

Bảo vệ chủ quyền tư tưởng trên không gian mạng trong kỷ nguyên trí tuệ nhân tạo tạo sinh: Thách thức từ các nền tảng xuyên biên giới và chiến lược thích ứng của báo chí đối ngoại Việt Nam

TS. PHẠM BÌNH DƯƠNG

Học viện Báo chí và Tuyên truyền; Email: binhduongajc@gmail.com

Nhận ngày 15 tháng 12 năm 2025; chấp nhận đăng tháng 01 năm 2026.

Tóm tắt: Năm 2025 là một năm với nhiều sự kiện quan trọng, đất nước đang chuẩn bị tâm thế cho kỷ nguyên vươn mình của dân tộc, sự bùng nổ của công nghệ Trí tuệ nhân tạo tạo sinh (Generative AI) đã làm thay đổi căn bản phương thức sản xuất và lan truyền thông tin. Bối cảnh mới đặt ra yêu cầu làm rõ nội hàm khái niệm “chủ quyền tư tưởng trên không gian mạng” trong tình hình mới; phân tích những thách thức phi truyền thống từ các nền tảng xuyên biên giới dưới sự hỗ trợ của thuật toán thông minh. Từ đó, đề xuất các chiến lược thích ứng cho báo chí đối ngoại Việt Nam nhằm chủ động chiếm lĩnh trận địa thông tin, bảo vệ nền tảng tư tưởng của Đảng Cộng sản Việt Nam trên trường quốc tế.

Từ khóa: chủ quyền tư tưởng; trí tuệ nhân tạo tạo sinh; báo chí đối ngoại; kỷ nguyên vươn mình; nền tảng xuyên biên giới.

Abstract: The year 2025 marks a period of significant events, as Vietnam prepares itself for the era of the national rise. The rapid advancement of Generative Artificial Intelligence (Generative AI) has fundamentally transformed the modes of information production and dissemination. This new context necessitates a clearer conceptualization of “ideological sovereignty in cyberspace” under new conditions, as well as an analysis of non-traditional challenges posed by cross-border platforms supported by intelligent algorithms. On that basis, the article proposes adaptive strategies for Vietnam’s external communications journalism, aiming to proactively dominate the information space and safeguard the ideological foundation of the Communist Party of Vietnam on the international stage.

Keywords: ideological sovereignty; generative artificial intelligence; external communications journalism; era of national rise; cross-border platforms.

1. Mở đầu

Kể từ sau Đại hội đại biểu toàn quốc lần thứ XIII của Đảng Cộng sản Việt Nam và các Hội nghị Trung ương khóa XIII, cụm từ “Kỷ nguyên vươn mình của dân tộc” đã trở thành kim chỉ nam cho mọi hành động chiến lược, đánh dấu bước chuyển mình từ giai đoạn “giữ gìn, ổn định” sang giai đoạn “bứt phá, khẳng định vị thế”. Tuy nhiên, sự vươn mình này diễn ra trong bối cảnh cuộc Cách mạng công nghiệp lần thứ tư đã bước sang

một trang mới với sự thống trị của trí tuệ nhân tạo tạo sinh (Generative AI).

Đến năm 2025, không gian mạng không còn là thế giới ảo mà đã trở thành “lãnh thổ số” - nơi diễn ra các cuộc đấu tranh gay gắt về tư tưởng và văn hóa. Các thế lực thù địch và phản động không chỉ sử dụng con người để chống phá mà đã tận dụng tối đa khả năng của trí tuệ nhân tạo để sản xuất thông tin sai lệch với quy mô và tốc độ chưa từng có. Trong bối cảnh đó, nhiệm vụ bảo vệ nền

tầng tư tưởng của Đảng không chỉ là ngăn chặn thông tin xấu độc, mà cốt lõi là phải khẳng định được chủ quyền tư tưởng quốc gia. Báo chí đối ngoại, với vai trò là lực lượng tiên phong đưa hình ảnh Việt Nam ra thế giới, đang đứng trước yêu cầu cấp bách phải đổi mới toàn diện để thích ứng với thách thức từ các nền tảng xuyên biên giới.

