

CÔNG CỤ AI

Vũ khí mạnh mẽ mới chống lạm dụng số liệu thống kê

Kate Wilkinson, Emily Barrington***

Tóm tắt:

Thông tin sai lệch – dù được phát tán một cách vô tình hay cố ý – đều đe dọa đến các xã hội trên khắp thế giới với những hậu quả nghiêm trọng. Khi số liệu thống kê chính thức bị lạm dụng hoặc phổ biến sai, niềm tin vào sản phẩm công nghệ này và các tổ chức tạo ra chúng sẽ bị hủy hoại. Hai nhóm, tổ chức kiểm chứng thông tin và cơ quan quản lý thống kê, có chung mục tiêu trong việc chống lại vấn đề này. Hỗ trợ công việc của họ là việc cung cấp các công cụ AI mới giúp giám sát khối lượng lớn phương tiện truyền thông, xác định các tuyên bố quan trọng trong cuộc tranh luận công khai và thông báo cho người dùng khi các tuyên bố mà họ biết là sai được lặp lại lần nữa. Bài viết này sẽ đề cập đến vấn đề do thông tin sai lệch gây ra, mục tiêu chung của cơ quan kiểm chứng thông tin và

**Full Fact, Johannesburg, Nam Phi*

** *Cơ quan Quản lý Thống kê, London, Vương quốc Anh*

cơ quan quản lý thống kê, sự phát triển của các công cụ AI và tác động của chúng.

1. Giới thiệu

Thông tin sai lệch, hoặc lan truyền thông tin sai lệch hoặc gây hiểu lầm, được xem là mối đe dọa đối với xã hội trên khắp thế giới. Chu kỳ tin tức 24 giờ và sự phát triển không ngừng của các nền tảng truyền thông xã hội đã khiến việc truyền bá thông tin sai lệch nhanh chóng và đến với nhiều người xem trở nên dễ dàng hơn bao giờ hết. Điều này có thể gây ra những hậu quả nghiêm trọng, từ việc gây hoang mang cho công chúng đến ảnh hưởng tới kết quả của các cuộc bầu cử, gây tổn hại đến kết quả sức khỏe và gây mất lòng tin vào nền dân chủ.

Hai nhóm, tổ chức kiểm chứng thông tin và cơ quan quản lý thống kê, chia sẻ các mục tiêu chung khi chống lại thông tin xấu. Tổ chức kiểm chứng thông tin đánh giá tính chính xác của các tuyên bố của các cá nhân,

➤ ➤ ➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

tổ chức và đảng phái chính trị bằng cách xác minh các nguồn và bằng chứng đằng sau tuyên bố của họ. Cơ quan quản lý thống kê giám sát và đảm bảo chất lượng dữ liệu thống kê được sử dụng bởi các cơ quan chính phủ, doanh nghiệp và công chúng.

Tổ chức kiểm chứng thông tin và cơ quan quản lý thống kê làm việc độc lập với những đối tượng sản xuất và sử dụng số liệu thống kê để xác định các lỗi quan trọng và điểm không chính xác. Cả hai tổ chức đều nhằm mục đích đảm bảo rằng thông tin được phổ biến cho công chúng là chính xác, đáng tin cậy, không thiên vị và thường xuyên hợp tác để đạt được mục tiêu này. Cùng với nhau, họ đóng một vai trò quan trọng trong việc thúc đẩy tính minh bạch và trách nhiệm giải trình, đồng thời bảo vệ niềm tin của công chúng vào số liệu thống kê chính thức.

Tại Vương quốc Anh (UK), việc kiểm chứng thông tin do tổ chức vận động từ thiện Full Fact đã xác minh các tuyên bố trong phạm vi công cộng kể từ năm 2009. Trong những năm gần đây, họ đã đầu tư vào các công cụ kiểm tra tính xác thực tự động được hỗ trợ bởi trí tuệ nhân tạo (AI). Những công cụ này cho phép người dùng xác định các tuyên bố quan trọng để xác minh trên quy mô lớn và cảnh báo người dùng về những hành vi sai trái được lặp đi lặp lại. Các công cụ này phần lớn đã được sử dụng bởi các tổ chức kiểm chứng thông tin; tuy nhiên, rõ ràng là những người dùng khác, chẳng hạn như cơ quan quản lý thống kê, cũng sẽ được hưởng lợi từ những tiến bộ trong không gian.

Cơ quan Quản lý Thống kê (OSR) là cơ quan quản lý độc lập của Cơ quan Thống kê Vương quốc Anh. Cơ quan này cung cấp các quy định độc lập về tất cả các số liệu thống kê chính thức, một tài sản công thiết yếu, được sản xuất tại Vương quốc Anh. OSR nhằm mục đích nâng cao niềm tin của công chúng vào số liệu thống kê do chính phủ tạo ra bằng cách đặt ra các tiêu chuẩn mà họ phải đáp ứng trong Quy tắc thực hành thống kê thông qua các tiêu chí Độ tin cậy, Chất lượng và Giá trị (TQV). Số liệu thống kê phải phục vụ lợi ích của công chúng và OSR thách thức mọi hành vi lạm dụng số liệu thống kê.

