

BỘ THÔNG TIN VÀ TRUYỀN THÔNG



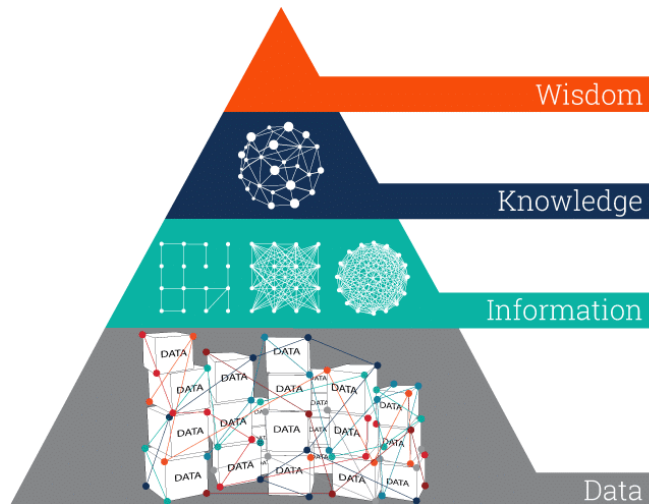
TÀI LIỆU

THUYẾT MINH CHUYÊN ĐỀ “KIẾN THỨC, KỸ NĂNG TẠO LẬP, SỬ DỤNG, KHAI THÁC, SỬ DỤNG DỮ LIỆU SỐ CỦA CÁN BỘ, CÔNG CHỨC”

Hà Nội, năm 2023

1. TỔNG QUAN VỀ DỮ LIỆU

Để hiểu và đi sâu hơn, trước hết cần tìm hiểu một số khái niệm ban đầu về thông tin, dữ liệu và tri thức:



1. Thông tin

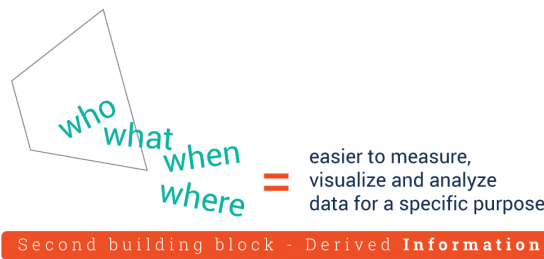
Thông tin là suy nghĩ, ý nghĩ chủ quan của con người truyền đạt hoặc mô tả thông qua nhiều phương tiện khác nhau, như thông tin liên lạc bằng văn bản, bằng miệng, hình ảnh và âm thanh. Đó là kiến thức được chia sẻ hoặc thu được thông qua học tập, hướng dẫn, tìm hiểu hoặc nghe tin tức và chia sẻ nó thông qua hành động giao tiếp, dù bằng lời nói, không lời nói, bằng hình ảnh hoặc thông qua chữ viết.

Thông tin có đặc điểm trừu tượng và vô hình, được nhìn nhận dưới nhiều góc nhìn khác nhau trong nhiều lĩnh vực khác nhau. Theo nghĩa thông thường, thông tin là tất cả các sự việc, sự kiện, ý tưởng, phán đoán làm tăng thêm sự hiểu biết của con người.

Ví dụ:

Khi tiếp nhận một bức ảnh, chúng ta thường sẽ tiếp nhận từ bức ảnh với những câu hỏi và câu trả lời:

- Cái gì trong bức ảnh vậy?
- Ai ở trong bức ảnh và ở đâu?
- Điều quan trọng nhất trong hình ảnh này là gì?
- Ý nghĩa của bức ảnh là gì?
-



Thông tin là thứ làm được tiếp thu qua nhận thức nên có sự không chắc chắn về thế giới qua lăng kính của mỗi người. Đó là câu trả lời cho những câu hỏi như ai, cái gì, ở đâu, bao nhiêu hoặc bao nhiêu. Về cơ bản, thông tin cung cấp sự rõ ràng về thế giới chúng ta đang sống.

Thông tin có thể được tạo ra từ dữ liệu. Chúng ta tạo ra thông tin bằng cách tổ chức, phân tích và giải thích dữ liệu. Việc tổ chức, phân tích và giải thích dữ liệu mang lại cho nó bối cảnh và ý nghĩa. Bối cảnh và ý nghĩa bổ sung này về cơ bản là những gì phân biệt dữ liệu với thông tin.

Ví dụ, hãy tưởng tượng rằng bác sĩ đã ghi lại lịch sử nhiệt độ cơ thể bình thường của con người trong vài năm qua. Họ phân tích dữ liệu lịch sử và tính toán rằng nhiệt độ cơ thể trung bình (nhiệt độ bình thường) của chúng ta là 37°C . Nhiệt độ trung bình 37°C này được gọi là thông tin.

Thông tin có ý nghĩa hơn các điểm dữ liệu riêng lẻ được sử dụng để tạo ra nó. Điều này làm cho thông tin trở nên khá hữu ích.

