

## TĂNG CƯỜNG ỨNG DỤNG THU THẬP DỮ LIỆU TRÊN THIẾT BỊ DI ĐỘNG VÀ WEBFORM TRONG ĐIỀU TRA THỐNG KÊ Ở VIỆT NAM

*ThS. Cao Quang Thành\**

### **Tóm tắt:**

*Hoạt động điều tra thống kê do ngành Thống kê tổ chức diễn ra trong bối cảnh Cách mạng công nghiệp 4.0 và chuyển đổi số ở Việt Nam. Cách mạng công nghiệp 4.0 và chuyển đổi số thúc đẩy việc sử dụng công nghệ thông tin - truyền thông, kỹ thuật số, Internet vạn vật kết nối (IoT), trí tuệ nhân tạo (AI), dữ liệu lớn... vào quản lý điều hành, hoạt động sản xuất kinh doanh của các cơ quan, tổ chức, cá nhân. Tận dụng các thành tựu của Cách mạng công nghiệp 4.0 và thực hiện chuyển đổi số là một nhiệm vụ quan trọng của ngành Thống kê. Bên cạnh đó Cách mạng công nghiệp 4.0 và chuyển đổi số tạo ra nhiều cơ hội để đổi mới hoạt động thu thập dữ liệu thống kê, đặc biệt là qua các ứng dụng trên thiết bị di động, webform trong điều tra thống kê.*

*Cũng kể từ năm 2020 đại dịch Covid-19 đã và đang làm thay đổi phương thức làm việc và giao tiếp truyền thống giữa người với người của toàn thế giới, trong đó có Việt Nam. Thực tế này đòi hỏi ngành Thống kê cần phải đẩy mạnh thu thập dữ liệu thống kê bằng phương pháp trực tuyến.*

### **1. Thuận lợi, thách thức trong xây dựng và ứng dụng thu thập dữ liệu trên thiết bị di động và webform trong điều tra thống kê**

• Việc thu thập dữ liệu thống kê bằng ứng dụng trên thiết bị di động và webform đem lại lợi ích to lớn. Thực tế cho thấy, việc ứng dụng thu thập dữ liệu thống kê trên thiết bị di động và webform sẽ đẩy nhanh tiến độ thu thập, kịp thời và giải quyết được nhiều vấn đề hơn so với cách thu thập truyền thống trước đây (thu thập bằng phiếu giấy) như: giảm thiểu gánh nặng cho người trả lời, điều tra viên; giảm thiểu nhiều khâu xử lý dữ liệu thống kê; giảm thiểu được một lượng

lớn nguồn nhân lực để số hóa dữ liệu thu thập được; giảm bớt công tác vận tải, trung chuyển... Dữ liệu thống kê thu thập từ ứng dụng trên thiết bị di động và webform đảm bảo hơn về độ chính xác, tính khách quan, cũng như thời gian thu thập nhanh hơn. Đồng thời nâng cao tính toàn vẹn của dữ liệu thống kê.

• Tiếp cận và nghiên cứu về công nghệ dữ liệu mới sẽ cho chúng ta có thêm phương án giải quyết việc thu thập dữ liệu thống kê, xử lý và đối phó với những thách thức sản xuất dữ liệu thống kê trong nước và quốc tế. Tuy nhiên việc áp dụng xây dựng ứng dụng trên thiết bị di động và webform cũng cần nhiều thời gian để học hỏi, xây dựng và thử nghiệm.

\* Cục phó Cục Thu thập dữ liệu và Ứng dụng công nghệ thông tin thống kê

## ➤➤➤ THỐNG KÊ VÀ CUỘC SỐNG

- Nghiên cứu phát triển thu thập dữ liệu thống kê trên thiết bị di động và webform thì cần phải có cơ sở hạ tầng CNTT hiện đại, đáp ứng các yêu cầu xử lý khối lượng dữ liệu lớn đồng thời truyền tải lên hệ thống máy chủ và xử lý trong thời gian ngắn.

- Xây dựng ứng dụng trên thiết bị di động và webform sẽ mở ra một thời kỳ mới về áp dụng công nghệ thông tin trong công tác thống kê của ngành. Xây dựng kiến thức mới cho đội ngũ cán bộ công nghệ thông tin. Tuy nhiên cũng yêu cầu cao hơn đối với đội ngũ công nghệ thông tin của ngành. Đội ngũ cán bộ này cần phải nhanh, nhạy trong lĩnh vực công nghệ thông tin, luôn tìm hiểu công nghệ mới, đòi hỏi phải có trình độ cao trong xây dựng ứng dụng.