Bài viết này sẽ cung cấp một cái nhìn tổng quan về các mối đe dọa do thông tin sai lệch gây ra và phản ứng của các tổ chức kiểm chứng thông tin và cơ quan quản lý thống kê. Sự phát triển của các công cụ kiểm chứng thông tin AI sẽ được thảo luận, cũng như tác động của chúng đối với hai lĩnh vực. Các ví dụ do Full Fact và OSR cung cấp sẽ minh họa tác động của các công cụ trong việc chống lạm dụng số liệu thống kê ở Vương quốc Anh.

2. Bối cảnh

Mỗi ngày ở các quốc gia trên thế giới, mọi người đưa ra quyết định cả lớn và nhỏ. Họ quyết định sẽ sống ở đâu, uống loại thuốc nào và bỏ phiếu cho ai. Ý kiến của họ về tình trạng nền kinh tế và hoạt động của chính phủ được thông báo bằng các tuyên bố công khai, và thường là các số liệu thống kê chính thức.

Thông tin sai lệch có tác động hủy hoại đối với các cá nhân và cộng đồng, kéo dài hận thù và chia rẽ thông qua các tuyên bố lừa đảo hoặc sai sự thật. Nó có thể làm suy yếu các giá trị dân chủ bằng cách làm xói mòn niềm tin vào các quan chức được bầu và các hệ thống chính trị. Nó cản trở việc ra quyết định sáng suốt bằng cách phá vỡ sự hiểu biết của công chúng về các vấn đề quan trọng như biến đổi khí hậu, chăm sóc sức khỏe và chi tiêu công.

Tổ chức kiểm chứng thông tin sẽ sửa chữa những hiểu lầm trong diễn ngôn công khai và quy trách nhiệm cho những người của công chúng về những gì họ nói. Các tổ chức này hoạt động dưới nhiều hình thức khác nhau trên thế giới. Một số là tổ chức từ thiện độc lập hoặc tổ chức phi lợi nhuận đã đăng ký, trong khi những tổ chức khác hoạt động trong các phương tiện truyền thông truyền thống. Mạng lưới kiểm chứng thông tin quốc tế (IFCN), bắt đầu vào năm 2015 tại viện truyền thông phi lợi nhuận và phòng tin tức Poynter, xác minh những tổ chức kiểm chứng thông tin tuân thủ quy tắc, nguyên tắc của họ. Tuy nhiên, nhiều tổ chức kiểm chứng thông tin hoạt động mà không có sự công nhận này.

Các phương pháp kiểm chứng thông tin khác nhau tùy theo tổ chức; tuy nhiên, có một cách tiếp cận chung. Các tổ chức theo dõi cuộc tranh luận công khai ở nhiều nguồn khác nhau bao gồm: phương tiện truyền thông trực tuyến, phương tiện in ấn, phương tiện phát sóng, phương tiện truyền thông xã hội và các ấn phẩm của chính phủ. Họ xác

định các tuyên bố - những tuyên bố có thể được xác minh là đúng hay sai - cần được kiểm tra tính xác thực. Những tổ chức kiểm chứng thông tin thường sẽ liên hệ với người đã đưa ra thông tin để yêu cầu bằng chứng của họ. Bước tiếp theo là xem xét thông tin có sẵn từ một số nguồn đáng tin cậy. Chúng có thể bao gồm các tạp chí học thuật, số liệu thống kê chính thức, trang web của chính phủ, trường đại học và các tổ chức nghiên cứu tư nhân. Những tổ chức kiểm chứng thông tin tham khảo ý kiến của các chuyên gia trong lĩnh vực này để giải mã luận điệu và thông tin họ đã thu thập được. Các chuyên gia thường có thể cung cấp bối cảnh, phân tích và cái nhìn sâu sắc. Các phát hiện được viết ra, xem xét và chỉnh sửa trước khi xuất bản.

Những tổ chức kiểm chứng thông tin được ví như "những người phản hồi đầu tiên" trong thời kỳ khủng hoảng thông tin, vì họ thường ở tuyến đầu để xác định các tuyên bố sai hoặc hiểu lầm. Cơ quan quản lý thống kê quốc gia đóng một vai trò tương tự.

Ở Vương quốc Anh, OSR có cách tiếp cận hai hướng để xây dựng niềm tin của công chúng vào số liệu thống kê. Đầu tiên, họ tập trung vào cách số liệu thống kê được tạo ra, xem xét độc lập liệu số liệu thống kê chính thức có tuân thủ Quy tắc thực hành thống kê và cho thấy độ tin cậy, chất lượng và giá trị hay không. Các cơ quan quản lý thống kê của tổ chức sử dụng quy trình đánh giá phù hợp để đánh giá mức độ mà số liệu thống kê đáp ứng các tiêu chuẩn được quy định trong Quy tắc. Nếu, theo

➤ ➤ ➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

đánh giá độc lập này, số liệu thống kê chính thức tuân thủ Bộ luật, thì chúng đáp ứng yêu cầu pháp lý để được gắn huy hiệu là Số liệu thống kê quốc gia.

