

# NÂNG CAO “SỨC ĐỀ KHÁNG” CỦA XÃ HỘI ĐỐI VỚI THÔNG TIN GIẢ, THÔNG TIN SAI LỆCH TRÊN KHÔNG GIAN MẠNG

PHẠM THỊ QUỲNH ANH  
Học viện Cảnh sát nhân dân

Nhận bài ngày 19/8/2026. Sửa chữa xong 19/9/2026. Duyệt đăng 25/9/2026.

## **Abstract**

*The rapid development of the Internet, social media, and artificial intelligence (AI) technologies has profoundly transformed how information is produced, distributed, and consumed. Alongside their socioeconomic benefits, digital environments have facilitated the dissemination of fake news, misinformation, and manipulated content at an unprecedented speed and scale. Such content not only affects individual perceptions but may also undermine social trust, influence collective behavior, disrupt economic activities, and distort decision-making processes. This article proposes measures to strengthen societal resilience to fake news and misinformation in cyberspace.*

**Keywords:** *Cyberspace, fake news, misinformation, societal resilience, solutions.*

## **1. Đặt vấn đề**

Sự phát triển nhanh chóng của Internet, mạng xã hội (MXH) và các nền tảng số đã tạo ra những thay đổi căn bản trong phương thức tiếp cận, sản xuất và chia sẻ thông tin của xã hội. Không gian mạng không chỉ mở rộng khả năng tiếp cận tri thức, thúc đẩy giao tiếp và tăng cường sự tham gia của người dân vào các hoạt động kinh tế - xã hội mà còn tạo điều kiện để thông tin được lan truyền với tốc độ và phạm vi chưa từng có. Tuy nhiên, cùng với những lợi ích đó, thông tin giả, thông tin sai lệch và các nội dung bị xuyên tạc, cắt ghép hoặc thao túng ngày càng xuất hiện với nhiều hình thức phức tạp. Sự phát triển của trí tuệ nhân tạo (AI) tạo sinh, công nghệ deepfake và các công cụ tự động hóa nội dung càng làm gia tăng khả năng tạo ra những thông tin có hình thức thuyết phục, gây khó khăn cho quá trình nhận diện và kiểm chứng. Tác động của thông tin giả không chỉ dừng lại ở việc tạo ra nhận thức không chính xác ở một bộ phận người sử dụng mà trong những trường hợp nhất định còn có thể ảnh hưởng đến hành vi cá nhân, tâm lý cộng đồng, niềm tin xã hội và hiệu quả của quá trình ra quyết định. Đáng chú ý, không phải mọi thông tin sai lệch đều xuất phát từ ý đồ cố ý lừa dối; nhiều nội dung được lan truyền do người dùng thiếu kỹ năng kiểm chứng, chia sẻ theo cảm xúc hoặc dựa trên những nguồn thông tin chưa được xác thực. Điều này cho thấy việc ứng phó với thông tin giả không thể chỉ dựa vào các biện pháp phát hiện, xử lý và gỡ bỏ nội dung mà cần đồng thời nâng cao khả năng tự bảo vệ của xã hội trước các nguy cơ thông tin. Từ góc độ đó, “sức đề kháng” của xã hội đối với thông tin giả và thông tin sai lệch có thể được hiểu là tổng hợp năng lực của cá nhân, cộng đồng và các thiết chế xã hội trong việc nhận diện, đánh giá, kiểm chứng, hạn chế sự lan truyền và giảm thiểu tác động tiêu cực của thông tin không chính xác. Xây dựng sức đề kháng cần được thực hiện trên nhiều phương diện, từ nâng cao năng lực số và năng lực thông tin của người dân, tăng cường cung cấp thông tin chính xác, minh bạch và kịp thời, phát triển hệ sinh thái kiểm chứng, đến nâng cao trách nhiệm của các nền tảng số và khả năng phối hợp giữa các chủ thể liên quan.

