

SỬ DỤNG DỮ LIỆU HÀNH CHÍNH PHỤC VỤ CÔNG TÁC THỐNG KÊ TRONG LĨNH VỰC DÂN SỐ CƠ HỘI, THÁCH THỨC VÀ GIẢI PHÁP

Nguyễn Thị Thanh Mai

Phó Vụ trưởng - Vụ Thống kê Dân số và Lao động - TCTK

Công tác thống kê dân số ở Việt Nam

Con người là vốn quý nhất, là động lực quan trọng cho sự thành công và thịnh vượng của mỗi quốc gia. Chính vì vậy, nhu cầu có được các bằng chứng đầy đủ và đáng tin cậy về nguồn lực này luôn là thiết yếu và được Đảng, Nhà nước và Chính phủ đặc biệt coi trọng vì đó là căn cứ để xây dựng hoạch định cũng như các đánh giá tác động của những chính sách, chiến lược phát triển kinh tế - xã hội. Đó là lý do tại sao, trong tổng số 230 chỉ tiêu thống kê quốc gia được quy định trong Luật Thống kê có 16 chỉ tiêu riêng về lĩnh vực dân số và 72 chỉ tiêu liên quan đến dân số.

Nguồn dữ liệu biên soạn các chỉ tiêu thống kê dân số

Về mặt lý thuyết, có hai nguồn dữ liệu có thể được sử dụng để biên soạn các chỉ tiêu thống kê nói chung và chỉ tiêu thống kê dân số nói riêng. Một là, nguồn dữ liệu dựa vào các cuộc Tổng điều tra và điều tra thống kê. Hai là, nguồn dữ liệu dựa vào hồ sơ hành chính. Nguồn dữ liệu dựa vào Tổng điều tra và điều tra thống kê là nguồn thông tin có được dựa trên việc tổ chức thực hiện các cuộc khảo sát, các cuộc điều tra thống kê. Nguồn dữ liệu dựa vào hồ sơ hành chính là nguồn dữ liệu được hình thành trong quá trình thực thi nhiệm vụ quản lý hành chính của các cơ quan nhà nước, được cơ quan quản lý nhà nước ghi chép,

lưu giữ, cập nhật trong các hồ sơ hành chính dạng giấy hoặc dạng điện tử. Hiện nay, để phục vụ công việc quản lý hiệu quả hơn, rất nhiều cơ quan quản lý nhà nước đã đẩy mạnh ứng dụng công nghệ thông tin trong công tác quản lý nên rất nhiều hồ sơ hành chính đã được số hóa và lưu trữ dưới dạng điện tử, vì vậy, nguồn dữ liệu dựa vào hồ sơ hành chính còn được gọi là nguồn dữ liệu dựa vào cơ sở dữ liệu hành chính. Ngoài hai nguồn này, trong Luật Thống kê còn quy định nguồn dữ liệu dựa vào chế độ báo cáo thống kê. Tuy nhiên, về mặt bản chất, chế độ báo cáo thống kê cũng được thực hiện dựa vào nguồn dữ liệu hành chính nói trên.

Trên thực tế, trong lĩnh vực thống kê dân số, hầu hết các chỉ tiêu thống kê đều được biên soạn và tổng hợp trên cơ sở kết quả thu thập được từ các cuộc Tổng điều tra và Điều tra dân số. Nguồn dữ liệu hành chính liên quan đến dân số vẫn chưa được tận dụng để phục vụ hiệu quả công tác thống kê. Lý do chủ yếu là trong giai đoạn trước đây các nguồn dữ liệu hành chính trong lĩnh vực này chưa có hệ thống, hồ sơ dữ liệu chưa được số hóa và được quản lý phân tán tại các địa phương. Ngoài ra, cơ sở hạ tầng công nghệ thông tin hỗ trợ quản lý dữ liệu cũng như kết nối chia sẻ dữ liệu giữa các cơ quan quản lý nhà nước còn hạn chế, các cơ chế về chia sẻ dữ liệu cũng chưa

được xây dựng. Chính vì vậy, các nguồn dữ liệu hành chính vẫn chưa được tận dụng cho công tác thống kê nói chung và công tác thống kê dân số nói riêng. Vì vậy, việc biên soạn chỉ tiêu thống kê dân số vẫn hoàn toàn dựa vào nguồn dữ liệu từ Tổng điều tra và điều tra dân số.