Thứ hai, OSR cũng đóng một vai trò phản hồi quan trọng trong việc thách thức khi lạm dụng số liệu thống kê. Tổ chức tích cực theo dõi và tìm kiếm các trường hợp sử dụng sai số liệu thống kê trên các phương tiện truyền thông, mạng xã hội và ở những nơi có ảnh hưởng lớn đến công chúng, chẳng hạn như các cuộc tranh luận tại quốc hội. Các trường hợp lạm dụng tiềm ẩn cũng có thể được chuyển đến tổ chức bởi các cá nhân muốn khiếu nại. Khi phát hiện ra có trường hợp lạm dụng, tổ chức sẽ xem xét phản hồi thích hợp để đảm bảo tính chính xác và toàn vẹn của số liệu thống kê trong phạm vi đại chúng.

Hai nhóm này - kiểm chứng thông tin và quản lý thống kê - chia sẻ một số mục tiêu chung. Cả hai đều theo dõi phương tiện truyền thông để phát hiện các lỗi thống kê và xác minh các khiếu nại được chia sẻ trong phạm vi đại chúng. Họ cũng làm việc để xác minh thông tin và yêu cầu các nhân vật của công chúng chịu trách nhiệm về thông tin họ chia sẻ. Họ cùng nhau tuyên truyền và giáo dục công chúng về việc lạm dụng số liệu thống kê và khuyến khích tính minh bạch trong dữ liệu thống kê.

Các công cụ do AI cung cấp có khả năng hỗ trợ các mục tiêu chung của tổ chức kiểm chứng thông tin và cơ quan quản lý thống kê. Các công cụ này có thể phân tích nhanh chóng và chính xác khối lượng lớn dữ liệu và

xác định các lỗi tiềm ẩn, sự không nhất quán hoặc thông tin sai lệch.

3. Công cụ AI

Hỗ trợ và tài trợ cho kiểm chứng thông tin tự động đã đạt được động lực vào giữa những năm 2010. Các tổ chức kiểm chứng thông tin như Full Fact và Duke Reporters Lab đã nhận được tài trợ cho các dự án từ một số nhà tài trợ, bao gồm Google, Quỹ Omidyar, Tổ chức Xã hội Mở, Tổ chức Hiệp sĩ, Dự án Báo chí Facebook và Quỹ Craig Newmark.

Hỗ trợ tài chính cho phép họ và những nhà đầu tư vào nghiên cứu và phát triển các công cụ kiểm chứng thông tin tự động. Có hai giai đoạn của quy trình xác minh tính xác thực mà các công cụ xác minh tính xác thực tự động đã cho thấy tiềm năng đặc biệt. Tác động của các công cụ đã được chứng minh là tiết kiệm thời gian và công sức của người dùng và cho phép họ làm việc trên quy mô lớn. Đây là một trong những phát hiện xác nhận quyền sở hữu và so khớp xác nhận quyền sở hữu.

3.1. Tìm các khiếu nại

Phát hiện khiếu nại – hoặc tìm các tuyên bố để điều tra và xác minh – là bước đầu tiên trong quy trình xác minh tính xác thực. Có hai thành phần cho giai đoạn này. Đầu tiên, tuyên bố phải có thể kiểm chứng được, tức là nó phải được diễn đạt theo cách cho phép xác nhận tính chính xác của nó.

Một tuyên bố như “Vương quốc Anh là quốc gia tốt nhất trên thế giới” không thể được xác minh vì nó dựa trên ý kiến cá nhân.

Cách hiểu và định nghĩa về “tốt nhất” sẽ khác nhau tùy theo từng người. Khi so sánh, tuyên bố “nền kinh tế Vương quốc Anh ghi nhận tốc độ tăng trưởng nhanh nhất trong G7” có thể được kiểm chứng. Có những phương pháp được chấp nhận trên toàn cầu để đo lường và so sánh quy mô của các nền kinh tế. Có nhiều cơ sở dữ liệu để tham khảo và các chuyên gia có thể tư vấn về cách tốt nhất để so sánh các biện pháp.

Thứ hai, tuyên bố phải “đáng để xác thực”. Đây là những tuyên bố mà “công chúng sẽ quan tâm đến việc tìm hiểu về tính xác thực của chúng”. Ví dụ, Gou et al. mô tả tuyên bố “hơn sáu triệu người Mỹ đã mắc COVID-19 vào tháng 1” là một tuyên bố đáng để xác thực. Trong khi đó, câu nói “nước ướt” thì không. Sẽ không có lợi ích rộng rãi trong việc xác định tính xác thực của một tuyên bố như vậy.

