối với một đối tác kinh doanh bất động sản, lợi thế lớn nhất có thể là xác nhận cơ sở dữ liệu của họ và ước tính chính xác phạm vi bao phủ của nó, vì trong thị trường nhà đất, thông tin có vai trò cực kỳ quan trọng.

L. Hợp tác với cơ quan hành chính công để hài hòa các lợi ích của hệ thống điện tử được phát triển từ cả hai phía

25. Trong vài năm qua, tất cả các tổ chức quốc gia và các đơn vị hành chính công đã thực hiện các hệ thống điện tử và chuyển sang các giải pháp quản trị điện tử. Song song với hệ thống thu thập dữ liệu điện tử HCSO (được gọi là ELEKTRA), cơ quan hành chính công đã đưa ra một số dự án điện tử để hiện đại hóa và tạo điều kiện thuận lợi cho công việc của các cơ quan xây dựng và cũng để làm cho các nhà xây dựng làm nhiệm vụ rẻ hơn, nhanh hơn và hiệu quả hơn. Số liệu thống kê xây dựng nhà ở được thu thập từ các đơn vị nhà hành chính địa phương thông qua ELEKTRA. Kết nối

ELEKTRA với các hệ thống quản trị tòa nhà điện tử sẽ giúp giảm bớt khối lượng công việc của nhân viên hành chính tại tòa nhà chính quyền địa phương.

M. Tham gia vào công việc của dự án do Trung tâm dịch vụ chăm sóc sức khỏe quốc gia dẫn đầu để số hóa giấy chứng tử

26. Một trong những mục tiêu quan trọng nhất của dự án này là số hóa giấy chứng tử (DC), hiện đang thực hiện dựa trên giấy tờ. Điểm đặc biệt của giấy chứng tử là một số tổ chức (nhà cung cấp dịch vụ chăm sóc sức khỏe, cơ quan nhà nước và thành phố) tham gia hoàn thành các chứng nhận. Hơn nữa, việc điền vào bảng hỏi là cho các mục đích khác nhau và điền thêm một số bản sao trên giấy. Người sử dụng các bản sao, gồm: HCSO, cơ quan đăng ký, công chứng viên (giải quyết vấn đề thừa kế), người thân của người quá cố (hoặc người tổ chức tang lễ), bác sĩ cấp giấy chứng tử, và Sở Y tế Công cộng của Văn phòng Chính phủ. Trong dự án, có thể thiết lập kết nối giữa các giao diện, tích hợp dữ liệu của các hệ thống chăm sóc sức khỏe và các hệ thống khác hỗ trợ các quy trình hành chính. Sự ra đời của kỹ thuật số hóa sẽ cải thiện hơn nữa tính kịp thời và tính chính xác của số liệu thống kê tử vong, giảm gánh nặng hành chính của các nhà cung cấp dữ liệu và giảm đáng kể thời gian nhập dữ liệu. Hệ thống e-DC mới dự kiến sẽ được giới thiệu vào năm 2020.

N. Phát triển hơn nữa việc truyền dữ liệu từ hệ thống đăng ký điện tử và giới thiệu thu thập dữ liệu điện tử về các sự kiện quan trọng hấp dẫn liên quan đến các tổ chức

27. HCSO được quyền thu thập số liệu thống kê dân số trên cơ sở Mục 30 Luật CLV năm 2016 về Thống kê nhà nước. Theo luật,