## **2. Nội dung nghiên cứu**

### **2.1. Về thông tin giả, thông tin sai lệch ở Việt Nam trên không gian mạng**

Trong những năm gần đây, sự phát triển mạnh mẽ của hạ tầng số đã làm thay đổi căn bản môi trường thông

Email: [phamthiquynhanht02@gmail.com](mailto:phamthiquynhanht02@gmail.com)

tin tại Việt Nam. Đầu năm 2025, Việt Nam có khoảng 79,8 triệu người sử dụng Internet, tương đương 78,8% dân số; số lượng social media user identities đạt khoảng 76,2 triệu, tương đương 75,2% dân số. Đáng chú ý, khoảng 95,4% người sử dụng Internet tại Việt Nam sử dụng ít nhất một nền tảng MXH. Quy mô kết nối này tạo ra một không gian thông tin có tốc độ lan truyền rất cao, trong đó mỗi người dùng đồng thời có thể là người tiếp nhận, sản xuất và phân phối nội dung. Đây là điều kiện thuận lợi cho tiếp cận tri thức và giao tiếp xã hội nhưng cũng làm gia tăng khả năng thông tin chưa được kiểm chứng được lan truyền nhanh chóng đến số lượng lớn người dùng. DataReportal lưu ý rằng social media user identities không hoàn toàn tương đương với số cá nhân duy nhất, bởi một người có thể sở hữu hoặc sử dụng nhiều tài khoản; do đó, số liệu cần được diễn giải thận trọng trong các nghiên cứu định lượng (DataReportal, 2025). Thực trạng hiện nay cho thấy MXH đã trở thành một bộ phận quan trọng của hệ sinh thái thông tin tại Việt Nam. Facebook có khoảng 76,2 triệu người trong phạm vi tiếp cận quảng cáo vào đầu năm 2025; TikTok có khoảng 40,9 triệu người dùng từ 18 tuổi trở lên trong dữ liệu quảng cáo của nền tảng. Những con số này không phải là số người dùng hoạt động hằng tháng nhưng cho thấy quy mô công chúng mà nội dung trên các nền tảng có thể tiếp cận. Đặc biệt, thông tin không còn vận động theo một chiều từ cơ quan báo chí đến công chúng mà được hình thành từ nhiều nguồn: báo chí, tài khoản cá nhân, nhóm cộng đồng, người có ảnh hưởng, video ngắn, livestream và các công cụ AI. Sự phân mảnh này làm cho việc xác định nguồn gốc và kiểm chứng thông tin trở nên phức tạp hơn, đồng thời tạo điều kiện để một nội dung sai lệch được sao chép và tái xuất hiện dưới nhiều hình thức khác nhau.

Thực trạng thông tin sai lệch tại Việt Nam là sự đa dạng về nội dung và phương thức thể hiện. Thông tin sai lệch có thể xuất hiện dưới dạng hoàn toàn bịa đặt, thông tin đúng nhưng bị đặt sai bối cảnh, hình ảnh hoặc video cũ được gắn với sự kiện mới, giả mạo phát ngôn, cắt ghép nội dung hoặc sử dụng tiêu đề gây hiểu nhầm. Các chủ đề liên quan trực tiếp đến đời sống như sức khỏe, dịch bệnh, thiên tai, giá cả, tài chính, sản phẩm tiêu dùng và các sự kiện xã hội thường dễ thu hút sự chú ý. Năm 2025, cơ quan quản lý cũng phải tăng cường cảnh báo trước các hình thức quảng cáo sai sự thật trên môi trường mạng, trong đó có quảng cáo liên quan đến thuốc chữa bệnh, sữa và thực phẩm bảo vệ sức khỏe giả. Điều này cho thấy ranh giới giữa thông tin sai lệch, quảng cáo gây hiểu nhầm và hành vi lừa đảo trên môi trường số ngày càng cần được nhận diện một cách hệ thống.

Vấn đề nổi bật khác là khoảng cách giữa tốc độ lan truyền thông tin và tốc độ kiểm chứng. Trên MXH, một nội dung có thể được chia sẻ, sao chép hoặc chuyển sang nền tảng khác trong thời gian rất ngắn, trong khi việc xác minh một thông tin phức tạp có thể đòi hỏi sự tham gia của cơ quan chuyên môn, nhà báo hoặc chuyên gia. Khi thông tin chính xác chưa được cung cấp kịp thời, nội dung chưa kiểm chứng có thể hình thành nhận thức ban đầu của công chúng. Đáng chú ý, việc xóa nội dung gốc cũng không đồng nghĩa với việc thông tin sai biến mất, bởi nội dung có thể đã được chụp lại, tải xuống hoặc sao chép sang nhiều tài khoản khác.