Khi một người có nhiều thông tin thì dần dần, thông tin đó sẽ biến thành kiến thức của người đó.

Với khả năng thay thế các nguồn tài nguyên khác, khả năng truyền với tốc độ rất cao và khả năng đem lại ưu thế cho người nắm giữ nó, thông tin đã thực sự trở thành cơ sở cho nhiều hoạt động xã hội như nghiên cứu, giáo dục, xuất bản, tiếp thị và cả hoạt động kinh tế, chính trị, văn hoá, xã hội.

Mối quan tâm của xã hội đối với việc sử dụng có hiệu quả các nguồn tài nguyên thông tin đã mở rộng từ các lĩnh vực truyền thống như thư viện, lưu trữ sang các tổ chức, các cơ quan, các ngành. Bây giờ ở bất kỳ đâu con người cũng quan tâm đến việc quản lý và khai thác các nguồn tài nguyên thông tin.

Thông tin có ở khắp mọi nơi. Chúng ta có thông tin trên thực đơn tại nhà hàng, trong sách trong thư viện và trên các biển báo, biển quảng cáo trên đường.

Ứng dụng công nghệ, máy móc để thu thập, lưu trữ, xử lý và truyền đưa thông tin thì người ta gọi là công nghệ thông tin.

Để lưu trữ, xử lý được thông tin, để truyền đưa tin từ người này sang người khác, từ nơi này sang nơi khác thì thông tin phải chuyển thành dữ liệu.

2. Dữ liệu

Dữ liệu (Data) là các số liệu hoặc các tài liệu thu thập được chưa qua xử lý, chưa được biến đổi (dạng thô) cho bất cứ một mục đích nào khác. Dữ liệu có thể biểu diễn dưới nhiều dạng khác nhau bao gồm: âm thanh, văn bản, hình ảnh,...

Ví dụ:

Khi ta muốn thông báo một thông tin cho người khác, ta phải thể hiện bằng văn bản. Khi đó, chữ trên giấy chính là dữ liệu.

Khi một doanh nghiệp bán một lô hàng sẽ xuất hiện các dữ liệu như số lượng hàng hóa bán ra, giá bán, địa điểm bán hàng, thời gian bán hàng, hình thức thanh toán, cách thức giao nhận hàng... Những dữ liệu này sau khi được thống kê sẽ được lưu trữ trên máy tính và quản lý bởi một chương trình máy tính để nhiều người có thể sử dụng với các mục đích khác nhau.

Khi nghĩ về dữ liệu, chúng ta có thể tưởng tượng ra một loạt các số 1 và 0 nằm bên trong máy tính, số liệu thống kê về thu hoạch mùa vụ hay số bệnh nhân tại bệnh viện được chữa trị qua các năm.

Nhưng chính xác thì dữ liệu là gì? Hoặc về mặt kỹ thuật, câu hỏi đúng ngữ pháp hơn là “dữ liệu là gì?”

Dữ liệu là tập hợp các ký hiệu mô tả các quan sát về thế giới xung quanh chúng ta. Chúng ghi lại những sự thật về thế giới tự nhiên mà chúng ta đang sống. Chúng bao gồm những mô tả về đặc tính của sự vật trong thế giới của chúng ta, chẳng hạn như màu sắc, hình dạng và kết cấu. Ngoài ra, chúng còn bao gồm các phép đo số lượng sự vật trong thế giới của chúng ta, chẳng hạn như kích thước, trọng lượng và vận tốc.

Dữ liệu được biểu diễn bằng các ký hiệu. Điều này bao gồm việc thể hiện phẩm chất của sự vật bằng cách sử dụng từ ngữ, ví dụ, màu của lá cây là màu xanh. Ngoài ra, điều này bao gồm việc biểu thị số lượng của sự vật bằng cách sử dụng các con số, ví dụ: quả cam có khối lượng 100 gam.

Hãy tưởng tượng chúng ta cảm thấy ốm và đi đến gặp bác sĩ. Bác sĩ đo nhiệt độ cơ thể bằng nhiệt kế. Nhiệt kế hiển thị 39°C. Con số 39 đó là dữ liệu. Dựa trên nhiệt độ này, rõ ràng là chúng ta đang bị sốt.

Hiện nay hầu hết mọi người sử dụng thuật ngữ “dữ liệu” ở cả dạng số ít và số nhiều, có thể thay thế cho nhau.

Về cơ bản, dữ liệu là thông tin được thể hiện dưới dạng ký hiệu tồn tại bên ngoài con người. Đó là cách chúng ta ghi lại thông tin về thế giới xung quanh.

Tuy nhiên, dữ liệu không chỉ được tạo ra từ thông tin. Dữ liệu còn có thể được sử dụng để tạo ra thông tin mới.

Trong máy tính, để ghi nhận lại dữ liệu, người ta dùng các con số 0 và 1.