các nhà đăng ký là một trong những nhà cung cấp dữ liệu quan trọng nhất của các sự kiện quan trọng. Hệ thống đăng ký điện tử được giới thiệu tại Hungary vào ngày 1 tháng 7 năm 2014; sau đó cũng có những thay đổi đáng kể trong việc thu thập dữ liệu thống kê các sự kiện quan trọng (cải thiện nội dung truyền dữ liệu từ hệ thống đăng ký điện tử và phát triển hệ thống CNTT cần thiết để truyền dữ liệu thường xuyên). Hiện tại, trong quá trình đăng ký, HCSO tiếp quản tất cả dữ liệu sự kiện quan trọng có sẵn từ hệ thống đăng ký điện tử. Tuy nhiên, cần cải thiện thêm để truyền bộ sưu tập dữ liệu điện tử hoàn chỉnh của số liệu thống kê các sự kiện quan trọng, vì một số thông tin nhân khẩu học cần cho số liệu thống kê các sự kiện quan trọng không được đưa vào hệ thống đăng ký, ví dụ: Giáo dục, hoạt động kinh tế, nghề nghiệp, dữ liệu sức khỏe của trẻ em mới sinh và bà mẹ. Để kiểm tra các cơ hội phát triển hơn nữa ngoài các nhà đăng ký, sự tham gia của các tổ chức khác - chủ yếu là các tổ chức chăm sóc sức khỏe - sẽ là cần thiết trong dịch vụ dữ liệu. Việc giới thiệu kỹ thuật số hóa hoàn toàn sẽ tiếp tục cải thiện chất lượng, tính kịp thời và tính chính xác của số liệu thống kê các sự kiện quan trọng, giảm gánh nặng hành chính của các nhà cung cấp dữ liệu và thời gian dành cho việc nhập dữ liệu sẽ giảm đi đáng kể.

O. Phát triển việc truyền dữ liệu sinh ở nước ngoài

28. HCSO không ngừng tìm kiếm các cách sử dụng nhiều nguồn dữ liệu hành chính hơn cho thống kê nhà nước. Đối với điều này, chúng tôi phải theo dõi các thay đổi lập pháp quốc gia khác nhau và tìm nguồn dữ liệu mới nào là tốt. Theo Luật sửa đổi kể từ ngày 1 tháng 1 năm 2018 liên quan đến phúc lợi gia đình, hai trong số các hình thức phúc lợi gia đình Hungary (trợ cấp thai sản, hỗ trợ

những người trẻ tuổi bắt đầu cuộc sống), cũng đã có sẵn cho người Hungary sống bên ngoài lãnh thổ Hungary. Đơn xin trợ cấp thai sản sau khi trẻ em là công dân Hungary được sinh ra ở nước ngoài được đánh giá tại Kho bạc Nhà nước Hungary. Mục đích của chúng tôi là tiếp quản dữ liệu có sẵn về trẻ em sinh ra ở nước ngoài có quốc tịch Hungary từ hệ thống trợ cấp thai sản của Kho bạc Nhà nước Hungary. Điều này sẽ cung cấp cho chúng tôi quyền truy cập vào dữ liệu mà hiện nay không có sẵn. Dữ liệu sẽ được nhận hàng năm vào cuối tháng 5 của năm sau năm nộp đơn. Dữ liệu quan trọng nhất được nhận là: Nơi sinh, thời gian sinh, giới tính, quyền công dân, nơi cư trú của con; nơi sinh và thời gian sinh và quốc tịch của mẹ. Sau khi xử lý dữ liệu, một báo cáo cũng sẽ được công bố, để đáp ứng nhu cầu ngày càng tăng về dữ liệu về các ca sinh nở xảy ra ở nước ngoài trong giai đoạn gần đây.

P. Dự án ESS.VIP.BUS.ADMIN để cải thiện việc đo lường di cư quốc tế bằng các nguồn dữ liệu hành chính

29. Như đã thấy ở trên, HCSO sử dụng nhiều nguồn dữ liệu hành chính. Mục đích chung của hành động này là thực hiện các bước tiếp theo trong việc sử dụng các nguồn dữ liệu hành chính để sản xuất số liệu thống kê di cư chính thức. Một trong những trọng tâm chính là mở rộng việc sử dụng dữ liệu của Quỹ Bảo hiểm Y tế quốc gia (NHIF) trong việc đo lường di cư quốc tế. NHIF đã là nhà cung cấp dữ liệu chính trong việc sản xuất số liệu thống kê di cư, hiện tại để đo lường lưu lượng di cư của công dân Hungary. Cơ sở dữ liệu NHIF chưa được sử dụng hết tiềm năng của nó. Chúng tôi dự định sử dụng cơ sở dữ liệu của các công dân không phải là người Hungary và sử dụng các phép đo theo thời gian để đáp ứng nhu cầu ngày càng tăng đối với việc đo lường việc di cư không ngừng

➤➤➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

thay đổi. Các hoạt động sau đây sẽ được thực hiện trong dự án:

- Xây dựng cơ sở dữ liệu phù hợp cho phân tích theo thời gian dựa trên dữ liệu từ NHIF;
- Xác định nhóm người tham gia di cư quốc tế bằng dữ liệu NHIF;
- Xem xét mức độ di cư vòng tròn trong nhóm sử dụng các định nghĩa và khuyến nghị đo lường gần đây của UNECE và Eurostat.