Các cách tiếp cận truyền thống để kiểm tra tính xác thực của xác nhận quyền sở hữu bao gồm xem xét thủ công các phương tiện truyền thông tin tức, mạng xã hội, thông cáo báo chí và các ấn phẩm của chính phủ. Một số tổ chức cũng khuyến khích các thành viên của công chúng gửi các tuyên bố mà họ tin rằng cần phải được xác minh. Các cách tiếp cận “công nghệ thấp” bao gồm đặt cảnh báo từ khóa, sử dụng các công cụ trực tuyến miễn phí như Google Alerts. Các giải pháp trả lại kết quả khác nhau. Họ sử dụng nhiều nhân lực vì họ yêu cầu các thành viên trong nhóm được chỉ định thực hiện giám sát thủ công. Chúng cũng có thể bị kém về mặt công nghệ, vì các tìm kiếm từ khóa đơn giản sẽ

không trả về các yêu cầu có cách diễn đạt khác.

Các công cụ kiểm tra tính xác thực tự động hỗ trợ người dùng trong giai đoạn đầu tiên của quy trình xác minh tính xác thực: tìm các câu khai mở để thẩm định và xác minh. Vấn đề xếp hạng hoặc ghi nhãn các tuyên bố theo mức độ xác thực của chúng hiện đang được phát triển tại Full Fact.

Các công cụ AI của Full Fact giám sát nhiều nguồn phương tiện bao gồm các trang web tin tức, phương tiện truyền thông xã hội, chương trình phát sóng trực tiếp và các ấn phẩm của chính phủ. Các nguồn này được thu thập hàng ngày và chia thành các câu riêng lẻ, là đơn vị phân tích. Sau đó, những câu này sẽ trải qua một số giai đoạn trong đó chúng được gắn nhãn và bổ sung để làm cho chúng hữu ích hơn cho người dùng.

Người dùng sau đó được hiển thị các câu yêu cầu. Đây là những câu có chứa một khẳng định có thể được xác minh hoặc kiểm chứng thông tin. Những tuyên bố này được gắn nhãn thêm để hỗ trợ người dùng theo dõi các cuộc tranh luận hoặc chủ đề mà họ quan tâm bằng cách sử dụng lược đồ được phát triển bởi những tổ chức kiểm chứng thông tin và các học giả. Tám nhãn loại khiếu nại được sử dụng trong các công cụ phát hiện khiếu nại của Full Fact AI bao gồm:

- Số lượng: Tuyên bố về giá trị hiện tại, thay đổi số lượng, so sánh số lượng và xếp hạng số lượng. Ví dụ: “Có nhiều người đi làm hơn trước khi xảy ra đại dịch”.

➤ ➤ ➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

- Tương quan/nhân quả: Khẳng định về mối tương quan, khẳng định về quan hệ nhân quả và khẳng định về việc không có liên kết. Ví dụ: “Uống thuốc tránh thai có thể “cứu sống bạn” vì ít phụ nữ có ý định tự tử hơn”.

- Dự đoán: Các tuyên bố và tuyên bố mang tính giả thuyết về tương lai. Những tuyên bố này thường không thể kiểm chứng được.

- Trải nghiệm cá nhân : Tuyên bố về trải nghiệm cá nhân, thường không thể kiểm chứng bằng thông tin có sẵn công khai. Ví dụ: “Tôi không đủ khả năng để tiết kiệm cho việc nghỉ hưu của mình.”

- Quy tắc: Yêu cầu về các thủ tục thể chế công và thay đổi quy tắc. Ví dụ: “Bạn không cần phải tuân theo các quy tắc của chính phủ sau khi tiêm phòng”.

- Hồ sơ biểu quyết: Tuyên bố về cách một người, nhóm hoặc đảng chính trị đã bỏ phiếu về một chủ đề. Ví dụ: “Chín trong 10 thành viên của NASUWT đã bỏ phiếu ủng hộ hành động đình công”.

- Hỗ trợ: Tuyên bố quan điểm về các chủ đề nhất định. Ví dụ: “Scotland không muốn trở thành thành viên của EU”.

- Ý kiến cá nhân: Tuyên bố về dư luận hoặc khảo sát về một vấn đề. Ví dụ: “Cứ mười thanh niên thì có một người dự định không bao giờ kiếm được việc làm”.

Người dùng cũng có thể lọc các câu bằng cách xác định từ khóa, phạm vi ngày, chủ đề và ấn phẩm. Các công cụ này cũng cho phép người dùng tìm kiếm các tuyên bố của các thực thể, chẳng hạn như người dân, tổ chức

và đảng phái chính trị. Hệ thống ghi nhãn và lọc này cho phép người dùng giảm hàng trăm nghìn khiếu nại thành số lượng có thể quản lý được để xem xét.

Các công cụ đã mở rộng đáng kể số lượng nguồn phương tiện mà tổ chức kiểm chứng thông tin có thể theo dõi. Tại Full Fact, khả năng phát hiện khiếu nại đạt khoảng 100 khiếu nại mỗi ngày. Sau khi giới thiệu công cụ này, tổ chức hiện có thể xem xét 100.000 xác nhận mỗi ngày. Trong mỗi quan hệ hợp tác gần đây với những tổ chức kiểm chứng thông tin ở Nigeria trước cuộc bầu cử tổng thống năm 2023 của đất nước, các công cụ AI của Full Fact có thể cung cấp trung bình 40.000 xác nhận để xem xét mỗi ngày từ 80 nguồn truyền thông.