Về quy mô xử lý, các cơ quan chức năng Việt Nam đã triển khai nhiều biện pháp phối hợp với các nền tảng số. Theo thông tin của Bộ Thông tin và Truyền thông công bố trước đây, trong năm 2024 Facebook đã chặn, gỡ 8.981 nội dung vi phạm pháp luật Việt Nam; YouTube chặn, gỡ 6.043 nội dung và TikTok chặn, gỡ 971 nội dung, với tỷ lệ xử lý theo yêu cầu của cơ quan quản lý lần lượt khoảng 94%, 91% và 93% (Cục Thông tin, Thống kê - Bộ Khoa học và Công nghệ, 2025). Các con số này không thể được coi là số lượng “tin giả”, bởi phạm vi bao gồm nhiều loại nội dung vi phạm khác nhau; tuy nhiên, chúng phản ánh quy mô đáng kể của hoạt động quản lý và xử lý nội dung trên không gian mạng. Đồng thời, việc Nhà nước phải thường xuyên phối hợp với các nền tảng cho thấy thông tin sai lệch và nội dung vi phạm đã trở thành một vấn đề quản trị thường xuyên trong môi trường số.

Thay đổi quan trọng trong quản lý môi trường mạng là tăng cường xác thực danh tính người dùng. Nghị định số 147/2024/NĐ-CP, có hiệu lực từ ngày 25/12/2024, quy định người sử dụng MXH phải xác thực tài khoản bằng số điện thoại di động tại Việt Nam hoặc số định danh cá nhân; chỉ tài khoản đã được xác thực mới được đăng tải, bình luận, livestream và chia sẻ thông tin (Chính phủ, 2024). Quy định này tạo cơ sở để nâng cao trách nhiệm của người sử dụng và hạn chế việc lợi dụng tài khoản ẩn danh để phát tán nội dung sai lệch. Tuy nhiên, xác thực danh tính chỉ giải quyết một khía cạnh của vấn

đề. Việc một tài khoản có danh tính xác thực không đồng nghĩa với việc mọi nội dung do tài khoản đó đăng tải đều chính xác. Thách thức mới là sự phát triển nhanh chóng của nội dung do AI tạo ra. AI tạo sinh cho phép tạo văn bản, hình ảnh, âm thanh và video với chi phí thấp, tốc độ cao và mức độ chân thực ngày càng lớn. Điều này làm cho việc kiểm chứng thông tin dựa đơn thuần vào quan sát trực quan trở nên khó khăn hơn. Các hình thức giả mạo khuôn mặt, giọng nói hoặc video có thể được sử dụng để tạo ra những nội dung gây hiểu nhầm, đặc biệt trong các lĩnh vực nhạy cảm như tài chính, thương mại, sức khỏe và danh tính cá nhân. Trong bối cảnh Việt Nam có gần 80 triệu người sử dụng Internet, sự kết hợp giữa AI tạo sinh và MXH quy mô lớn tạo ra một thách thức mới đối với năng lực nhận diện và kiểm chứng của người dùng.

## **2.2. Những yếu tố làm gia tăng nguy cơ lan truyền thông tin sai lệch ở Việt Nam trên không gian mạng**

Tại Việt Nam, nguy cơ lan truyền thông tin sai lệch ngày càng gia tăng cùng với quá trình phổ cập Internet, điện thoại thông minh và MXH. Theo Digital 2025: Vietnam, đầu năm 2025 Việt Nam có khoảng 79,8 triệu người sử dụng Internet, tương đương 78,8% dân số; số lượng định danh người dùng MXH đạt khoảng 76,2 triệu, tương đương 75,2% dân số. Đáng chú ý, khoảng 95,4% số người sử dụng Internet tại Việt Nam sử dụng ít nhất một nền tảng MXH. Quy mô kết nối rất lớn này tạo ra điều kiện thuận lợi cho thông tin được truyền tải với tốc độ nhanh, phạm vi rộng và chi phí thấp. Đồng thời, nó cũng làm gia tăng “bề mặt lan truyền” của thông tin sai lệch, bởi một nội dung không chính xác có thể được tạo ra ở một tài khoản nhưng nhanh chóng được sao chép, bình luận, cắt ghép và chia sẻ trên nhiều nền tảng khác nhau.