- Áp dụng thu thập dữ liệu thống kê trên thiết bị di động và webform mang lại tiến độ thu thập nhanh hơn, đáp ứng được nhu cầu cấp thiết của việc cần dữ liệu thống kê để phân tích sự phát triển của kinh tế, dự báo tình hình kinh tế - xã hội... Nhưng bên cạnh đó thì yêu cầu về đồng bộ nguồn lực và vật lực, quy trình thu thập dữ liệu cũng phải thay đổi.

### **2. Kết quả của việc xây dựng và ứng dụng thu thập dữ liệu thống kê bằng thiết bị di động và webform của một số cuộc điều tra và tổng điều tra**

Ngành Thống kê Việt Nam đã và đang có những bước chuyển mình rõ rệt, từ quyết tâm mạnh mẽ của đội ngũ lãnh đạo ngành đến ý thức của mỗi cá nhân hướng tới việc đẩy mạnh ứng dụng công nghệ thông tin trong mọi lĩnh vực từ quản lý điều hành đến chuyên môn nghiệp vụ.

Cùng với sự phát triển của khoa học công nghệ, để tận dụng những lợi thế về công nghệ thông tin đáp ứng các yêu cầu về

chất lượng, đảm bảo thời gian và sự minh bạch, hiệu quả về chi phí... của dữ liệu thống kê, các hình thức thu thập dữ liệu điều tra đã được cải tiến theo hướng ứng dụng công nghệ thông tin. Các phương pháp thu thập dữ liệu có ứng dụng công nghệ thông tin bao gồm: Phỏng vấn trực tiếp, ghi chép bằng thiết bị điện tử thông minh (Computer Assisted Personal Interviewing, viết tắt là CAPI), phỏng vấn bằng điện thoại ghi chép bằng thiết bị điện tử thông minh (Computer Assisted Telephone Interviewing, viết tắt là CATI); tự cung cấp dữ liệu bằng thiết bị điện tử thông minh (Computer Assisted Self-Interviewing, viết tắt là CASI), phỏng vấn bằng tai nghe và tự cung cấp dữ liệu bằng thiết bị điện tử thông minh (Audio Capi, viết tắt là ACASI); điều tra trực tuyến bằng webform.

Trong vài năm trở lại đây, các đơn vị nghiệp vụ thuộc Tổng cục và một số Cục Thống kê địa phương đã đẩy mạnh nghiên cứu, thử nghiệm ứng dụng một số phương thức mới trong thu thập dữ liệu như: Sử dụng thiết bị điện tử thông minh (CAPI) trong điều tra giá tiêu dùng, lao động - việc làm, điều tra trực tuyến bằng webform trong điều tra doanh nghiệp; ứng dụng phần mềm blockchaine điều tra chăn nuôi; ứng dụng phần mềm điều tra chỉ số giá tiêu dùng, chỉ số giá bất động sản bằng dữ liệu lớn hay ứng dụng bài giảng điện tử trong công tác tập huấn điều tra mức sống dân cư...

Đặc biệt, việc áp dụng hình thức thu thập dữ liệu bằng CAPI trong các cuộc điều tra mẫu vài năm gần đây và mới nhất là trong Tổng điều tra dân số và nhà ở 2019 cho thấy việc sử dụng thiết bị điện tử di động như máy tính bảng, điện thoại thông minh để thu thập dữ liệu sẽ làm tăng chất lượng dữ liệu do kiểm soát được các lỗi logic trong quá trình thu thập dữ liệu (khi điều tra viên tiến

hành phỏng vấn tại hộ, phần mềm trong thiết bị sẽ phát hiện những lỗi sai logic trong quá trình hộ trả lời phỏng vấn và có cảnh báo để điều tra viên kiểm tra lại dữ liệu kỹ hơn), giám sát được việc điều tra viên thống kê có thực tế đến hộ phỏng vấn hay không; bên cạnh đó tiết kiệm được thời gian hoàn thành phiếu điều tra (điều tra viên kết thúc phỏng vấn tại hộ, dữ liệu thu thập được sẽ gửi thẳng lên hệ thống máy chủ lưu trữ và xử lý dữ liệu), nhờ đó rút ngắn thời gian xử lý và công bố thông tin (do không phải làm sạch, vận chuyển phiếu và nhập dữ liệu trên phiếu vào hệ thống như áp dụng việc thu thập dữ liệu cũ trên phiếu giấy). Phương pháp này sẽ tiết kiệm một khoản kinh phí lớn để in phiếu giấy, vận chuyển, thuê nhân công nhập tin và các chi phí liên quan khác. Đã có trên 99,9% hộ dân cư được điều tra bằng hình thức CAPI trong Tổng điều tra dân số và nhà ở 2019.