3.2. So khớp tính xác thực

Lĩnh vực phát triển thứ hai trong các công cụ kiểm tra tính xác thực của AI là “so khớp tính xác thực”. Tổ chức kiểm chứng thông tin và những người dùng khác cần được thông báo khi điều gì đó mà họ biết là sai được lặp lại lần nữa. Điều này cho phép họ xem xét sự giả dối lặp đi lặp lại và quyết định xem họ nên làm gì. Những hành động này có thể bao gồm xuất bản một bản xác minh tính xác thực mới, liên hệ với người đã lặp lại hành vi sai trái hoặc nộp đơn khiếu nại với cơ quan quản lý.

Các cách tiếp cận truyền thống để so khớp tính xác thực bao gồm theo dõi và xem xét phương tiện theo cách thủ công, điều này rất tốn thời gian. Các công cụ như Google Alerts có thể tìm kiếm hàng ngày cho một số từ khóa nhất định. Tuy nhiên, điều này sẽ bỏ

lỡ một số lượng đáng kể các xác nhận quyền sở hữu phù hợp vì ngôn ngữ là động và có nhiều cách khác nhau để đưa ra cùng một xác nhận quyền sở hữu. Cuối cùng, các ý kiến đóng góp từ công chúng có thể hữu ích nhưng chúng không đáng tin cậy hoặc nhất quán. Các công cụ so khớp tính xác thực do AI cung cấp giúp giảm thời gian, công sức và hạn chế của các phương pháp này.

Các xác nhận tính chính xác bị loại bỏ và gắn nhãn trong bước phát hiện xác nhận quyền sở hữu trước đó được so sánh với bộ dữ liệu gồm các xác nhận đã được kiểm tra hoặc giám sát tính xác thực trước đó. Công cụ cảnh báo của AI Full Fact sử dụng mô hình kiểu BERT để dự đoán câu khớp/không khớp cho các câu. Điều này được kết hợp với tìm kiếm từ khóa, ước tính độ tương tự ngữ nghĩa và so khớp thực thể. Các xác thực đáp ứng ngưỡng so khớp tối thiểu sẽ được hiển thị cho người dùng để xem xét. Người dùng quyết định cách hành động đối với những cảnh báo này.

Ví dụ: đã có sự nhầm lẫn của công chúng về những thay đổi đối với Bộ luật Đường cao tốc của Vương quốc Anh khi nó được cập nhật vào năm 2022. Những kiểm chứng thông tin của Full Fact đã viết về một số thay đổi, bao gồm cả báo cáo của một tờ báo rằng "người đi xe đạp ở hai bên xe được ưu tiên khi ô tô đang rẽ". Một tờ báo khác đưa tin về cùng mục đã viết: "Người đi xe đạp có thể vượt bạn ở bên trái cũng như bên phải khi bạn bị tắc đường: Người lái xe cần phải tỉnh táo trên các tuyến đường tắc nghẽn, vì bản cập nhật của Bộ luật Đường cao tốc hiện cho biết người đi xe đạp được phép vượt họ khi

giao thông di chuyển chậm hoặc dừng ở cả bên phải và bên trái".

Cả hai câu đều đề cập đến "người đi xe đạp" nhưng câu trước đề cập đến "phương tiện" và "ô tô", trong khi câu sau đề cập đến "người lái xe mô tô". Cái trước dùng để chỉ "một trong hai bên của phương tiện" và cái sau dùng "cả bên phải và bên trái". Những khác biệt này chứng minh cách tìm kiếm đơn giản dựa trên các từ cụ thể sẽ không bao giờ tìm thấy kết quả khớp này. Các công cụ AI của Full Fact đã nhận ra hai câu này khớp nhau dù được diễn đạt bằng các từ khác nhau.

Công cụ này cho phép Full Fact hành động dựa trên nhiều lần lặp lại tuyên bố gây hiểu lầm do Thủ tướng Vương quốc Anh lúc bấy giờ là Boris Johnson đưa ra vào ngày 24 tháng 11 năm 2021. Trong các câu hỏi của Thủ tướng, ông tuyên bố rằng hiện có nhiều người làm việc hơn so với trước khi đại dịch Covid-19 bắt đầu.

Theo nội dung trên, Johnson đang đề cập đến số lượng công nhân Vương quốc Anh trong biên chế của chủ lao động là 29,3 triệu vào tháng 10 năm 2021, cao hơn mức 29 triệu vào tháng 2 năm 2020, trước khi bắt đầu đại dịch. Nhưng điều này không bao gồm tất cả người đang làm việc, vì nó loại trừ những người tự kinh doanh.

Dữ liệu tư vấn từ Cơ quan Thống kê Quốc gia, bao gồm tất cả những người làm công việc được trả lương ở Vương quốc Anh, cho thấy con số này vẫn thấp hơn mức ngay trước đại dịch. Nó đã giảm từ khoảng 33 triệu trong khoảng thời gian từ tháng 12 năm 2019

➤ ➤ ➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

đến tháng 2 năm 2020 xuống còn 32,5 triệu theo số liệu mới nhất có sẵn vào thời điểm đó (tháng 7 đến tháng 9 năm 2021).