2. Nội dung

2.1. Một số khái niệm mới

Để giải quyết thấu đáo vấn đề, trước hết cần làm rõ một số nội hàm khái niệm dưới góc độ lý luận chính trị và truyền thông hiện đại.

Chủ quyền tư tưởng trên không gian mạng

Theo quan điểm truyền thống, chủ quyền quốc gia gắn liền với biên giới lãnh thổ, vùng trời và vùng biển. Tuy nhiên, trong kỷ nguyên số, khái niệm chủ quyền cần được mở rộng. Chủ quyền tư tưởng trên không gian mạng được hiểu là quyền tối cao và tuyệt đối của quốc gia trong việc quản lý, định hướng và bảo vệ hệ giá trị tư tưởng, văn hóa cốt lõi của dân tộc trên môi trường Internet, độc lập trước sự can thiệp, xâm lấn của các trào lưu tư tưởng ngoại lai hoặc thù địch.

Tại Hội nghị Trung ương 10 khóa XIII (năm 2024), Đảng Cộng sản Việt Nam đã nhấn mạnh yêu cầu kiên định chủ nghĩa Mác - Lênin, tư tưởng Hồ Chí Minh, đồng thời phải chủ động định hướng dư luận trong bối cảnh toàn cầu hóa. Như vậy, chủ quyền tư tưởng không có nghĩa là đóng cửa, cô lập thông tin, mà là năng lực làm chủ dòng chảy thông tin, đảm bảo dòng chủ lưu của báo chí cách mạng luôn giữ vai trò dẫn dắt, chi phối nhận thức xã hội, ngay cả trên các nền tảng số do nước ngoài cung cấp.

Trí tuệ nhân tạo tạo sinh và “Thực tế tổng hợp”

Khác với Trí tuệ nhân tạo truyền thống chỉ dừng lại ở việc phân tích dữ liệu, trí tuệ nhân tạo tạo sinh (Generative AI) có khả năng tự tạo ra nội dung mới (văn bản, hình ảnh, âm thanh, video) với độ chân thực gần như tuyệt đối.

Trong lĩnh vực chính trị - truyền thông, sự xuất

hiện của Trí tuệ nhân tạo tạo sinh dẫn đến nguy cơ hình thành một “Thực tế tổng hợp”. Đây là trạng thái môi trường thông tin mà ở đó ranh giới giữa sự thật và hư cấu bị xóa nhòa. Một video giả mạo (Deepfake) về một nhà lãnh đạo phát biểu những nội dung đi ngược lại đường lối của Đảng có thể được tạo ra trong vài phút và lan truyền tới hàng triệu người trước khi cơ quan chức năng kịp đính chính. Đây là một khái niệm mới mà lý luận báo chí cần bổ sung để nhận diện đúng bản chất của cuộc đấu tranh tư tưởng hiện nay.

2.2. Thách thức từ các nền tảng xuyên biên giới trong bối cảnh mới

Tính đến tháng 10 năm 2025, các nền tảng mạng xã hội xuyên biên giới không còn đóng vai trò là nhà cung cấp hạ tầng trung lập, mà đã trở thành những “người gác cổng thuật toán”. Sự kết hợp giữa các nền tảng này và trí tuệ nhân tạo tạo sinh tạo ra ba thách thức lớn đối với công tác bảo vệ nền tảng tư tưởng của Đảng, chuyển dịch từ các nguy cơ an ninh truyền thống sang an ninh phi truyền thống với độ phức tạp cao.