Tiếp nối thành công từ Tổng điều tra dân số và nhà ở năm 2019, điều tra thu thập dữ liệu về thực trạng kinh tế - xã hội của 53 dân tộc thiểu số năm 2019 cũng ứng dụng CNTT trong tất cả các khâu điều tra. Và điểm đặc biệt hơn trong cuộc điều tra này là việc kết nối các dữ liệu sẵn có của Tổng điều tra dân số và nhà ở năm 2019. Có khoảng hơn 20 trường dữ liệu từ kết quả của Tổng điều tra dân số và nhà ở năm 2019 được kết nối với chương trình điều tra thu thập dữ liệu về thực trạng kinh tế - xã hội của 53 dân tộc thiểu số năm 2019. Các dữ liệu này được kiểm tra, đối chiếu trong quá trình thu thập dữ liệu tại địa bàn của điều tra thu thập dữ liệu về thực trạng kinh tế - xã hội của dân tộc thiểu số năm 2019 và được sử dụng để tổng hợp số liệu và phân tích báo cáo kết quả điều tra. Đây là lần đầu tiên ngành thống kê thực hiện kết nối chuỗi dữ liệu mảng (panel data) ngay trong quá trình

thu thập dữ liệu tại địa bàn giúp thuận tiện cho quá trình điều tra, nâng cao chất lượng thông tin, giảm bớt thời gian để thu thập các thông tin đã sẵn có.

Có thể thấy, áp dụng hình thức thu thập dữ liệu bằng CAPI và webform là một bước tiến trong cải tiến phương pháp điều tra thống kê, giúp nâng cao chất lượng số liệu, giảm thời gian xử lý và công bố kết quả và phù hợp với xu hướng phát triển chung của thống kê trên thế giới và trong khu vực.

Ứng dụng phiếu điều tra điện tử: Năm 2020, phương án điều tra doanh nghiệp đã có sự thay đổi để phù hợp với tình hình mới. Đó là ứng dụng triệt để công nghệ thông tin trong thu thập, xử lý, tổng hợp, phân tích dữ liệu điều tra doanh nghiệp. Theo đó, khâu thu thập dữ liệu được thực hiện bằng phương pháp điều tra trực tuyến, doanh nghiệp cung cấp thông tin vào bảng hỏi điện tử trên website. Đây là bước cải tiến đột phá tạo thuận lợi cho doanh nghiệp chủ động cung cấp thông tin, tiết kiệm thời gian, công sức.

Từ kết quả đạt được và những kinh nghiệm rút ra trong triển khai thu thập dữ liệu trực tuyến của các cuộc điều tra, tổng điều tra trước. Tổng cục Thống kê cũng đã rút ra được những bài học kinh nghiệm về phát triển ứng dụng công nghệ thông tin, cũng như những kinh nghiệm về xử lý trong thời gian triển khai thu thập dữ liệu. Dựa trên những kinh nghiệm và bài học quý giá đó và phát huy những ưu điểm, Tổng cục Thống kê đang dần áp dụng triệt để việc ứng dụng công nghệ thông tin vào tất cả các khâu thực hiện cho các cuộc điều tra và tổng điều tra.

### **3. Một số vấn đề cần lưu ý khi ứng dụng CAPI và webform để thu thập dữ liệu trong điều tra thống kê**

*Chuẩn bị đầy đủ:* Phải dành đủ thời gian

## ➤➤➤ THÔNG KÊ VÀ CUỘC SỐNG

để thiết kế và soạn thảo bảng câu hỏi điện tử cũng như kiểm tra tổng thể và gỡ lỗi phần mềm. Đặc biệt là bảng câu hỏi của các cuộc điều tra phức tạp, vì vậy điều quan trọng là phải bảo đảm rằng các bước nhảy và logic của dữ liệu phải hoạt động chính xác trước khi triển khai.

Lựa chọn thiết bị, phần mềm nền tảng để phát triển ứng dụng CAPI và webform

Việc lựa chọn phần mềm nền tảng để xây dựng ứng dụng cũng như việc lựa chọn thiết bị là một yếu tố rất quan trọng để triển khai được công tác thu thập dữ liệu. Ở Việt Nam, tỷ lệ người dùng thiết bị di động có hệ điều hành Android chiếm tỷ trọng cao, vì vậy nên chú trọng việc xây dựng phần mềm trên

*Lựa chọn và đào tạo điều tra viên:* Việc lựa chọn thích hợp và đào tạo kỹ lưỡng điều tra viên là rất quan trọng cho sự thành công của chương trình. Điều tra viên được lựa chọn ngoài kinh nghiệm về phỏng vấn thì cũng cần kỹ năng sử dụng thiết bị di động thông minh.