Đặc điểm đáng chú ý của Việt Nam là sự phụ thuộc tương đối lớn của người dùng vào các nền tảng MXH có khả năng tạo và phân phối nội dung theo thời gian thực. Theo số liệu được Bộ Thông tin và Truyền thông công bố, đến tháng 6/2024, Việt Nam có khoảng 76,5 triệu người dùng Zalo thường xuyên hằng tháng, trong khi Facebook có khoảng 72 triệu, TikTok 67 triệu và YouTube 63 triệu người dùng (T. Thủy, 2025). Các con số này cho thấy MXH đã trở thành một bộ phận quan trọng của hệ sinh thái thông tin mà người dân sử dụng hằng ngày. Đặc biệt, sự hiện diện đồng thời của nhiều nền tảng khiến một thông tin sai lệch không nhất thiết chỉ tồn tại ở một không gian mà có thể di chuyển từ Facebook sang TikTok, YouTube, Zalo hoặc các nhóm, diễn đàn trực tuyến khác. Bộ Thông tin và Truyền thông cũng nhận định sức ảnh hưởng của các nền tảng xuyên biên giới tại Việt Nam vẫn lớn hơn đáng kể so với các nền tảng MXH trong nước, qua đó làm cho việc kiểm soát nguồn phát tán, xác định chủ thể và xử lý nội dung sai lệch trở nên phức tạp hơn.

Sự thay đổi căn bản trong cấu trúc sản xuất và phân phối thông tin. Nếu trong mô hình truyền thông truyền thống, nội dung thường phải trải qua các khâu lựa chọn, biên tập và kiểm chứng trước khi đến công chúng thì trên MXH, mỗi cá nhân đều có thể đồng thời là người tiếp nhận, sản xuất và phân phối thông tin. Tại Việt Nam, sự phát triển của các nền tảng video ngắn, livestream, nhóm cộng đồng và tài khoản cá nhân có lượng người theo dõi lớn đã làm cho ranh giới giữa “nguồn tin” và “người sử dụng thông tin” ngày càng mờ nhạt. Trong môi trường đó, mức độ phổ biến của một nội dung không đồng nghĩa với mức độ chính xác của nội dung. Một thông tin được nhiều tài khoản chia sẻ có thể tạo ra hiệu ứng “nhiều người cùng nói nên có vẻ đúng”, trong khi nguồn gốc ban đầu và bằng chứng xác thực lại ít được kiểm tra. Bên cạnh đó, cơ chế vận hành của các nền tảng số dựa nhiều vào mức độ tương tác của người dùng cũng có thể góp phần khuếch đại những nội dung gây chú ý mạnh. Những nội dung có tiêu đề gây sốc, hình ảnh nổi bật, ngôn ngữ cực đoan hoặc kích thích tò mò thường dễ tạo ra lượt xem, bình luận và chia sẻ. Trong trường hợp thông tin sai lệch được thiết kế nhằm thu hút sự chú ý, chính hành vi tương tác của người dùng có thể vô tình giúp nội dung tiếp tục được phân phối tới nhiều người hơn. Vấn đề này đặc biệt đáng chú ý tại Việt Nam khi số lượng người dùng MXH rất lớn và thời gian dành cho các nền tảng số ngày càng tăng như Facebook có khoảng 76,2 triệu người có thể tiếp cận thông tin qua hệ thống quảng cáo tại Việt Nam; YouTube có khoảng 62,3 triệu và TikTok có khoảng 40,9 triệu người dùng từ 18 tuổi trở lên theo số liệu tiếp cận quảng cáo của nền tảng (Nielsen, R. K, 2025).

Yếu tố tâm lý và hành vi của người sử dụng cũng đóng vai trò quan trọng trong quá trình lan truyền thông tin sai lệch. Người dùng thường có xu hướng chú ý và chia sẻ những nội dung phù hợp với nhận