Sau khi công bố kiểm chứng thông tin, Full Fact đã liên hệ với văn phòng Thủ tướng để yêu cầu đính chính về tuyên bố này. Tổ chức này cũng đã viết thư cho OSR vào tháng 1 năm 2022 để nêu rõ việc Thủ tướng sử dụng sai dữ liệu. Tại thời điểm gửi thư, các công cụ so khớp tính xác thực do AI cung cấp đã cảnh báo Full Fact về ba trường hợp nữa Johnson lặp lại yêu cầu không chính xác vào ngày 1 tháng 12 năm 2021, ngày 15 tháng 12 năm 2021 và ngày 5 tháng 1 năm 2022.

Công cụ AI cũng xác định những người khác đã đưa ra các tuyên bố. Thành viên đảng Bảo thủ của quốc hội Lee Rowley cho biết tại Quốc hội vào ngày 11 tháng 1 năm 2022: “Có hơn 400.000 người có việc làm so với trước đại dịch”. Một thành viên Đảng Bảo thủ khác của quốc hội Paul Scully cũng chia sẻ thống kê thiếu sót tương tự. Nó cũng được chia sẻ trên Twitter bởi Mark Spencer, Suella Braverman và Nadine Dorries. Công cụ so khớp tính xác thực của AI có thể giám sát phương tiện truyền thông trực tuyến và phương tiện truyền thông xã hội hàng ngày để biết các lần lặp lại xác thực và cảnh báo cho nhóm Full Fact. Điều này đã biến những gì có thể là cả ngày giám sát thành một đánh giá ngắn gọn về các so khớp tính xác thực có thể xảy ra.

Ed Humpherson, Tổng Giám đốc Quy định tại Cơ quan quản lý Thống kê, đã trên Downing Street vào ngày 01 tháng 02 năm 2022 để nói rằng, không đúng khi tuyên bố rằng có nhiều người làm việc vào cuối giai

đoạn này hơn so với đầu kỳ. Công việc của Full Fact là một trong nhiều yếu tố mà nhóm tại OSR đã cân nhắc khi xem xét phản hồi này.

Kể từ ngày 07 tháng 3 năm 2023, công cụ AI của Full Fact đã xác định được 27 lần lặp lại tuyên bố này mà có thể đã không bị phát hiện. Điều này đã cho phép tổ chức làm được hai việc. Đầu tiên, nó có thể theo dõi sự lan truyền của các thông tin theo thời gian. Một phần của việc chống lại khi sử dụng sai số liệu thống kê là hiểu rõ ai sử dụng sai chúng, khi nào, ở đâu và như thế nào. Khả năng lập bản đồ mức độ lan rộng và tác động của việc sử dụng sai số liệu thống kê là một cách mạnh mẽ để thu hút sự chú ý đến hậu quả của nó. Thứ hai, nó cho phép có nhiều cơ hội hơn để thu hút các nhân vật của công chúng về việc sử dụng (hoặc lạm dụng) số liệu thống kê của họ và yêu cầu chỉnh sửa hoặc cam kết khi thích hợp. Full Fact đã có thể thực hiện chín hành động can thiệp riêng biệt đối với việc sử dụng sai số liệu thống kê chính thức này.

4. Việc sử dụng các công cụ AI của Cơ quan quản lý thống kê chính thức

OSR cũng thực hiện giám sát thường xuyên các nền tảng công khai, bao gồm cả phương tiện truyền thông và phương tiện truyền thông xã hội, để xác định khả năng lạm dụng số liệu thống kê. Việc giám sát này cũng cho phép tổ chức tiến hành “quét theo chiều ngang”, nơi họ phát hiện và hiểu các xu hướng mới nổi trong số liệu thống kê và các câu chuyện xung quanh chúng. Điều này cung cấp cái nhìn sâu sắc về cách tốt nhất để hỗ trợ các nhà sản xuất số liệu thống kê

trong việc truyền đạt số liệu thống kê có tác động cao và hỗ trợ sự hiểu biết của công chúng về câu chuyện mà số liệu thống kê kể. Thực hiện cả hai hoạt động này theo cách thủ công đi kèm với nhiều thách thức mà các công cụ hỗ trợ bởi AI có thể giúp vượt qua.

Thứ nhất, có rất nhiều thông tin để theo dõi. Lấy phương tiện truyền thông làm ví dụ, có rất nhiều nền tảng truyền thông chính thống. Nhưng cũng có nhiều cơ quan độc lập, nhỏ hơn sẽ chỉ báo cáo về số liệu thống kê liên quan đến một lĩnh vực nhất định, chẳng hạn như số liệu thống kê về giáo dục. Kiểm tra từng số liệu này yêu cầu một quy trình nhất quán và tài liệu tốt về các cụm từ tìm kiếm để sử dụng cho từng cụm từ. Mức độ nhất quán này khó đạt được theo cách thủ công vì trí nhớ của con người là một yếu tố quan trọng. Tuy nhiên, khi sử dụng các công cụ AI, các nền tảng và thuật ngữ tìm kiếm được sử dụng có thể được mã hóa cứng và do đó không có tìm kiếm nào bị lãng quên.