Q. Trao đổi dữ liệu vi mô trong lĩnh vực thống kê di cư giữa Cơ quan Thống kê Áo và HCSO

30. Theo số liệu thống kê có sẵn, Áo là nước đứng thứ hai mà các công dân Hungary di cư đến. Việc đo lường di cư và đặc biệt là đo lường di cư của những người có quyền di cư tự do là một trong những thách thức lớn nhất đối với các cơ quan thống kê trong ESS. Điều này phù hợp với lợi ích của những người di cư đăng ký vào đăng ký hành chính của các quốc gia đến vì đăng ký này được kết nối với một số quyền hoặc lợi ích, nhưng những người này lại hiếm khi hủy đăng ký khi rời khỏi quốc gia ban đầu của họ. Để có bức tranh chi tiết và rõ ràng hơn về sự di cư giữa Áo và Hungary, HCSO đang lên kế hoạch tiến hành trao đổi dữ liệu liên quan đến di cư theo hai giai đoạn. Trong giai đoạn đầu tiên, HCSO đang thay đổi dữ liệu thống kê tổng hợp với Cơ quan Thống kê Áo ở cấp độ NUTS 2 và trong bước thứ hai, HCSO đang lên kế hoạch thực hiện trao đổi dữ liệu vi mô. Việc trao đổi dữ liệu tổng hợp sẽ chứa dữ liệu của tất cả các loại công dân tham gia vào quá trình di cư, nhưng trao đổi dữ liệu vi mô sẽ chỉ chứa dữ liệu của công dân Hungary và Áo. Mục tiêu chính của việc trao đổi dữ liệu này là cải thiện số liệu thống kê di cư dựa trên các nguồn dữ liệu hành chính. Việc trao

đổi dữ liệu vi mô được ESS hỗ trợ mạnh mẽ để cải thiện số liệu thống kê nhà nước. Theo kinh nghiệm của dự án này, HCSO sẽ có thể xác nhận tốt hơn số liệu thống kê di cư quốc tế của mình.

R. Web-scraping trong thống kê giá tiêu dùng

31. Quy trình tính chỉ số giá tiêu dùng ở Hungary hiện không dựa trên dữ liệu được thu thập tự động, mặc dù một số công cụ CNTT đã được phát triển để quét các trang web. Nghiên cứu ban đầu về cách sử dụng dữ liệu giá được quét trên web để phát triển và sản xuất số liệu thống kê nhà nước đã bắt đầu gần đây tại HCSO; các câu hỏi về tính toán trọng số và một số vấn đề khác với các điểm bán hàng và các sản phẩm lấy mẫu vẫn còn là vấn đề mở. Do dữ liệu giá được quét trên web có nhiều thông tin bổ sung về chất lượng sản phẩm hơn dữ liệu điều tra ban đầu, điều này tạo ra cơ hội mới để điều chỉnh chất lượng và tập trung vào khía cạnh tính toán chỉ số này. Với những dữ liệu này, chúng ta sẽ hồi quy theo mô hình Hedonic như một phương pháp điều chỉnh chất lượng. Dựa trên những kinh nghiệm có hạn của HCSO cùng với các giải pháp quét trên web, việc tìm kiếm và phát triển các công cụ CNTT hiệu quả để hỗ trợ các phương pháp quét trên web không phải là điều hiển nhiên và dễ dàng. Điều này đặt ra một thách thức cho toàn bộ dự án.

S. Dữ liệu máy quét trong thống kê giá tiêu dùng

32. Quy trình tính chỉ số giá tiêu dùng ở Hungary hiện nay không dựa trên dữ liệu máy quét. Từ năm 2019, chúng tôi đã có một chương trình giới thiệu dữ liệu mới liên quan đến dữ liệu máy quét trong Chương trình điều tra thống kê quốc gia; do đó, báo cáo là chỉ định bắt buộc đối với các nhà bán lẻ.