*Xây dựng và bảo đảm hạ tầng CNTT:* Hệ thống phải bảo đảm đủ năng lực cho hoạt động điều tra thu thập dữ liệu được ổn định và thông suốt. Tại Việt Nam, các cuộc điều tra đều có yêu cầu cao về bảo mật dữ liệu. Do đó hệ thống máy chủ lưu trữ cần được trang bị hệ thống bảo đảm an toàn an ninh thông tin.

### 4. Đề xuất quy trình xây dựng ứng dụng thu thập dữ liệu trên thiết bị di động và webform trong điều tra thống kê

Quy trình xây dựng ứng dụng thu thập dữ liệu trên thiết bị di động gồm 10 bước sau:



Cụ thể các bước của quy trình như sau:

Bước 1: Nghiên cứu phương án điều tra

Nghiên cứu phương án điều tra nhằm giúp cho lập trình viên có cái nhìn tổng quan

về cuộc điều tra mà ứng dụng sẽ được phát triển, đặc biệt là có được thông tin về những nội dung cần thiết trực tiếp cho việc xây dựng ứng dụng; đó là:

- Đối tượng, đơn vị, phạm vi điều tra: Giúp cho việc xây dựng ứng dụng thu thập dữ liệu

- Nội dung điều tra và thiết kế phiếu điều tra: Đây là vấn đề đóng vai trò then chốt trong xây dựng ứng dụng nên cần được nghiên cứu kỹ lưỡng trước khi bắt tay xây dựng ứng dụng.

- Phương pháp xử lý thông tin và tổng hợp kết quả điều tra: Giúp xác định cách thức dữ liệu được thu thập sẽ được xử lý như thế nào; cách thức tổng hợp kết quả điều tra (có phải xây dựng phần mềm tổng hợp kết quả hay không và nếu có thì có được tích hợp với ứng dụng thu thập dữ liệu trên thiết bị di động)

- Kế hoạch điều tra: Giúp cho việc lập và kiểm soát tiến độ thực hiện xây dựng ứng dụng gắn với tiến độ chung của cuộc điều tra.

Bước 2: Xác định đặc điểm của dữ liệu, thiết lập mô hình chức năng và luồng dữ liệu

Trên cơ sở nghiên cứu phương án điều tra xác định đặc điểm của dữ liệu cần thu thập như: định dạng dữ liệu của các câu hỏi; các câu hỏi và lựa chọn trả lời, các bước nhảy và logic của lựa chọn trả lời các loại phiếu điều tra; các danh mục và bảng phân loại thống kê sử dụng. Từ đây hình thành lên mô hình chức năng và luồng dữ liệu.

Bước 3: Xác định yêu cầu đối với đối tượng sử dụng ứng dụng

Đối tượng sử dụng ứng dụng thường là các điều tra viên và giám sát viên các cấp. Cần làm rõ yêu cầu về trình độ sử dụng thiết bị số và sử dụng ứng dụng trên thiết bị số đối với đối tượng sử dụng ứng dụng.

Bước 4: Xác định yêu cầu cấu hình của thiết bị cầm tay và hạ tầng công nghệ thông tin

Xác định yêu cầu tối thiểu của thiết bị cầm tay sẽ được cài đặt ứng dụng để thu thập và giám sát thu thập dữ liệu của cuộc điều tra. Đồng thời xác định yêu cầu tối thiểu của hạ tầng công nghệ thông tin phục vụ hoạt động thu thập dữ liệu tại địa bàn (như đường truyền, máy chủ tiếp nhận dữ liệu)

Bước 5: Xác định các chức năng của ứng dụng

Từ nghiên cứu phương án điều tra và đặc điểm của dữ liệu cần thu thập, tiến hành xác định các chức năng của ứng dụng và các tiện ích để đáp ứng yêu cầu của tổ chức thực hiện điều tra (công tác chuẩn bị, điều tra thực địa, xử lý thông tin, chỉ đạo thực hiện).

Bước 6: Xây dựng giao diện chức năng, câu hỏi và câu trả lời

Trên cơ sở các nội dung có liên quan của phương án điều tra, tiến hành thiết kế giao diện cho từng loại phiếu điều tra. Vì kích cỡ màn hình trên thiết bị di động nhỏ nên cần phải sắp xếp câu hỏi và câu trả lời hợp lý và có tính quy luật để điều tra viên dễ sử dụng, thao tác nhanh.