Dữ liệu trên máy tính thì chỉ được lưu dưới dạng 0 và 1 tương ứng với trạng thái bật tắt của thiết bị gọi là hệ nhị phân. Chuỗi 8 số 0 hoặc 1 sẽ tạo thành một byte tương đương với một chữ cái hoặc con số 0-10. 1.024 byte tạo thành 1 Kilobyte do máy tính sử dụng hệ nhị phân.

Để thuận tiện cho con người, hiện nay chúng ta làm tròn là 1000 byte thành 1 kilobyte.

thực tế hiện nay, công nghệ phát triển nên khả năng lưu trữ rất lớn, và chúng ta có những đơn vị lớn hơn nhiều.

1 Megabyte = 1 triệu byte (1.048.576 byte)

1 Gigabyte = 1 tỷ byte (1.073.741.824 byte)

1 Terabyte = 1 nghìn tỷ byte (1.099.511.627.776 byte)

1 Petabyte = 1 triệu tỷ byte (1.125.899.906.842.624 byte)

1 Exabyte = 1 tỷ tỷ byte (1.152.921.504.606.846.976 byte)

1 Zettabyte = 1000 tỷ tỷ byte (1.180.591.620.717.411.303.424 byte)

1 Yottabyte = 1 triệu tỷ tỷ byte (1.208.925.819.614.629.174.706.176 byte)

Và hiện nay thì chưa thấy bất cứ thứ gì lớn hơn yottabyte.

Để dễ hình dung:

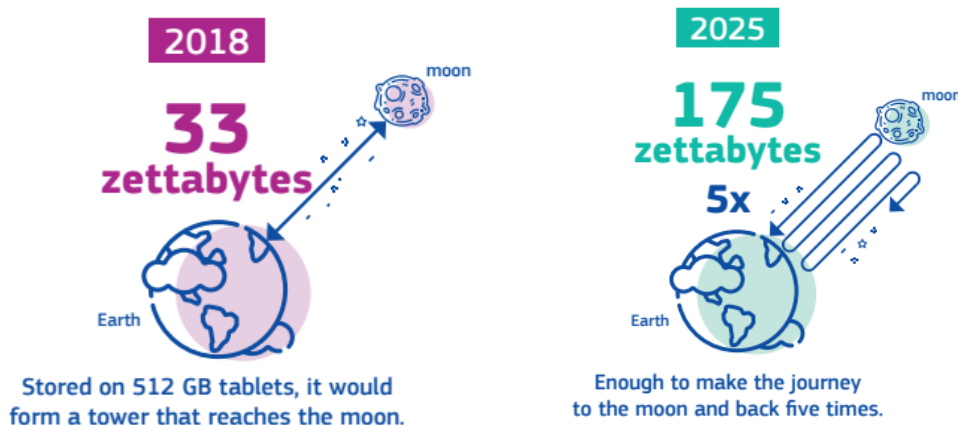
Một quyển sách 800 trang có cả chữ và hình ảnh tương đương 10MB dữ liệu.

Để lưu toàn bộ 1000 đầu sách mà mỗi quyển sách 800 trang thì chỉ cần 10Gigabyte dữ liệu.

Một cái điện thoại bình thường hiện nay chúng ta cầm trên tay có dung lượng 64G thì đã có thể lưu trữ tương đương với 6 thư viện sách như trên.

Dữ liệu hiện nay trên thế giới lớn cỡ nào?

Chúng ta có một hình dung về khối dữ liệu của toàn bộ thế giới hiện nay như sau: Nếu dữ liệu được lưu trong các máy tính bảng có dung lượng 512GB (máy có dung lượng lớn hiện nay- tương đương với 50 thư viện 1000 đầu sách) thì tổng số các máy tính bảng cần thiết để lưu trữ toàn bộ dữ liệu của thế giới sẽ xếp chồng lên nhau sẽ được một tòa nhà cao tới mặt trăng. Đây mới là dữ liệu tính đến năm 2018.



Và dự tính đến năm 2025 thì dữ liệu sẽ lớn đến mức là số máy tính bảng được xếp chồng lên nhau tương ứng với 5 toà nhà cao tới mặt trăng.

Thông tin thì ở trong đầu của mỗi người còn dữ liệu chính là thông tin được ghi lại bên ngoài con người. Nếu chỉ ghi trên sách, giấy thì không ai có thể tìm được vì quá lớn. Chúng ta phải lưu trên máy tính, thiết bị số.

3. Kiến thức

Chúng ta đã biết về thông tin và dữ liệu, chúng ta cũng đã biết là muốn hiểu biết phải học, phải tìm hiểu để có thông tin. Hiểu biết ở đây chính là có kiến thức.

Có thông tin là chúng ta đã biết có ai, cái gì, ở đâu, khi nào. Còn kiến thức chính là chúng ta biết được “làm thế nào”



“Làm thế nào” chính là bước tiếp theo chúng ta sử dụng thông tin được lấy từ dữ liệu được thu thập có liên quan đến mục tiêu của chúng ta?