Sự tấn công của “Vũ khí thông tin tự động hóa” và chiến tranh tâm lý diện rộng

Các thế lực thù địch hiện nay không cần duy trì một đội ngũ nhân sự lớn để viết bài chống phá theo phương thức “tác chiến mạng” thủ công như trước đây. Thay vào đó, chúng sử dụng các mô hình ngôn ngữ lớn được tinh chỉnh để tự động sản xuất hàng nghìn bài viết, bình luận với nội dung xuyên tạc lịch sử, phủ nhận thành quả cách mạng của Việt Nam mỗi ngày. Nguy hiểm hơn, công nghệ giả mạo sâu - deepfake đã cho phép tạo ra các bằng chứng giả mạo có độ chân thực cao để tấn công vào niềm tin của công chúng.

Sự chuyển dịch này thể hiện rõ qua sự nâng cấp của các chiến dịch chống phá. Nếu như năm 2024, thế giới chứng kiến chiến dịch “Doppelgänger”, được biết đến dưới tên “Kẻ song trùng” sử dụng trí tuệ nhân tạo để giả mạo giao diện các tờ báo uy tín tại châu Âu, thì đến năm 2025, kịch bản tương tự đã xuất hiện nhắm vào Việt Nam.

Các hệ thống giám sát an ninh mạng đã phát hiện các trang tin giả mạo giao diện của các báo điện tử chủ lực với tên miền thay đổi rất nhỏ, đăng tải các bài viết về “khủng hoảng tài chính nội bộ” do trí tuệ nhân tạo biên soạn với văn phong báo chí chuẩn xác. Mục tiêu không chỉ là gây nhiễu loạn thông tin mà còn nhằm kích động tâm lý hoang mang, gây bất ổn thị trường chứng khoán và kinh tế vĩ mô.

Bên cạnh đó, công nghệ sao chép giọng nói cũng đang bị lạm dụng triệt để. Từ bài học về các cuộc gọi tự động giả giọng chính trị gia trong các cuộc bầu cử tại phương Tây, các đối tượng phản động đã sử dụng công nghệ này để giả giọng lãnh đạo địa phương tại Việt Nam, thực hiện các cuộc gọi kích động người dân tại các điểm nóng về đất đai, gây khó khăn lớn cho công tác dân vận và ổn định tình hình cơ sở.

Cơ chế “Buồng vang thông tin” và sự phân cực xã hội qua các tiểu văn hóa số

Thuật toán gợi ý của các nền tảng video ngắn như: TikTok, YouTube Shorts, Reels... hoạt động dựa trên việc tối ưu hóa thời gian giữ chân người dùng. Các nghiên cứu chỉ ra rằng nội dung gây sốc, đánh vào cảm xúc tiêu cực hoặc sự phẫn nộ thường được thuật toán ưu tiên lan truyền mạnh mẽ hơn tin tức tích cực. Trí tuệ nhân tạo sẽ học hành vi của người dùng và liên tục “mớm” cho họ những thông tin một chiều, tạo ra các “buồng vang thông tin”.

Hiện tượng “gây nghiện bất bình - phẫn nộ” đang tác động trực tiếp đến giới trẻ - thế hệ Gen Z và Gen Alpha. Các dữ liệu xã hội học truyền thông năm 2024 - 2025 cho thấy, nếu một người dùng trẻ dừng lại xem một video có nội dung than vãn, chê bai chế độ đãi ngộ xã hội khoảng 15 giây, thuật toán sẽ ngay lập tức đề xuất liên tiếp hàng chục video khác có nội dung tương tự với mức độ cực đoan tăng dần. Đây là cơ chế hình thành nên các “nhóm cô lập tư tưởng”, nơi các quan điểm đối lập bị triệt tiêu, tạo điều kiện cho các tổ chức phản động gieo rắc tư tưởng ly khai, biệt lập với

dòng chảy chính trị chủ lưu.