Phải khẳng định rằng, số liệu thống kê tính toán từ kết quả Tổng điều tra và điều tra dân số đã giúp Việt Nam có được bức tranh tổng quan về tình hình dân số và các đặc điểm nhân khẩu học của dân số, làm cơ sở để Việt Nam xây dựng và thực hiện nhiều chiến lược và chính sách quan trọng.

Tuy nhiên, việc thu thập dữ liệu dựa vào các cuộc điều tra thống kê thời gian qua cũng đã bộc lộ một số hạn chế như: Chi phí cao; Số liệu chụp ảnh thông tin vào thời điểm nhất định trong giai đoạn trước khi tiến hành điều tra; Mức độ đại diện thường chỉ đến cấp tỉnh. Một số số liệu tính toán từ kết quả Tổng điều tra thì đại diện ở cấp thấp hơn nhưng tần suất 10 năm/lần; Rủi ro về tỷ lệ không trả lời ngày càng cao.

Cơ hội sử dụng dữ liệu hành chính phục vụ công tác thống kê trong lĩnh vực dân số

Cuộc cách mạng công nghệ lần thứ tư đã đem lại cho nhiều quốc gia cơ hội tốt trong việc số hóa các nguồn dữ liệu hành chính phục vụ công việc quản lý của các bộ, ngành. Tại Việt Nam, các nguồn dữ liệu hành chính có thể sử dụng trong lĩnh vực thống kê

dân số được lấy từ Cơ sở dữ liệu quốc gia về dân cư; Cơ sở dữ liệu hộ tịch điện tử; Kho dữ liệu chuyên ngành dân số; Cơ sở dữ liệu chuyên ngành bảo hiểm... Chính vì vậy, nghiên cứu khả năng tận dụng cơ sở dữ liệu hành chính phục vụ công tác thống kê nói chung và thống kê dân số nói riêng không chỉ có ý nghĩa giúp nâng cao hiệu quả công tác thống kê dân số mà còn phù hợp với định hướng phát triển ngành thống kê giai đoạn 2021-2030.

Đối với cơ sở dữ liệu quốc gia về dân cư: Đây là tập hợp thông tin cơ bản của tất cả công dân Việt Nam được chuẩn hóa, số hóa, lưu trữ, quản lý bằng cơ sở hạ tầng thông tin để phục vụ quản lý nhà nước và giao dịch của cơ quan, tổ chức, cá nhân. Cơ sở dữ liệu quốc gia về dân cư là cơ sở dữ liệu dùng chung, do Bộ Công an quản lý. Những thông tin được thu thập và cập nhật vào Cơ sở dữ liệu quốc gia về dân cư gồm: Họ, chữ đệm và tên khai sinh; Ngày, tháng, năm sinh; Giới tính; Nơi đăng ký khai sinh; Quê quán; Dân tộc; Tôn giáo; Quốc tịch; Tình trạng hôn nhân; Nơi thường trú; Nơi ở hiện tại; Nhóm máu (không bắt buộc); Họ và tên, số định danh cá nhân, quốc tịch của cha, mẹ, vợ, chồng hoặc người đại diện hợp pháp; Họ và tên, số định danh cá nhân của chủ hộ, quan hệ với chủ hộ; Ngày, tháng, năm chết/ mất tích...

Hiện nay Việt Nam đã xây dựng thành công và chính thức vận hành cơ sở dữ liệu quốc gia về dân cư từ ngày 22/6/2021. Đến thời điểm hiện tại, về cơ bản, cơ sở dữ liệu quốc gia về dân cư đã cập nhật tương đối đầy đủ thông tin về toàn bộ công dân Việt Nam. Đây là nguồn dữ liệu, được thu thập, cập nhật thường xuyên trên nguyên tắc “đúng, đủ, sạch, sống”.