Bước 7: Xây dựng các tiện ích

Ngoài các chức năng thu thập thì cũng cần thiết kế xây dựng các chức năng tiện ích giúp cho điều tra viên có thể thao tác tìm kiếm thông tin một cách nhanh chóng dễ dàng. Các chức năng này như: tìm kiếm thông tin, di chuyển nhanh đến câu hỏi, truy cập vào giải thích câu hỏi....

Bước 8: Xây dựng các báo cáo

## ➤ ➤ ➤ THÔNG KÊ VÀ CUỘC SỐNG

Xây dựng các báo cáo như: báo cáo tiến độ thu thập, báo cáo số lượng đối tượng điều tra, đơn vị điều tra đã được thu thập... Các báo cáo này sẽ giúp cho người đi thu thập có thể thống kê, kiểm soát dữ liệu trong quá trình thu thập; cân đối thời gian, đẩy nhanh tiến độ, chủ động về mặt thời gian cũng như lượng dữ liệu đã thu thập được.

Bước 9: Xây dựng chức năng bảo mật và trao đổi dữ liệu với hệ thống

Chức năng bảo mật là chức năng quan trọng nhất trong thu thập dữ liệu, thông tin thống kê. Để bảo đảm tính bí mật của dữ liệu thì cần phải thiết kế các chức năng bảo mật chặt chẽ. Ngoài bảo mật về mặt nội dung của dữ liệu đã thu thập thì còn phải bảo đảm tính toàn vẹn của dữ liệu đã thu thập trên đường truyền từ thiết bị di động tới hệ thống máy chủ.

Bước 10: Đưa ứng dụng lên kho ứng dụng hoặc hosting

Sau khi hoàn tất việc thiết kế và xây dựng phần mềm thì đến bước cuối cùng là việc đóng gói và phần mềm đến với người sử dụng. Đối với ứng dụng cho hệ điều hành Android thì đưa lên kho Google Play, đối với hệ điều hành iOS thì đưa lên AppStore, webform thì đưa ứng dụng lên hosting. Thực hiện thông báo địa chỉ tải hoặc truy cập ứng dụng, cung cấp tài khoản và mật khẩu truy cập cho người dùng.

### 5. Kết luận

Sử dụng các ứng dụng thu thập dữ liệu trên thiết bị di động và webform là một đòi hỏi khách quan và xu thế tất yếu đang diễn

ra trên thế giới. Đối với nước ta, việc sử dụng hình thức điều tra CAPI và trực tuyến bằng webform là rất cần thiết trong điều kiện cả nước tận dụng các cơ hội do Cách mạng Công nghiệp 4.0 mang lại, đẩy mạnh chuyển đổi số, tăng cường giao tiếp trực tuyến như là một trong các biện pháp cần thiết, hữu hiệu để đạt mục tiêu kép vừa phòng chống đại dịch Covid-19, vừa bảo đảm phát triển kinh tế - xã hội. Điều này ngày càng trở nên cấp thiết trong điều kiện ngành Thống kê đang nỗ lực đổi mới trên mọi mặt để thực hiện tốt sứ mệnh của mình là cung cấp thông tin trung thực, khách quan, chính xác, đầy đủ, kịp thời, minh bạch phục vụ đánh giá, dự báo tình hình, hoạch định chiến lược, xây dựng kế hoạch, quy hoạch chính sách, quản lý, điều hành phát triển kinh tế - xã hội; đáp ứng yêu cầu kiểm tra, giám sát việc thực hiện chiến lược, kế hoạch, chính sách phát triển kinh tế - xã hội; đáp ứng nhu cầu sử dụng thông tin thống kê của cơ quan, tổ chức, cá nhân trong và ngoài nước.

### Tài liệu tham khảo

1. Quốc Hội (2015), *Luật số 89/2015/QH13 ban hành Luật thống kê*, ngày 23 tháng 11 năm 2015;
2. Thủ tướng Chính Phủ (2018), *Quyết định số 501/QĐ-TTg về việc phê duyệt Đề án ứng dụng công nghệ thông tin - truyền thông trong Hệ thống thống kê Nhà nước giai đoạn 2017 - 2025, tầm nhìn đến năm 2030 xác định mục tiêu đến năm 2025*, ngày 10 tháng 5 năm 2018;
3. Báo cáo các cuộc điều tra và Tổng điều tra dân số và nhà ở năm 2019.