“Làm thế nào” các phần thông tin này được kết nối với các phần khác để tăng thêm ý nghĩa và giá trị? Và có lẽ quan trọng nhất là chúng ta có thể áp dụng thông tin đó “làm thế nào” để đạt được mục đích của mình?

Ví dụ: qua nghe báo, xem sách, tìm kiếm trên mạng Internet chúng ta đã biết được cây trồng gì năng suất, mua giống ở đâu, trồng vào mùa vụ gì, quy trình chăm bón ra sao thì kiến thức sẽ là cách vận dụng các thông tin này để làm sao chúng ta có thể canh tác được cây trồng đó trên đồng rộng của chúng ta. Tức là chúng ta trả lời cho câu hỏi “Làm thế nào”.

Khi chúng ta không chỉ xem thông tin dưới dạng mô tả các dữ kiện thu thập được mà còn hiểu cách áp dụng nó để đạt được mục tiêu của mình, chúng ta sẽ biến nó thành kiến thức.

Kiến thức này thường là lợi thế mà doanh nghiệp có được so với đối thủ cạnh tranh.

Kiến thức là sự hiểu biết về thế giới tự nhiên xung quanh chúng ta. Nó giải thích những quan sát mà chúng ta nhìn thấy và tại sao mọi thứ lại thế, như thế. Ngoài ra, nó cho phép chúng ta dự đoán hành vi của các hiện tượng trong thế giới của chúng ta. Cả hai điều đó đều khá có lợi cho sự sống còn và khả năng phát triển của chúng ta trên thế giới.

Kiến thức được tạo ra từ thông tin. Về cơ bản, nó là một tập hợp thông tin được tổ chức để mang lại sự hiểu biết nhất quán và gắn kết về một chủ đề cụ thể. Từ quan điểm thực dụng hơn, kiến thức được sử dụng để giải quyết các vấn đề trong một lĩnh vực cụ thể. Kiến thức cho phép chúng ta đưa ra quyết định để có thể thực hiện hành động dẫn chúng ta đến một mục tiêu nào đó.

Ví dụ, hãy tưởng tượng rằng khi bác sĩ khám bệnh. Bác sĩ biết rằng khi nhiệt độ cơ thể của một người tăng lên 39°C thì có khả năng cơ thể đang chống lại virus. Mối quan hệ giữa sự gia tăng nhiệt độ cơ thể và sự hiện diện của virus là những gì chúng ta gọi là kiến thức. Trong trường hợp ví dụ này, bác sĩ sẽ sử dụng kiến thức này để quyết định rằng rất có thể cơ thể của chúng ta đã bị nhiễm virus và cơ thể đang chống lại nó nên mới sinh ra nhiệt làm nóng cơ thể lên. Do hiểu biết này, bác sĩ có thể sẽ khuyên nghị chúng ta nên nghỉ ngơi, uống nhiều nước và có thể tiến hành xét nghiệm thêm nếu cần thiết.

Chỉ kiến thức thôi thì không đủ để giải quyết vấn đề. Để giải quyết các vấn đề trong thế giới thực, chúng ta cần sử dụng kiến thức kết hợp với thông tin mới. Chính sự kết hợp giữa kiến thức hiện có và thông tin mới sẽ dẫn đến giải pháp cho một vấn đề. Ví dụ: bác sĩ sử dụng sự kết hợp giữa kiến thức y tế hiện có của họ và thông tin mới về nhiệt độ cơ thể hiện tại để quyết định rằng đề xuất một phương pháp điều trị.

4. Trí tuệ

Chỉ khi chúng ta sử dụng kiến thức và hiểu biết sâu sắc thu được từ thông tin để đưa ra quyết định chủ động, chúng ta mới có thể nói rằng chúng ta đã đạt đến bước cuối cùng – ‘trí tuệ’.

Trí tuệ là bậc cao nhất và để đạt được điều đó, chúng ta phải trả lời các câu hỏi như 'tại sao phải làm điều gì đó' và 'điều gì là tốt nhất'. Nói cách khác, trí tuệ là kiến thức được áp dụng vào hành động.

Chúng ta cũng có thể nói rằng, nếu dữ liệu và thông tin giống như sự nhìn lại quá khứ thì kiến thức và trí tuệ gắn liền với những gì chúng ta làm hiện tại và những gì chúng ta muốn đạt được trong tương lai.



2. VAI TRÒ CỦA DỮ LIỆU SỐ

Việc thu thập dữ liệu, xử lý dữ liệu để chuyển tải thành thông tin rồi sang tri thức, trí tuệ giờ đây máy tính có thể được hỗ trợ con người rất sâu, thay thế con người rất nhiều bước, công đoạn.