thức sẵn có, đáp ứng nhu cầu cảm xúc hoặc liên quan trực tiếp đến lợi ích của bản thân. Trong những tình huống xảy ra sự kiện bất thường, thiên tai, dịch bệnh, biến động thị trường hoặc các vụ việc xã hội được quan tâm, nhu cầu tìm kiếm thông tin tăng nhanh trong khi thông tin chính xác có thể chưa được công bố đầy đủ. Khoảng trống thông tin này tạo điều kiện để các nội dung chưa được kiểm chứng xuất hiện và lan truyền. Đáng chú ý, không phải mọi hành vi phát tán thông tin sai đều xuất phát từ mục đích cố ý lừa dối. Một bộ phận người dùng có thể chia sẻ thông tin vì tin rằng thông tin đó là đúng, muốn cảnh báo người khác hoặc đơn giản chỉ vì thấy nội dung được nhiều người chia sẻ. Chính cơ chế “chia sẻ do tin tưởng” khiến thông tin sai lệch có khả năng lan truyền tự phát và khó kiểm soát. Sự tồn tại của các tài khoản ẩn danh, tài khoản giả mạo và các hình thức sử dụng danh tính không xác thực. Khi người phát tán thông tin không phải chịu trách nhiệm rõ ràng về danh tính, rào cản tâm lý đối với việc đăng tải hoặc chia sẻ thông tin chưa được kiểm chứng có thể giảm xuống. Thực tiễn quản lý tại Việt Nam cho thấy đây là một vấn đề được cơ quan quản lý đặc biệt quan tâm. Nghị định số 147/2024/NĐ-CP, có hiệu lực từ ngày 25/12/2024 quy định MXH phải xác thực người dùng bằng số điện thoại hoặc số định danh cá nhân và chỉ tài khoản đã xác thực mới được đăng tải, bình luận, livestream và chia sẻ thông tin trên MXH (Chính phủ, 2024). Việc ban hành quy định này cho thấy quản lý danh tính số được xem là một công cụ nhằm hạn chế việc lợi dụng tính ẩn danh để phát tán tin giả, lừa đảo và các nội dung vi phạm pháp luật. Đáng chú ý khác là khoảng cách giữa tốc độ lan truyền thông tin trên MXH và tốc độ kiểm chứng của các thiết chế chính thức. Một nội dung chỉ cần vài phút để được chia sẻ rộng rãi nhưng việc xác minh một vấn đề phức tạp có thể cần nhiều giờ hoặc nhiều ngày. Đặc biệt, khi thông tin liên quan đến những lĩnh vực chuyên môn như tài chính, y tế, pháp luật, khoa học hoặc các sự kiện xã hội phức tạp, người dùng phổ thông thường khó tự xác định tính chính xác nếu không có nguồn chuyên môn đáng tin cậy. Vì vậy, khi thông tin chính xác được cung cấp chậm hoặc khó tiếp cận, thông tin chưa kiểm chứng có thể chiếm ưu thế trong giai đoạn đầu của quá trình hình thành nhận thức xã hội. Thực tế xử lý thông tin vi phạm tại Việt Nam cho thấy quy mô của vấn đề không nhỏ. Theo Bộ Thông tin và Truyền thông (2024), Facebook đã chặn, gỡ 8.981 nội dung vi phạm pháp luật Việt Nam, Google chặn, gỡ 6.043 nội dung trên YouTube và TikTok chặn, gỡ 971 nội dung, với tỷ lệ xử lý lần lượt khoảng 94%, 91% và 93% theo yêu cầu của cơ quan quản lý. Các số liệu này không phải là thống kê toàn bộ thông tin sai lệch trên Internet và cũng không đồng nhất với số lượng “tin giả”, bởi phạm vi xử lý bao gồm nhiều loại nội dung vi phạm khác nhau. Tuy nhiên, chúng phản ánh quy mô đáng kể của nhu cầu phát hiện, xác minh và xử lý nội dung có vấn đề trên các nền tảng lớn tại Việt Nam.

### **2.3. Giải pháp nâng cao “sức đề kháng” của xã hội đối với thông tin giả, thông tin sai lệch trên không gian mạng**

#### **2.3.1. Nâng cao năng lực số và năng lực kiểm chứng thông tin của người dân**

Nâng cao năng lực số và năng lực kiểm chứng thông tin cần được xem là giải pháp nền tảng để hình thành “sức đề kháng” của xã hội. Việt Nam hiện có khoảng 79,8 triệu người sử dụng Internet và 76,2 triệu định danh người dùng MXH vào đầu năm 2025; khoảng 95,4% người sử dụng Internet có sử dụng ít nhất một nền tảng MXH (Reuters Institute for the Study of Journalism, 2025). Quy mô người dùng lớn cho thấy việc nâng cao năng lực tự bảo vệ của cá nhân có ý nghĩa đặc biệt quan trọng, bởi không thể xử lý toàn bộ thông tin sai lệch chỉ bằng cơ chế quản lý hoặc gỡ bỏ nội dung. Bộ Thông tin và Truyền thông trước đây đã xây dựng khung chương trình nâng cao kỹ năng số cho người dân, trong đó bao gồm kỹ năng tìm kiếm và khai thác thông tin trên Internet, giao tiếp và phối hợp trên môi trường số, sáng tạo nội dung, an toàn thông tin và giải quyết vấn đề. Do đó, cần chuyển từ cách tiếp cận “phổ biến kiến thức chung” sang đào tạo kỹ năng thực hành. Người dân cần được hướng dẫn một quy trình đơn giản trước khi chia sẻ thông tin, chẳng hạn: kiểm tra nguồn phát hành; kiểm tra thời gian và bối cảnh; đối chiếu với ít nhất một nguồn độc lập; kiểm tra hình ảnh, video và trích dẫn; sau đó mới quyết định chia sẻ. Đặc biệt, giáo dục kỹ năng thông tin cần được triển khai theo từng nhóm đối tượng, chú ý đến học sinh, sinh viên, người cao tuổi, người dân ở khu vực nông thôn và những nhóm ít có điều kiện tiếp cận các nguồn thông tin chuyên môn.