Trong cơ sở dữ liệu quốc gia về dân cư, có 17 nội dung nội dung

thông tin được thu thập, là các trường thông tin cố định và được thực hiện theo quy định của Luật cư trú. Căn cứ để thu thập thông tin là: Sổ sách quản lý về cư trú, tàng thư hồ sơ hộ khẩu, tàng thư căn cước công dân; Cơ sở dữ liệu về cư trú, Cơ sở dữ liệu căn cước công dân, Cơ sở dữ liệu hộ tịch; Cơ sở dữ liệu quốc gia, cơ sở dữ liệu chuyên ngành khác; Thu thập từ công dân. Việc thu thập và cập nhật thông tin trong cơ sở dữ liệu quốc gia về dân cư đòi hỏi thông tin phải được kiểm tra và đảm bảo tính chính xác. Trường hợp không thống nhất về nội dung thông tin thì cơ quan có thẩm quyền có trách nhiệm kiểm tra tính pháp lý của các thông tin trước khi thu thập, cập nhật vào Hệ thống. Người cung cấp, thu thập và cập nhật thông tin phải chịu trách nhiệm pháp lý.

Nếu tận dụng được nguồn thông tin từ cơ sở dữ liệu quốc gia về dân cư để biên soạn các chỉ tiêu thống kê quốc gia về dân số thì ngành Thống kê có thể cung cấp thông tin cho người sử dụng với mức độ đại diện lớn hơn (có thể đến cấp xã) và tần suất nhiều hơn (6 tháng, quý, tháng, thậm chí là hàng ngày nếu dữ liệu được chia sẻ tức thời khi phát sinh sự kiện). Không những thế, nguồn thông tin này cũng sẽ rất có giá trị nếu áp dụng để phục vụ công tác phân chia địa bàn, chọn mẫu hộ cho các cuộc điều tra mẫu và thực hiện Tổng điều tra dân số và nhà ở theo phương thức hiện đại, tiết kiệm và hiệu quả.

Đối với cơ sở dữ liệu hộ tịch điện tử: Đây là cơ sở dữ liệu ngành Tư pháp, được lập trên cơ sở tin học hóa công tác đăng ký hộ tịch, dựa trên căn cứ Luật hộ tịch năm 2014 và triển khai căn cứ theo Đề án “Cơ sở dữ liệu hộ tịch điện tử toàn quốc”, do Bộ Tư pháp quản lý. Tính đến ngày 19/8/2021, tất cả 63 tỉnh, thành phố trực thuộc Trung ương

đã tham gia triển khai, áp dụng và cập nhật thông tin vào Cơ sở dữ liệu hộ tịch điện tử phục vụ công tác đăng ký, quản lý hộ tịch tại địa phương.

Cơ sở dữ liệu hộ tịch điện tử hiện nay cũng đã được kết nối với Cơ sở dữ liệu quốc gia về dân cư để thực hiện đăng ký khai sinh, cấp Sổ định danh cá nhân và đồng bộ các thông tin thay đổi, cải chính, bổ sung hộ tịch, xác định lại dân tộc và các thay đổi hộ tịch khác trên thông tin khai sinh của công dân đã có Sổ định danh cá nhân.

Thông tin trong Cơ sở dữ liệu hộ tịch điện tử được hình thành trước tiên từ các thông tin hộ tịch của cá nhân được xác lập khi đăng ký khai sinh. Tiếp theo đó, các thông tin hộ tịch khác của cá nhân (như đăng ký kết hôn; ghi nhận ly hôn; đăng ký khai tử và các thông tin hộ tịch khác) tiếp tục được cập nhật vào thông tin hộ tịch của cá nhân, từng bước hình thành “bức tranh” tổng thể về thông tin hộ tịch của mỗi công dân.