Việc thay thế này chính là ứng dụng công nghệ thông tin. Và giờ đây, chúng ta phát triển đến mức độ cao hơn là chuyển đổi số: xây dựng chính phủ điện tử, chính phủ số, phát triển kinh tế số, xã hội số.

Đối với cơ quan nhà nước, dữ liệu mang lại nhiều lợi ích, có thể tổng kết dữ liệu giúp:

Giúp cơ quan nhà nước, cán bộ hiểu đúng, đủ về đối tượng cần quản lý: quản lý hạ tầng điện đường trường trạm thì biết chính xác được số liệu, hiện trạng. Có biết đúng, biết đủ, kịp thời mới quản lý được.

Nhà nước hiểu rõ về người dân hơn, chủ động phục vụ tốt hơn: Trước đây, khi chưa có dữ liệu, nhà nước chỉ cung cấp những thủ tục hành chính. Công dân nào có nhu cầu thì đến viết đơn đề nghị. Quy trình áp dụng cho mọi người là giống nhau. Khi có đủ dữ liệu, nhà nước sẽ biết mỗi công dân đang cần gì, khi nào cần phải giải quyết thủ tục gì và phục vụ đúng theo yêu cầu của từng người.

Dữ liệu giúp tăng hiệu quả của cơ quan nhà nước: Như ví dụ trên, để tổng hợp số liệu từ địa phương lên trung ương tốn rất nhiều thời gian vì phải qua từng cấp. Nếu chúng ta có sẵn cơ sở dữ liệu tập trung, chỉ cần vài giây là có sẵn số liệu chính xác mà không phải huy động toàn bộ các cơ quan hành chính vào cuộc. Khi có sẵn dữ liệu chính xác và đầy đủ, các năng suất hoạt động của các cơ quan sẽ được thực hiện nhanh chóng và chính xác hơn, giảm thiểu các rủi ro và sai sót trong hoạt động quản lý nhà nước.

Nâng cao năng lực quản lý của các cán bộ công chức, viên chức: Giúp cán bộ, CCVC đỡ vất vả hơn trong các hoạt động chuyên môn của mình. Khi không có dữ liệu, để tìm kiếm thông tin làm báo cáo, cán bộ phải lục trong kho hồ sơ

giấy tờ không lờ mờ tổng hợp thông tin. Khi có dữ liệu trong máy, ví dụ bảng tính chẳng hạn, chỉ cần cập nhật thông tin, dữ liệu tổng hợp tự động điều chỉnh theo.

Cải cách hành chính tối ưu, tinh gọn. Một công việc tổng hợp số liệu trước kia mất rất nhiều thời gian thì khi có dữ liệu máy có thể tự động tổng hợp không tốn nhiều nhân công. Từ đó, cắt giảm được công chức, viên chức xử lý công việc không cần thiết.

Nhà nước minh bạch hơn, giảm tham nhũng: Khi số liệu có sẵn, máy tính tự động xử lý, không ai có thể can thiệp vào công việc. Từ đó, mọi hoạt động phải được minh bạch, giảm ý chí chủ quan của con người. Dữ liệu là công cụ hỗ trợ cho việc đánh giá tác động của quyết định. Nếu có sẵn dữ liệu chính xác, Chính phủ có thể đánh giá tác động của các quyết định và điều chỉnh khi cần thiết. Dữ liệu công bố trên các trang thông tin điện tử sẽ giúp mọi công dân giám sát được hoạt động của cơ quan nhà nước.

Chính sách nhà nước sát với thực tế hơn, giải quyết được các bài toán thực tế. Để xây dựng được hệ thống hành lang pháp lý, lập kế hoạch và triển khai chính sách hiệu quả thì dữ liệu đóng vai trò đầu vào quan trọng. Dữ liệu cung cấp thông tin cho việc lập kế hoạch, triển khai chính sách và hoạt động quản lý công, đảm bảo tính hiệu quả và tính bền vững của các hoạt động này. Sử dụng dữ liệu còn giúp cho các quyết định được đưa ra nhanh chóng và chính xác hơn, từ đó đảm bảo tính minh bạch và trách nhiệm trong hoạt động của cơ quan nhà nước

Tại Việt Nam, Văn phòng Chính phủ đã đưa vào vận hành Trung tâm Thông tin, chỉ đạo điều hành của Chính phủ; hơn 50 bộ, ngành, địa phương đã triển khai Trung tâm giám sát, điều hành thông minh của mình. Các hệ thống này là công cụ đắc lực để hỗ trợ Lãnh đạo Chính phủ, lãnh đạo các cấp trong chỉ đạo điều hành và ra quyết định dựa trên dữ liệu.

Tăng cường khả năng dự đoán và phòng ngừa rủi ro. Quản trị dữ liệu giúp Chính phủ đánh giá và dự đoán các rủi ro trong hoạt động quản trị công, từ đó phòng ngừa các vấn đề tiềm tàng và tăng cường khả năng ứng phó với các tình huống khẩn cấp. Ví dụ, dữ liệu về tình hình kinh tế, tài chính, và dự báo thị trường được sử dụng để xác định chính sách tài chính và ngân sách của Chính phủ. Ngoài ra, dữ liệu về môi trường, khí hậu, và tình trạng thiên tai được sử dụng để phát triển các kế hoạch phòng ngừa và ứng phó với các tình huống khẩn cấp về thiên tai.