### 2.3.2. Chủ động cung cấp thông tin chính xác, kịp thời và dễ tiếp cận

Một xã hội có khả năng chống chịu tốt trước thông tin sai lệch không chỉ cần biết “tin nào sai” mà còn phải dễ dàng tìm được thông tin đúng. Vì vậy, các cơ quan nhà nước, tổ chức chuyên môn và cơ quan báo chí cần tăng cường cung cấp thông tin chủ động, đặc biệt đối với những vấn đề có khả năng gây hoang mang hoặc tác động trực tiếp đến đời sống người dân. Khoảng cách giữa thời điểm xuất hiện một thông tin gây chú ý và thời điểm xuất hiện thông tin chính xác có thể tạo ra “khoảng trống thông tin”, qua đó tạo điều kiện để nội dung chưa được kiểm chứng chiếm ưu thế. Do đó, phản hồi truyền thông cần được thiết kế theo nguyên tắc nhanh - chính xác - dễ hiểu - có bằng chứng - có nguồn để kiểm tra. Thực tiễn năm 2025 cho thấy các cơ quan quản lý ở Việt Nam đã phải chủ động sử dụng nhiều kênh truyền thông để cảnh báo người dân trước thông tin sai lệch, trong đó có các nội dung quảng cáo liên quan đến thuốc chữa bệnh, sữa và thực phẩm bảo vệ sức khỏe giả. Bộ Văn hóa, Thể thao và Du lịch cho biết đang phối hợp với Bộ Công an, Bộ Y tế và các cơ quan liên quan để rà soát, xử lý các sai phạm về quảng cáo trên báo chí, môi trường mạng và nền tảng xã hội, đồng thời triển khai việc gửi tin nhắn khuyến cáo người dân nâng cao cảnh giác.

### 2.3.3. Xây dựng và mở rộng hệ sinh thái kiểm chứng thông tin

Cần hình thành một hệ sinh thái kiểm chứng thông tin đa chủ thể, trong đó cơ quan báo chí, cơ quan quản lý, cơ sở giáo dục, viện nghiên cứu, chuyên gia và các tổ chức kiểm chứng có thể phối hợp để xác minh những nội dung có khả năng tác động lớn đến xã hội. Hệ thống này không nên chỉ tập trung vào việc tuyên bố một nội dung “đúng” hay “sai” mà cần giải thích căn cứ, nguồn dữ liệu và phương pháp kiểm chứng để người dân có khả năng tự đánh giá. Đây cũng là cách tăng tính minh bạch và góp phần nâng cao niềm tin vào thông tin chính xác. Digital News Report 2025 cho thấy khi người dân nghi ngờ một thông tin trực tuyến là sai hoặc gây hiểu nhầm, các nguồn tin báo chí đáng tin cậy, nguồn chính thức và công cụ tìm kiếm là những nơi thường được lựa chọn để kiểm tra; các tổ chức kiểm chứng cũng giữ một vai trò đáng kể. Tuy nhiên, báo cáo đồng thời lưu ý rằng dữ liệu là tự khai báo và có thể không phản ánh hoàn toàn hành vi thực tế. Đối với Việt Nam, điều này gợi ý cần làm cho các sản phẩm kiểm chứng dễ tìm kiếm trên Google và các nền tảng xã hội, ngắn gọn, trực quan và xuất hiện gần với nội dung sai lệch, thay vì chỉ đăng trên một trang web chuyên biệt mà người dùng ít tiếp cận.