Chúng tôi đã thành công trong việc ký kết thỏa thuận hợp tác với chuỗi cửa hàng bán lẻ về việc truyền dữ liệu của họ về dữ liệu doanh số và khối lượng. Trong giai đoạn phát triển hiện nay, thách thức lớn nhất là sự kết nối của dữ liệu được thu thập từ các nguồn dữ liệu truyền thống và các nguồn dữ liệu mới và thực hiện các giải pháp CNTT và phương pháp phù hợp.

III. Kết luận

33. Bài viết nêu bật một số phát triển chính của HCSO để chia sẻ kinh nghiệm sử dụng các nguồn dữ liệu không thông thường để giảm gánh nặng của các nhà cung cấp dữ liệu và cải thiện chất lượng dữ liệu. Để sử dụng hiệu quả các nguồn dữ liệu mới, chúng tôi cần chiến lược dữ liệu không chỉ ở cấp quốc gia thành viên mà còn ở cấp quốc tế.

34. Tình hình thống kê hiện nay được đặc trưng bởi tính năng động mạnh mẽ, nhu cầu thông tin đa dạng và kịp thời ở nhiều cấp độ và thống kê nhà nước cần đáp ứng hiệu quả bằng cách đảm bảo và chuyển đổi hệ thống quản lý dữ liệu. Vì lý do này, NSI cần đổi mới mô hình tác nghiệp để duy trì sự phù hợp trong một thế giới thay đổi nhanh chóng, nơi các công nghệ và truyền thông tiên tiến áp đặt các cải cách cơ cấu đầy thách thức trong quá trình sản xuất thông tin thống kê. Thống kê thử nghiệm là một bước đệm quan trọng trong việc giới thiệu các nguồn dữ liệu mới. Chủ đề ưu tiên cũng rất quan trọng trong quá trình cải cách, do có một khoảng cách ngày càng tăng giữa các yêu cầu bên ngoài và nguồn lực của các tổ chức thống kê. Chúng ta không có tài nguyên để lãng phí, chúng ta phải tối ưu hóa chúng. Chúng ta phải đảm bảo rằng các định hướng chiến lược của chúng ta phản ánh các nhu cầu

quan trọng nhất. Mặt khác, đạt được sự hiện đại hóa là một nhiệm vụ kết hợp các kỹ năng khác nhau. Những thách thức chính trong việc thực hiện chiến lược dữ liệu bao gồm những thay đổi liên quan đến pháp luật phù hợp với việc truy cập và sử dụng dữ liệu, cũng như tiếp tục đảm bảo chất lượng dữ liệu cùng với tính toàn vẹn và sự phù hợp của các phương pháp thống kê trong khuôn khổ phương pháp cải tiến. Hơn nữa, đầu tư vào truyền thông và phổ biến là rất quan trọng cả trong việc tăng văn hóa thống kê giữa người dùng và thúc đẩy sự hợp tác quốc tế trong khi thu hút các bên liên quan trong các quyết định, tạo mối quan hệ đối tác với cộng đồng khoa học, trường đại học và khu vực tư nhân để làm phong phú thêm sức mạnh tổng hợp toàn cầu.

35. Ý nghĩa của chiến lược dữ liệu là việc áp dụng khi bối cảnh bên ngoài đang thay đổi nhanh chóng và khả năng tiếp tục sản xuất thông tin thống kê có chất lượng đặt lên hàng đầu. Nhiều thách thức khác nhau xuất hiện từ sự thay đổi nhu cầu về dữ liệu thống kê, từ sự sẵn có của thông tin đến các công cụ và công nghệ mới. Số liệu thống kê chính thức cần phải tiếp tục thay đổi từ chỉ "đưa ra các con số đơn thuần về số lượng" hướng tới "sáng tạo của kiến thức và giá trị mới cao hơn". Sử dụng các cơ hội hiện tại của Thế kỷ XXI và chiến lược dữ liệu, chúng ta có thể mở rộng chân trời thống kê trong tương lai.

Lan Phương (dịch)

Nguồn: Hội nghị các nhà thống kê châu Âu năm 2019 về nguồn dữ liệu mới - khả năng truy cập và sử dụng, http://www.unecce.org/fileadmin/DAM/stats/documents/ece/ces/2019/ECE_CES_2019_40_1907808E.pdf