Dự báo thời tiết chúng ta nghe hằng ngày là một ví dụ, để có được bản tin dự báo chuẩn xác nhất, nhiều trạm khí tượng, vệ tinh, ra đa phải thu thập dữ liệu. Sau đó hệ thống máy tính mới phân tích và dự báo, dự đoán biến đổi thời tiết để cung cấp thông tin cho chúng ta ứng phó.

Nhận thức được vai trò của dữ liệu phục vụ quản lý nhà nước, những năm qua Chính phủ đã triển khai nhiều giải pháp xây dựng và khai thác dữ liệu phục vụ nâng cao năng lực quản trị công và đạt được nhiều kết quả rõ rệt. Dữ liệu đã từng bước khẳng định được vai trò quan trọng, cốt lõi trong cải cách hành chính và nâng cao năng lực của Chính phủ. Điểm nổi bật có thể kể đến là Cơ sở dữ liệu quốc gia về dân cư và 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 2030 đang được triển khai mạnh mẽ.

Vai trò của dữ liệu phục vụ nâng cao hiệu quả quản trị của cơ quan nhà nước được khẳng định trong nhiều văn bản chương trình, chiến lược. Cụ thể là chương trình chuyển đổi số quốc gia, chiến lược phát triển chính phủ điện tử hướng tới chính phủ số và chiến lược phát triển kinh tế số, xã hội số và nhiều chiến lược chuyên ngành đều lấy trọng tâm xây dựng các cơ sở dữ liệu nhằm tạo nền tảng và động lực để cải cách và đổi mới hoạt động, mang lại hiệu quả cao hơn.

Đối với người dân, doanh nghiệp

Dữ liệu cũng giúp cho người dân đạt được nhiều lợi ích:

Người, dân, doanh nghiệp được phục vụ tốt hơn, đơn giản hoá giấy tờ. Những giấy tờ do cơ quan nhà nước đã cấp sẽ được lưu lại và sử dụng cho những mục đích khác mà người dân không cần nộp lại. Những dữ liệu do cơ quan nhà nước đã lưu trữ thì người dân không cần phải kê khai.

Tăng cường sự tương tác với người dân và huy động sự tham gia của người dân vào hoạt động của cơ quan nhà nước. Dữ liệu cho phép Chính phủ tương tác và giao tiếp với cộng đồng thông qua việc cung cấp thông tin và dữ liệu công khai, từ đó tăng cường sự tham gia và đóng góp của cộng đồng trong hoạt động quản trị công.

4. CÁC QUY ĐỊNH PHÁT TRIỂN DỮ LIỆU

Nhà nước đã sớm nhận ra cần phải phát triển các cơ sở dữ liệu để phục vụ mục đích hỗ trợ hoạt động của cơ quan nhà nước và cải cách hành chính.

Năm 2006, Luật Công nghệ thông tin ra đời đã xác lập trong cơ quan nhà nước cần phải xây dựng các cơ sở dữ liệu quốc gia và cơ sở dữ liệu của bộ, ngành, địa phương.

Theo đó:

Cơ sở dữ liệu quốc gia là tập hợp thông tin của một hoặc một số lĩnh vực kinh tế - xã hội được xây dựng, cập nhật và duy trì đáp ứng yêu cầu truy nhập và sử dụng thông tin của các ngành kinh tế và phục vụ lợi ích công cộng. Tổ chức,

cá nhân có quyền truy nhập và sử dụng thông tin trong cơ sở dữ liệu quốc gia, trừ trường hợp pháp luật có quy định khác.

Cơ sở dữ liệu của bộ, ngành, địa phương là tập hợp thông tin được xây dựng, cập nhật và duy trì đáp ứng yêu cầu truy nhập, sử dụng thông tin của mình và phục vụ lợi ích công cộng. Tổ chức, cá nhân có quyền truy nhập và sử dụng thông tin trong cơ sở dữ liệu của bộ, ngành, địa phương, trừ trường hợp pháp luật có quy định khác.

Năm 2023, Luật Giao dịch điện tử sửa đổi đã quy định cụ thể hơn:

Dữ liệu trong cơ quan nhà nước được tổ chức thống nhất, được phân cấp quản lý theo trách nhiệm quản lý của cơ quan nhà nước nhằm thúc đẩy giao dịch điện tử; được chia sẻ phục vụ hoạt động của cơ quan nhà nước, người dân, doanh nghiệp theo quy định của pháp luật. Cơ sở dữ liệu dùng chung trong cơ quan nhà nước bao gồm cơ sở dữ liệu quốc gia, cơ sở dữ liệu của Bộ, ngành, địa phương.