### 2.3.4. Nâng cao trách nhiệm của các nền tảng số và tăng cường xác thực tài khoản

Các nền tảng MXH giữ vị trí quan trọng trong việc phân phối nội dung nên cần được xem là một chủ thể trung tâm trong hệ thống phòng ngừa thông tin sai lệch. Các biện pháp cần tập trung vào xác thực tài khoản, xử lý tài khoản giả mạo, hạn chế hành vi thao túng có tổ chức, cải thiện cơ chế báo cáo nội dung và tăng tính minh bạch của hệ thống đề xuất. Đồng thời, cần phân biệt giữa thông tin sai do nhầm lẫn, nội dung cố ý lừa dối và ý kiến cá nhân, qua đó lựa chọn biện pháp xử lý tương ứng. Tại Việt Nam, Nghị định 147/2024/NĐ-CP có hiệu lực từ ngày 25/12/2024 đã tạo cơ sở pháp lý quan trọng cho việc xác thực người sử dụng MXH (Chính phủ, 2024). Theo quy định, chỉ tài khoản đã xác thực mới được đăng tải, bình luận, livestream và chia sẻ thông tin; với một số trường hợp livestream thương mại, yêu cầu xác thực bằng số định danh cá nhân được áp dụng. Nghị định cũng quy định cơ chế khóa tạm thời hoặc vĩnh viễn đối với các tài khoản, trang, nhóm hoặc kênh nội dung thường xuyên vi phạm. Tuy nhiên, xác thực danh tính cần được coi là một lớp bảo vệ, không phải giải pháp duy nhất; bởi tài khoản xác thực vẫn có thể đăng tải thông tin không chính xác.

### 2.3.5. Tăng cường năng lực nhận diện và ứng phó với nội dung do AI tạo ra

Sự phát triển của AI tạo sinh đòi hỏi Việt Nam phải bổ sung một lớp năng lực mới vào chương trình giáo dục công dân số. Các công cụ AI hiện có khả năng tạo văn bản, hình ảnh, âm thanh và video với chi phí thấp, làm giảm đáng kể rào cản kỹ thuật đối với việc sản xuất nội dung giả hoặc gây hiểu nhầm. World Economic Forum nhận định AI đang làm gia tăng khối lượng nội dung sai lệch và khiến việc phân biệt nội dung do AI tạo ra với nội dung do con người tạo ra khó khăn hơn; đồng thời, misinformation và disinformation tiếp tục nằm trong nhóm rủi ro ngắn hạn hàng đầu trong Global Risks Report 2025. Do đó, cần phổ cập cho người dân nguyên tắc “nội dung nhìn thấy không mặc nhiên là bằng chứng”. Đối với hình ảnh hoặc video có dấu hiệu bất thường, người dùng cần kiểm tra nguồn ban đầu, thời gian xuất hiện, bối cảnh và đối chiếu với các nguồn độc lập. Về phía các cơ quan báo chí và nền tảng cần nghiên cứu các

giải pháp xác định nguồn gốc nội dung, đánh dấu hoặc cung cấp thông tin về quá trình tạo và chỉnh sửa nội dung. Đồng thời, tránh phụ thuộc tuyệt đối vào các công cụ “phát hiện AI”, bởi chính các công cụ này cũng có giới hạn về độ chính xác và có thể tạo ra kết quả dương tính hoặc âm tính giả.

### 2.3.6. Tăng cường phối hợp liên ngành và đo lường sức đề kháng thông tin

UNESCO nhấn mạnh quản trị nền tảng số và bảo đảm tính toàn vẹn của môi trường thông tin cần sự tham gia của Nhà nước, nền tảng công nghệ, báo chí, giới học thuật, xã hội dân sự và cộng đồng kỹ thuật; các hướng dẫn của UNESCO được xây dựng thông qua quá trình tham vấn với hơn 10.000 ý kiến từ 134 quốc gia, cho thấy tính chất đa chủ thể của vấn đề này (Reuters Institute for the Study of Journalism, 2025). Đối với Việt Nam, cơ chế phối hợp cần bao gồm ít nhất bốn chức năng: phát hiện sớm thông tin có nguy cơ gây tác động lớn; xác minh nhanh với cơ quan chuyên môn; truyền tải thông tin chính xác qua nhiều kênh; và đánh giá mức độ lan truyền sau khi đã có thông tin chính xác. Bên cạnh đó, cần xây dựng bộ chỉ số đo lường “sức đề kháng” của xã hội. Không nên chỉ đánh giá hiệu quả bằng số lượng nội dung bị gỡ bỏ hoặc số tài khoản bị xử lý. Các chỉ số quan trọng hơn có thể bao gồm tỷ lệ người dân biết kiểm tra nguồn tin, tỷ lệ người thực hiện kiểm chứng trước khi chia sẻ, thời gian từ khi phát hiện thông tin sai đến khi có thông tin xác minh, khả năng tiếp cận thông tin chính xác và mức độ nhận diện nội dung do AI tạo ra.