Cơ quan thống kê cần nghiên cứu tận dụng cơ sở dữ liệu hộ tịch điện tử để biên soạn các chỉ tiêu thống kê dân số như: Tổng tỷ suất sinh; tỷ suất sinh thô, tỷ số giới tính khi sinh, tỷ lệ trẻ em dưới 5 tuổi đã đăng ký khai sinh; Tỷ suất chết thô, tỷ suất chết của trẻ em dưới 1 tuổi, tỷ suất chết của trẻ em dưới 5 tuổi, tỷ suất chết mẹ, tỷ suất chết đặc trưng theo tuổi, tuổi thọ bình quân; tình trạng hôn nhân của dân số, tuổi kết hôn trung bình lần đầu. Cơ sở dữ liệu hộ tịch điện tử đồng thời có thể được sử dụng để biên soạn báo cáo thống kê hộ tịch theo khuyến nghị của Liên hợp quốc. Cũng giống như cơ sở dữ liệu dân cư quốc gia, thông tin thống kê biên soạn từ nguồn dữ liệu hộ tịch điện tử có thể công bố với tần suất cao cùng với các phân tổ chi tiết ở phạm vi sai số không đáng kể.

Việc sử dụng cơ sở dữ liệu hành chính cho công tác thống kê được đánh giá là sẽ mang lại nhiều lợi ích. Thứ nhất, đây là một nguồn dữ liệu liên tục và lâu dài; có thể tạo ra dữ liệu ở cấp hành chính thấp nhất với tần suất công bố thông tin không giới hạn. Dữ liệu được thu thập trong cơ sở dữ liệu dân cư quốc gia và cơ sở dữ liệu hộ tịch điện tử không có sai sót nhớ lại, không có sai sót chọn mẫu, do đó việc ước lượng thông tin từ nguồn dữ liệu này không bị chệch. Bên cạnh đó, việc sử dụng cơ sở dữ liệu hành chính có chi phí thấp nhất, tiết kiệm đáng kể ngân sách nhà nước.

Việc sử dụng cơ sở dữ liệu hành chính cho công tác thống kê được cho là rất khả quan vì Cơ sở dữ liệu quốc gia về dân cư dân cư và Cơ sở dữ liệu hộ tịch điện tử đã được xây dựng và tin học hóa thành công. Mức độ đăng ký khai sinh cao (98%) và có sự kết nối chặt chẽ giữa Cơ sở dữ liệu hộ tịch điện tử với Cơ sở dữ liệu quốc gia về dân cư trong việc đăng ký khai sinh, cấp Sổ định danh cá nhân và đồng bộ các thông tin thay đổi, cải chính,... trên thông tin khai sinh của công dân đã có Sổ định danh cá nhân. Bên cạnh đó, tất cả các trung tâm đăng ký đều đã có phần mềm hỗ trợ đăng ký và số hóa dữ liệu; dữ liệu được chuyển đến máy chủ trung tâm theo thời gian phát sinh sự kiện. Bên cạnh đó là những văn bản pháp lý cho phép Tổng cục Thống kê khai thác cơ sở dữ liệu hành chính cho công tác thống kê theo tinh thần của Luật Thống kê và hướng dẫn tại Nghị định số 94/2022/NĐ-CP. Cụ thể, hành lang pháp lý về việc khai thác và chia sẻ thông tin trong Cơ sở dữ liệu quốc gia về dân cư cũng đã được quy định rõ tại Khoản 2, Điều 10 Luật Căn cước công dân và được hướng dẫn cụ thể tại Điều 8, Nghị định số 137/2015/NĐ-CP. Hành lang pháp lý về việc khai thác và chia sẻ

thông tin trong Cơ sở dữ liệu hộ tịch điện tử được quy định tại Điều 8, Nghị định 87/2020/NĐ-CP. Hơn nữa, Tổng cục Thống kê (GSO) đã chứng minh được kinh nghiệm và khả năng xử lý dữ liệu quy mô lớn (như dữ liệu của Tổng điều tra dân số và nhà ở, dữ liệu các cuộc điều tra dân số giữa kỳ) đảm bảo nguyên tắc bảo mật thông tin.