Cùng với đó là các nghị định của Chính phủ quy định về quản lý, kết nối, chia sẻ dữ liệu của cơ quan nhà nước:

Một số nguyên tắc về pháp lý dữ liệu thực hiện đã được xác lập:

Dữ liệu hình thành trong hoạt động của cơ quan nhà nước được chia sẻ phục vụ các hoạt động của cơ quan nhà nước hướng tới phục vụ người dân, doanh nghiệp tuân thủ quy định của pháp luật trong việc tạo lập, quản lý và sử dụng dữ liệu.

Cơ quan nhà nước phải có trách nhiệm chia sẻ dữ liệu với cơ quan, tổ chức, cá nhân khác theo quy định của pháp luật; không cung cấp thông tin qua hình thức văn bản đối với thông tin đã được khai thác qua hình thức kết nối, chia sẻ dữ liệu giữa các hệ thống thông tin; không được thu thập, tổ chức thu thập lại dữ liệu hoặc yêu cầu người dân, doanh nghiệp cung cấp thông tin, dữ liệu trong quá trình giải quyết các thủ tục hành chính nếu dữ liệu này đã được cơ quan nhà nước khác cung cấp, sẵn sàng cung cấp thông qua kết nối, chia sẻ dữ liệu, trừ trường hợp dữ liệu đó không bảo đảm yêu cầu về chất lượng theo tiêu chuẩn, quy chuẩn chuyên ngành hoặc pháp luật có quy định khác.

Một trong những chính sách rất mới về dữ liệu là dữ liệu mở, Dữ liệu mở của cơ quan nhà nước là dữ liệu được cơ quan nhà nước có thẩm quyền công bố rộng rãi cho cơ quan, tổ chức, cá nhân tự do sử dụng, tái sử dụng, chia sẻ. Cơ quan nhà nước công bố dữ liệu mở để cơ quan, tổ chức, cá nhân tự do sử dụng, tái sử dụng, chia sẻ nhằm thúc đẩy giao dịch điện tử, chuyển đổi số, phát triển kinh tế số và xã hội số.

Cơ quan, tổ chức, cá nhân được phép tự do sao chép, chia sẻ, trao đổi, sử dụng dữ liệu mở hoặc kết hợp dữ liệu mở với dữ liệu khác; sử dụng dữ liệu mở

vào sản phẩm, dịch vụ thương mại hoặc phi thương mại của mình, trừ trường hợp pháp luật có quy định khác.

Dữ liệu mở là bước tiến của việc cung cấp thông tin rộng rãi của CQNN trên cổng thông tin trên Internet, điểm mới là dữ liệu mở được sử dụng để máy móc sử dụng, để người dân, doanh nghiệp sáng tạo các sản phẩm công nghệ số mới.

5. PHÁT TRIỂN DỮ LIỆU TRONG CÁC CƠ QUAN NHÀ NƯỚC

Việc thực hiện phát triển dữ liệu trong cơ quan nhà nước thời gian qua đã đạt được những kết quả lớn: Đã xác lập được các cơ sở dữ liệu quan trọng nhất:

Cơ sở dữ liệu quốc gia về dân cư và căn cước công dân

Bộ Công an đã hoàn thành xây dựng cơ sở dữ liệu quốc gia về dân cư. Kết quả đã thu nhập vào Cơ sở dữ liệu quốc gia về dân cư gần 100 triệu nhân khẩu

Bộ Công an cũng đã phối hợp với Bộ Tư pháp, Bộ Nội vụ rà soát, cập nhật trên trường hợp thi quốc tịch, trường hợp thay đổi thông tin trong hộ tịch để phục vụ việc làm sạch dữ liệu; cấp số định danh cá nhân cho trẻ em mới sinh.

Cơ sở dữ liệu quốc gia về đăng ký doanh nghiệp

Cơ sở dữ liệu quốc gia về đăng ký doanh nghiệp là tập hợp dữ liệu về đăng ký doanh nghiệp trên phạm vi toàn quốc. Thông tin trong hồ sơ đăng ký doanh nghiệp và tình trạng pháp lý của doanh nghiệp lưu giữ tại cơ sở dữ liệu quốc gia có giá trị pháp lý là thông tin gốc về doanh nghiệp. Số lượng doanh nghiệp trong cơ sở dữ liệu khoảng 1,5 triệu doanh nghiệp. Cơ sở dữ liệu quốc gia về đăng ký doanh nghiệp đã kết nối, chia sẻ dữ liệu cho các bộ, ngành, địa phương.

Cơ sở dữ liệu quốc gia về bảo hiểm

Bảo hiểm xã hội Việt Nam thu thập dữ liệu qua các hoạt động nghiệp vụ, từ đó hình thành dữ liệu trong Cơ sở dữ liệu quốc gia về bảo hiểm. Dữ liệu do Bảo hiểm xã hội Việt Nam quản lý bao gồm gần 32 triệu hộ gia đình, hơn 6 triệu người tham gia bảo hiểm xã hội và hơn 80 triệu người tham gia bảo hiểm y tế.