Đáng lo ngại hơn là sự xuất hiện của trào lưu “lịch sử thay thế” trên các nền tảng chia sẻ video. Các đối tượng xấu sử dụng trí tuệ nhân tạo tạo sinh hình ảnh để minh họa lại các sự kiện lịch sử Việt Nam nhưng lồng ghép các chi tiết sai lệch về trang phục, bối cảnh văn hóa hoặc bóp méo sự thật lịch sử như các cuộc chiến tranh biên giới hoặc chủ quyền lãnh thổ. Khi người xem tranh luận trái chiều, thuật toán lại càng đẩy các video này lên xu hướng, khiến nhận thức lịch sử của một bộ phận giới trẻ bị lệch lạc, dẫn đến nguy cơ “xâm lăng văn hóa” ngay trên không gian số.

2.3. Thách thức về “Chủ quyền dữ liệu” và thiên kiến văn hóa trong đào tạo trí tuệ nhân tạo

Một vấn đề ít được chú ý nhưng mang tính chiến lược là việc các mô hình trí tuệ nhân tạo toàn cầu được huấn luyện chủ yếu dựa trên dữ liệu phương Tây hoặc dữ liệu không được kiểm chứng từ các nguồn mở thiếu tin cậy. Khi người dùng Việt Nam tra cứu các vấn đề lịch sử, chính trị, chủ quyền biển đảo trên các công cụ này, kết quả trả về thường mang thiên kiến, thiếu chính xác hoặc sai lệch nghiêm trọng.

Thực tế đã chứng minh sự tồn tại của “lỗi thiên kiến văn hóa”. Khi thử nghiệm yêu cầu các công cụ Trí tuệ nhân tạo phổ biến vẽ hình ảnh về văn hóa truyền thống Việt Nam, ví dụ như ngày Tết Nguyên đán, kết quả thường bị pha trộn các yếu tố văn hóa nước ngoài như trang phục, kiến trúc của các quốc gia lân cận... do thiếu hụt dữ liệu đầu vào thuần Việt.

Nghiêm trọng hơn là vấn đề “ảo giác trí tuệ nhân tạo” liên quan đến chủ quyền quốc gia. Trong một số bản cập nhật bản đồ số hoặc ứng dụng du lịch tích hợp trợ lý ảo quốc tế giai đoạn 2024 - 2025, khi người dùng truy vấn về chủ quyền hai quần đảo Hoàng Sa và Trường Sa, hệ thống đã tự động trích xuất dữ liệu từ các nguồn không chính thống và đưa ra câu trả lời sai lệch, xâm phạm chủ quyền của Việt Nam. Điều này cho thấy, nếu chúng ta không nắm giữ “chủ quyền dữ

liệu”, tức là quyền kiểm soát nguồn tri thức dùng để dạy cho máy móc, chúng ta sẽ đối mặt với nguy cơ bị nô dịch về mặt nhận thức ngay từ gốc rễ, khi thể hệ tương lai nhìn nhận đất nước qua lăng kính méo mó của các thuật toán ngoại lai.

2.4. Chiến lược thích ứng và giải pháp cho báo chí đối ngoại Việt Nam

Trong kỷ nguyên vươn mình của dân tộc, báo chí đối ngoại không thể mãi ở thế “phòng ngự, phản bác”. Chúng ta cần chuyển sang thế “tấn công, dẫn dắt” bằng cách tận dụng chính công nghệ Trí tuệ nhân tạo.

Chuyển đổi mô hình: Từ “thông tin một chiều” sang “đối thoại đa chiều, cá nhân hóa”

Báo chí đối ngoại truyền thống thường nặng về đưa tin nghị quyết, văn bản, khô cứng và thiếu hấp dẫn. Với sự hỗ trợ của Trí tuệ nhân tạo, các cơ quan báo chí chủ lực như: Thông tấn xã Việt Nam, Báo Nhân Dân, Đài Truyền hình Việt Nam cần xây dựng các bản tin được cá nhân hóa cho từng khu vực địa lý và đối tượng độc giả quốc tế.