Thứ hai, thông tin được theo dõi thay đổi nhanh chóng. Mỗi nền tảng cần được kiểm chứng ít nhất hai lần một tuần để đảm bảo phản hồi kịp thời và không bỏ sót điều gì. Đối với mạng xã hội, điều này cần phải diễn ra thường xuyên hơn nữa vì tốc độ trao đổi thông tin trên các nền tảng này cao hơn nhiều. Một hoạt động thường xuyên và kỹ lưỡng như vậy chiếm rất nhiều thời gian, khiến bạn có ít thời gian hơn để tập trung vào các ưu tiên khác. Các công cụ AI có thể chạy các kiểm chứng này theo định kỳ, tự động giải phóng thời gian của cơ quan quản lý để dành cho các nhiệm vụ khác.

Cuối cùng, việc kiểm chứng thông tin sử dụng các từ và cụm từ cụ thể rất mệt mỏi và dễ xảy ra lỗi do con người. Khi kiểm chứng từng nền tảng, các cơ quan quản lý OSR đang tìm kiếm các tài liệu tham khảo về số liệu thống kê chính thức có thể được trích dẫn trực tiếp hoặc, trong trường hợp có khả năng lạm dụng, các tài liệu tham khảo thống kê không liên quan đến các nguồn chính thức. Các tài liệu tham khảo này có thể ở một trong nhiều dạng và điều đó đòi hỏi các cơ quan quản lý phải có con mắt tinh tường và mức độ tập trung cao trên tất cả các nền tảng khi tìm kiếm. Các công cụ AI loại bỏ yếu tố mệt mỏi này của con người. Các công cụ AI có thể kiểm tra nhiều dạng cụm từ khác nhau và làm như vậy mà không mệt mỏi. Điều này tạo ra mức độ hài lòng trong công việc cao hơn cho các nhà quản lý bằng cách tránh các quy trình lặp đi lặp lại.

Các công cụ AI cũng đã tăng cường các lĩnh vực khác trong công việc của OSR. Chúng cho phép tổ chức tự động kiểm tra khi việc sử dụng sai số liệu thống kê được lặp lại bởi cùng một người/nhóm ban đầu hoặc bởi những người khác. Một số trường hợp lạm dụng cần có sự can thiệp của công chúng, trong khi những trường hợp khác yêu cầu cách tiếp cận cá nhân hơn. Bất kể hành động được thực hiện là gì, các công cụ này cho phép tổ chức đánh giá mức độ lạm dụng một cách nhanh chóng và dễ dàng.

Các công cụ cho phép tổ chức lọc các câu và câu dựa trên việc chúng có chứa thông tin thống kê hay không. Điều này được thực hiện bằng cách sử dụng loại yêu cầu "số lượng" được mô tả trước đó. Các công cụ AI thu

➤ ➤ ➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

thập nội dung trên các nền tảng có sẵn công khai và sử dụng các mô hình Xử lý ngôn ngữ tự nhiên (NLP) được tập huấn trước để phát hiện câu nào chứa ngôn ngữ thống kê như "phần trăm", "tăng" hoặc thông tin số. Bằng cách này, OSR đã có thể thu thập thông tin từ Hansard (địa điểm chính thức của tất cả các cuộc tranh luận tại Nghị viện Vương quốc Anh) và chỉ xem xét các câu có chứa thông tin thống kê.

Cuối cùng, các công cụ AI giúp trả lời các câu hỏi quan trọng liên quan đến tầm nhìn của tổ chức. Ví dụ, là một phần của chương trình nghiên cứu OSR, tổ chức muốn hiểu thêm về hàng hóa công và cách các bên trung gian, chẳng hạn như phương tiện truyền thông, mô tả hàng hóa công. Điều này giúp bổ sung cho các công việc khác trong không gian này, chẳng hạn như công việc tương tác và đánh giá tài liệu. Trong ví dụ này, OSR đã sử dụng các công cụ AI để tìm kiếm các tuyên bố của các cá nhân có chứa từ "hàng hóa công" và "số liệu thống kê" hoặc "dữ liệu" từ tháng 01 năm 2020 đến tháng 10 năm 2022. Kết quả của việc này đã được đưa vào hội thảo của chúng tôi để giúp xác định hàng hóa công của số liệu thống kê và nhìn chung đã cung cấp một chế độ xem xác định, không chỉ dựa vào một khoảng thời gian.

5. Thách thức và hạn chế

Các công cụ do AI cung cấp đã hứa hẹn và có tác động đối với những tổ chức kiểm chứng thông tin và cơ quan quản lý giám sát việc lạm dụng số liệu thống kê chính thức. Tuy nhiên, có những hạn chế để xem xét.