Dữ liệu bảo hiểm cũng đã được sử dụng để xây dựng ứng dụng VssID phục vụ người dân. Đã có hơn 25 triệu tài khoản giao dịch điện tử cá nhân với cơ quan BHXH, người dân đã có thể thực hiện các dịch vụ công với tài khoản bảo hiểm của mình từ điện thoại cầm tay.

Cơ sở dữ liệu quốc gia về đất đai

Cơ sở dữ liệu quốc gia về đất đai bao gồm các dữ liệu thành phần:

- Cơ sở dữ liệu địa chính: đã tích hợp dữ liệu các tỉnh, huyện thị, xã, phường chứa thông tin thửa đất; giấy chứng nhận và hồ sơ tiếp nhận liên quan đến đất đai.

Dữ liệu đất đai cũng được thu thập: Cơ sở dữ liệu Thống kê, kiểm kê đất đai: tập hợp, thống kê các khoanh đất; Cơ sở dữ liệu Giá đất bao gồm dữ liệu khung giá đất thời kỳ 2015-2019, thời kỳ 2020-2025.

- Cơ sở dữ liệu Quy hoạch, kế hoạch sử dụng đất bao gồm cả quy hoạch kế hoạch sử dụng đất trước năm 2010 và năm 2010.

- Cơ sở dữ liệu Điều tra cơ bản đất đai: Số liệu điều tra cấp vùng, cấp quốc gia năm 2015 bao gồm: chất lượng đất; chế độ nước; độ phì nhiêu đất: giảm độ phì; hiệu quả kinh tế: hiệu quả môi trường:

- Cơ sở dữ liệu khác liên quan đất đai: cơ sở dữ liệu người định cư ở nước ngoài sở hữu giấy chứng nhận

Tất cả đang từng bước hoàn thiện và đưa vào cơ sở dữ liệu quốc gia để dùng chung cho toàn quốc.

6. KỸ NĂNG CƠ BẢN VỀ KHAI THÁC, SỬ DỤNG DỮ LIỆU SỐ

Để phát triển kỹ năng về dữ liệu số, các cán bộ, công chức, viên chức luôn ý thức về những nội dung sau:

- **Tư duy về số liệu:** Trong mọi hoạt động để quyết định một vấn đề gì đó, ta luôn cần tư duy về số liệu giúp ta quyết định. Cách đơn giản nhất là tự xác định tỷ lệ phần trăm thành công trong quyết định của chúng ta. Từ đó, ta cần thu thập số liệu để tính tỉ lệ phần trăm đó. Như vậy, mọi quyết định của chúng ta sẽ hạn chế tính chủ quan hơn và sẽ có cơ hội xem xét toàn diện các vấn đề và yếu tố ảnh hưởng.

Thu thập dữ liệu: Thu thập tất cả dữ liệu khi có thể. Dữ liệu nếu không thu thập sẽ mất đi theo thời gian và không thể thu thập lại được. Vì vậy, mọi dữ liệu chúng ta có luôn thu thập và ghi lại như lưu nhật ký, sổ tay của mình. Nhưng chúng ta hình dung tới một cái sổ tay lớn hơn, nhiều người sử dụng hơn và trên máy tính hoặc điện thoại.

Lưu trữ: Hãy cố gắng lưu trữ một cách có hệ thống nhất, luôn tư duy theo bảng có cột và hàng và số liệu được đưa vào cột và hàng đó.

Xử lý: Dữ liệu thô sẽ mang lại rất nhiều điều cho chúng ta, vì vậy, việc tìm kiếm thông tin trong đó phụ thuộc vào nhận thức và kiến thức của chúng ta. Công nghệ số có thể giúp chúng ta tìm rất nhanh, khám phá rất sâu lượng dữ liệu khổng lồ. Vì vậy, nên tìm các công cụ phù hợp. Công cụ đơn giản nhất là bảng tính Excel cần phải được sử dụng thành thạo.

Chia sẻ: Luôn chia sẻ và góp phần vào xây dựng một khối dữ liệu dùng chung, càng nhiều người tham gia dữ liệu càng lớn, kiến thức càng sâu. Chúng ta

có thể chia sẻ theo nhiều cách nhưng đơn giản nhất là sử dụng một nguồn dữ liệu dùng chung và mọi người cùng cập nhật vào đó và sử dụng nó.

Khai thác, sử dụng: Mọi hoạt động đều cần đến dữ liệu, mọi quyết định đều cần dữ liệu làm sở cứ, ta có thể khai thác trên dữ liệu hiện có hoặc có thể tìm nguồn dữ liệu từ nhiều nơi để khai thác./.

BỘ THÔNG TIN VÀ TRUYỀN THÔNG