Việc sử dụng cơ sở dữ liệu hành chính cho công tác thống kê sẽ là nền tảng vững chắc để Tổng cục Thống kê xây dựng các phương án đổi mới Tổng điều tra dân số theo hướng hiện đại, tiết kiệm, phù hợp với tinh thần của Quyết định số 06/QĐ-TTg ngày 6/1/2022 của Thủ tướng Chính phủ về việc phê duyệt Đề án Phát triển ứng dụng dữ liệu về dân cư, định danh và xác thực điện tử phục vụ chuyển đổi số quốc gia giai đoạn 2022-2025, tầm nhìn đến năm 2050.

Tuy nhiên, hiện nay vẫn tồn tại những rào cản khiến việc sử dụng cơ sở dữ liệu hành chính cho công tác thống kê cũng gặp không ít khó khăn. Đó là: (i) Tầm quan trọng của việc sử dụng cơ sở dữ liệu hành chính như là nguồn dữ liệu quan trọng cho công tác thống kê vẫn chưa nhận được sự quan tâm thích đáng của lãnh đạo quản lý cấp cao; (ii) Cơ chế hợp tác chia sẻ dữ liệu giữa Bộ Công an, Bộ Tư pháp với Tổng cục Thống kê cũng chưa được thiết lập; (iii) Tổng cục Thống kê cũng chưa có nhiều kinh nghiệm trong việc hợp tác, chuyển giao cũng như sử dụng hiệu quả các nguồn dữ liệu được chuyển giao để phục vụ công tác thống kê gồm cả biên soạn số liệu và sử dụng để thu thập thông tin. (iv) Tâm lý ngại thay đổi, sợ mất dữ liệu vẫn còn tồn tại (ở cả trong cơ quan thống kê và ở các Bộ ngành).

Nhưng dù thế nào thì chúng ta cũng cần xác định Cơ sở dữ liệu quốc gia về dân cư và Cơ sở dữ liệu hộ tịch điện tử là hai nguồn dữ liệu

hành chính quan trọng cần được tận dụng để tính toán các chỉ tiêu thống kê trong lĩnh vực dân số. Trước mắt cần ưu tiên thời gian và nguồn nhân lực cần thiết để phối hợp với Bộ Công an, Bộ Tư pháp xây dựng cơ chế phối hợp và chia sẻ thông tin. Tăng cường học tập kinh nghiệm quốc tế trong việc chia sẻ và khai thác dữ liệu hành chính cho công tác thống kê. Triển khai các hoạt động thí điểm chia sẻ thông tin từ trên phạm vi nhỏ, trên cơ sở đó rút kinh nghiệm và triển khai nhiệm vụ trên phạm vi lớn hơn và trên toàn quốc, đảm bảo vừa thực hiện tốt nhiệm vụ thống kê và công bố số liệu, vừa kiểm soát, lưu trữ, bảo mật, bảo đảm an toàn thông tin đối với nguồn dữ liệu liên quan đến mỗi người dân.

Tóm lại, sự thành công trong việc xây dựng, số hóa và đưa vào sử dụng các cơ sở dữ liệu hành chính, trong đó có Cơ sở dữ liệu quốc gia về dân cư; Cơ sở dữ liệu hộ tịch điện tử là điều kiện thuận lợi cho việc chia sẻ và hỗ trợ khai thác nguồn dữ liệu này phục vụ công tác thống kê. Việc sử dụng cơ sở dữ liệu hành chính trong công tác thống kê là xu thế tất yếu của thống kê hiện đại. Điều này sẽ góp phần nâng cao năng lực, tận dụng, phát huy nguồn lực của các cơ quan trong việc quản lý đảm bảo duy trì chất lượng đầu vào của của các cơ sở dữ liệu hành chính, vừa đảm bảo quyền và lợi ích hợp pháp của người dân, vừa góp phần nâng cao chất lượng nguồn dữ liệu, vừa đảm bảo sử dụng nguồn dữ liệu hợp lý, hiệu quả và tiết kiệm. Về cơ bản, hành lang pháp lý và các điều kiện cơ bản của việc chia sẻ các cơ sở dữ liệu này phục vụ cho công tác thống kê đã được thiết lập. Tổng cục Thống kê sẵn sàng đổi mới và cũng chứng tỏ năng lực về tiếp nhận, quản lý, bảo mật dữ liệu theo đúng quy định pháp luật./.