Sử dụng trí tuệ nhân tạo để phân tích dữ liệu lớn (Big Data) về mối quan tâm của công chúng tại các nước sở tại. Ví dụ, đối với độc giả tại Hợp chủng quốc Hoa Kỳ, báo chí đối ngoại cần tập trung vào các câu chuyện về hợp tác kinh tế, đổi mới sáng tạo và nhân quyền tại Việt Nam thông qua các câu chuyện của con người cụ thể, thay vì chỉ đưa tin về các cuộc họp. Trí tuệ nhân tạo có thể giúp biên tập lại nội dung, giọng văn sao cho phù hợp với văn hóa đọc của từng quốc gia mà vẫn giữ vững định hướng chính trị.

Ứng dụng trí tuệ nhân tạo tạo sinh để sản xuất nội dung đa phương tiện chất lượng cao

Để cạnh tranh với thông tin xấu độc trên các nền tảng xuyên biên giới, thông tin chính thống phải hấp dẫn hơn, nhanh hơn và đẹp hơn. Các cơ quan báo chí đối ngoại cần triển khai các MC ảo có khả năng dẫn chương trình bằng 30 - 50 ngôn ngữ khác nhau, hoạt động 24/7. Điều này giúp thông tin về chủ trương, chính sách của Đảng đến được với bạn bè quốc tế ngay tức thì, xóa bỏ rào

cản ngôn ngữ. Đồng thời, sử dụng Trí tuệ nhân tạo để tạo ra các sản phẩm báo chí dữ liệu, báo chí đồ họa trực quan hóa thành tựu phát triển kinh tế - xã hội của Việt Nam, bác bỏ các luận điệu xuyên tạc về “kinh tế Việt Nam sụp đổ” hay “bất ổn chính trị”.

Xây dựng “hệ miễn dịch số” và liên minh báo chí quốc tế

Bảo vệ chủ quyền tư tưởng không thể đơn độc. Báo chí đối ngoại Việt Nam cần chủ động kết nối, chia sẻ dữ liệu với các hãng thông tấn lớn trên thế giới và các nền tảng mạng xã hội để thiết lập cơ chế “dán nhãn tin giả”.

Cần kiến nghị xây dựng hành lang pháp lý buộc các nền tảng xuyên biên giới phải minh bạch thuật toán và hợp tác gỡ bỏ các nội dung deepfake vi phạm pháp luật Việt Nam trong thời gian thực. Song song đó, việc giáo dục năng lực truyền thông số cho công chúng, đặc biệt là giới trẻ, phải được coi là một nhiệm vụ chính trị, giúp họ có khả năng tự đề kháng trước các thông tin sai lệch do trí tuệ nhân tạo tạo ra.

3. Kết luận

Kỷ nguyên vươn mình của dân tộc đặt ra cho công tác tư tưởng những yêu cầu mới, cao hơn và phức tạp hơn. Việc bảo vệ nền tảng tư tưởng của Đảng trên không gian mạng hiện nay không đơn thuần là cuộc đấu tranh giữa con người với con người, mà là cuộc chạy đua về năng lực công nghệ và quản trị dữ liệu.

Trí tuệ nhân tạo tạo sinh vừa là thách thức, nhưng cũng là cơ hội để báo chí đối ngoại Việt Nam hiện đại hóa, nâng cao sức cạnh tranh và khả năng lan tỏa. Để làm được điều đó, chúng ta cần một tư duy mới: Lấy công nghệ làm công cụ, lấy bản sắc văn hóa và lý luận chính trị làm nền tảng, lấy lợi ích quốc gia - dân tộc làm mục tiêu tối thượng. Chỉ khi làm chủ được công nghệ và chủ động chiếm lĩnh không gian mạng, chúng ta mới có thể bảo vệ vững chắc chủ quyền tư tưởng, góp phần hiện thực hóa khát vọng hùng cường của đất nước vào năm 2045./.