### 3. Kết luận

Thông tin giả, thông tin sai lệch trên không gian mạng đang trở thành một thách thức đáng chú ý đối với đời sống xã hội trong bối cảnh chuyển đổi số và sự phát triển nhanh chóng của các công nghệ mới. Đặc điểm lan truyền nhanh, phạm vi tiếp cận rộng, khả năng tạo lập nội dung ngày càng dễ dàng và sự hỗ trợ của các công nghệ AI khiến việc nhận diện, kiểm chứng và hạn chế tác động của thông tin sai lệch ngày càng phức tạp. Vì vậy, ứng phó với thông tin giả không thể chỉ được xem là nhiệm vụ của cơ quan quản lý hoặc các nền tảng công nghệ mà cần trở thành trách nhiệm và năng lực chung của toàn xã hội. Nâng cao “sức đề kháng” của xã hội cần được thực hiện theo hướng chủ động, lâu dài và đồng bộ. Trong đó, nâng cao năng lực số, năng lực thông tin và tư duy phản biện của người dân là nền tảng quan trọng; đồng thời cần xây dựng môi trường thông tin, trong đó các nguồn tin chính xác, minh bạch, kịp thời và dễ tiếp cận được cung cấp cho cộng đồng. Ngoài ra, cần phát triển các cơ chế kiểm chứng thông tin, tăng cường trách nhiệm của nền tảng số, nâng cao khả năng nhận diện nội dung do AI tạo ra và hoàn thiện cơ chế phối hợp giữa cơ quan quản lý, báo chí, cơ sở giáo dục, doanh nghiệp công nghệ, chuyên gia và cộng đồng. Hiệu quả của việc phòng ngừa thông tin sai lệch không nên chỉ được đánh giá thông qua số lượng nội dung bị phát hiện hoặc xử lý mà cần chú trọng đến khả năng tự kiểm chứng và tự điều chỉnh của người sử dụng, tốc độ cung cấp thông tin chính xác, mức độ minh bạch của hệ sinh thái thông tin và khả năng phục hồi của cộng đồng sau các sự cố thông tin. Khi các yếu tố này được củng cố đồng thời, xã hội sẽ có khả năng chủ động hơn trước những biến đổi của môi trường truyền thông số. Do đó, xây dựng “sức đề kháng” thông tin cần được xác định là một thành tố quan trọng của năng lực công dân số và năng lực thích ứng xã hội trong kỷ nguyên số.

#### Tài liệu tham khảo

- Bộ Thông tin và Truyền thông (2024). *Thông tin điện tử năm 2024: Đóng góp tích cực cho đời sống xã hội*. <https://mst.gov.vn/thong-tin-dien-tu-nam-2024-dong-gop-tich-cuc-cho-doi-song-xa-hoi-197241225172400673.htm>, ngày 21/12/2024.
- Chính phủ (2024). *Nghị định số 147/2024/NĐ-CP ngày 09/11/2024 quản lý, cung cấp, sử dụng dịch vụ Internet và thông tin trên mạng*.
- Cục Thông tin, Thống kê - Bộ Khoa học và Công nghệ (2025). *Nghiên cứu xây dựng Bộ chỉ số đánh giá các hành vi ứng xử trên mạng xã hội và đề xuất áp dụng tại Việt Nam*. <https://vista.gov.vn/vi/news/ket-qua-nghien-cuu-trien-khai/nghien-cuu-xay-dung-bo-chi-so-danh-gia-cac-hanh-vi-ung-xu-tren-mang-xa-hoi-va-de-xuat-ap-dung-tai-viet-nam-11715.html>, ngày 11/8/2025.
- DataReportal (2025). *Digital 2025: Vietnam*. Báo cáo số liệu về Internet, mạng xã hội và các nền tảng số tại Việt Nam năm 2025.
- Nielsen, R. K. (2025). *How the public checks information it thinks might be wrong*. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025/how-public-checks-information-it-thinks-might-be-wrong>, ngày 17/6/2025.
- Reuters Institute for the Study of Journalism. (2025). *Digital News Report 2025*. University of Oxford.
- T. Thủy (2025). *Gửi tin nhắn khuyến cáo người dân nâng cao cảnh giác đối với các thông tin quảng cáo*. <https://bvhttdl.gov.vn/gui-tin-nhan-kyuyen-cao-nguoi-dan-nang-cao-canh-giac-doi-voi-cac-thong-tin-quang-cao-2025052608560936.htm>, ngày 26/5/2025.