Có những lợi ích cho các cơ quan quản lý thống kê để thực hiện giám sát phương tiện truyền thông của họ theo cách thủ công. Việc yêu cầu các cơ quan quản lý đọc và giám sát thông tin theo cách thủ công dẫn đến nhận thức và hiểu biết chung về một tình huống có thể có lợi cho việc xây dựng kiến thức đúng chuẩn.

Các nhà quản lý con người cũng có nhiều khả năng xác định thông tin được diễn đạt hơi khác so với những gì một công cụ AI tìm thấy.

Các công cụ, khi được sử dụng bởi tổ chức kiểm chứng thông tin hoặc cơ quan quản lý thống kê, vẫn cần có sự giám sát của con người. Mặc dù chúng có thể xác định tính chính xác và so khớp phù hợp với tiềm năng, nhưng chúng không thể hiểu ngữ cảnh hoặc sắc thái. Một con người vẫn cần thiết để đảm bảo rằng kết quả là chính xác nhất và có liên quan.

Tuy nhiên, OSR đã phát hiện ra rằng những lợi ích chung của các công cụ AI đã vượt xa giới hạn này. Tổ chức sử dụng chúng, cùng với chuyên môn của cơ quan quản lý, để mở rộng và nâng cao khả năng giám sát việc sử dụng số liệu thống kê. Tương tự, Full Fact AI không coi các công cụ AI là sự thay thế cho công cụ kiểm chứng thông tin và công việc mà chúng làm. Thay vào đó, các công cụ bổ sung kỹ năng và giải phóng thời gian của họ để thực hiện các nhiệm vụ quan trọng khác không thể tự động hóa được.

6. Kết luận

Thông tin sai lệch đặt ra các mối đe dọa cho các xã hội trên khắp thế giới. Thông tin xấu có thể được tạo ra để đánh lạc hướng, không có bất kỳ cơ sở thực tế hoặc bằng chứng nào. Trong các trường hợp khác, số liệu thống kê chính thức bị trình bày sai hoặc sử dụng sai. Hệ lụy của việc này là nghiêm trọng.

Thứ nhất, công chúng bị nhầm lẫn. Điều này có thể ảnh hưởng đến nhận thức và hiểu biết của công chúng về các vấn đề quan trọng. Thứ hai là sự mất niềm tin vào số liệu thống kê chính thức và các tổ chức tạo ra chúng.

Một số bên liên quan đang làm việc để chống lại việc sử dụng sai số liệu thống kê chính thức và buộc các nhân vật công khai phải chịu trách nhiệm về việc gây hiểu lầm cho công chúng. Chúng bao gồm các tổ chức kiểm chứng thông tin độc lập và cơ quan quản lý thống kê. Họ chia sẻ một số mục tiêu chung, bao gồm: phát hiện các lỗi thống kê, xác minh các khiếu nại được chia sẻ trong phạm vi công chúng, yêu cầu các nhân vật của công chúng chịu trách nhiệm về thông tin họ chia sẻ, giáo dục công chúng về việc sử dụng sai số liệu thống kê và khuyến khích tính minh bạch trong dữ liệu thống kê. Những mục tiêu được chia sẻ này thường dẫn đến những nỗ lực phối hợp với tư cách là quan hệ đối tác, như trường hợp của Full Fact và OSR.

Quy mô của vấn đề mà họ đang giải quyết là rất lớn. Thông tin sai lệch có thể được tạo ra và phổ biến với tốc độ chưa từng

thấy trước đây. Các phương pháp theo dõi và phát hiện thủ công truyền thống không phù hợp với khối lượng thông tin được chia sẻ mỗi ngày.

Một giải pháp cho vấn đề này là trao quyền cho những tổ chức kiểm chứng thông tin và cơ quan quản lý thống kê bằng công nghệ tiên tiến, để cho phép họ nâng cao nỗ lực của mình. Các công cụ AI, ban đầu được thiết kế cho những tổ chức kiểm tra tính xác thực, đã cho thấy giá trị của chúng với các cơ quan quản lý. Việc cung cấp các công cụ hiện tại cho phép cả hai nhóm theo dõi khối lượng lớn các cuộc tranh luận công khai, tìm các tuyên bố quan trọng và được thông báo khi điều gì đó mà họ biết là sai được lặp lại.

Các giải pháp công nghệ này không nhằm mục đích thay thế các công cụ kiểm chứng thông tin hoặc cơ quan quản lý thống kê. Các công cụ AI đã được phát triển để hỗ trợ những tổ chức kiểm chứng thông tin và cơ quan quản lý thống kê trong nỗ lực của họ, cho phép họ tập trung vào công việc phân tích phức tạp hơn trong khi các công cụ này đảm nhiệm các công việc thường ngày. Những công cụ này có sẵn cho người dùng trên toàn thế giới và có khả năng mở rộng đáng kể tác động của công việc của các nhóm này, giúp chống lại việc lạm dụng số liệu thống kê chính thức và quảng bá thông tin đáng tin cậy.

Đỗ Ngát (lược dịch)

Nguồn: <https://content.iospress.com/download/statistical-journal-of-the-iaos/sji230020?id=statistical-journal-of-the-iaos%2Fsjj230020>