

Bảo vệ nền tảng tư tưởng của Đảng trước thách thức trí tuệ nhân tạo (Kỳ 1: Cơ sở lý luận và nhận diện thách thức từ trí tuệ nhân tạo trên không gian mạng)

Ngày nhận bài: 27/04/2026 | Ngày phản biện: 08/05/2026 | Ngày xuất bản: 01/07/2026

Tác giả: Nguyễn Thị Ngọc Hoa (Học viện viện Báo chí và Tuyên truyền)

DOI: 10.66057/llcttt-29750

Tóm tắt: Sự bùng nổ của trí tuệ nhân tạo (AI) trên không gian mạng đang tạo ra một "chiến trường tư tưởng" mới, nơi các thế lực thù địch khai thác deepfake, mạng lưới bot, hệ thống tạo nội dung tổng hợp và thuật toán phân cực để tấn công nền tảng tư tưởng của Đảng Cộng sản Việt Nam với quy mô, tốc độ và mức độ tinh vi chưa từng có. Bài viết, đăng thành ba kỳ liên tiếp trên Tạp chí Lý luận chính trị và Truyền thông, nghiên cứu vấn đề này một cách hệ thống. Kỳ 1 dựa trên phương pháp luận duy vật biện chứng và duy vật lịch sử để phân tích cơ sở lý luận của chủ nghĩa Mác - Lênin, tư tưởng Hồ Chí Minh và chủ trương của Đảng; phác họa thực trạng không gian mạng Việt Nam; nhận diện và phân tích năm nhóm thủ đoạn tấn công tư tưởng mới sử dụng AI, cùng những hạn chế trong công tác ứng phó hiện nay trên ba phương diện: tính chủ động, tính kịp thời và tính sắc bén.

1. Đặt vấn đề

Lịch sử đấu tranh tư tưởng của Đảng Cộng sản Việt Nam là lịch sử của những thử thách liên tiếp và không ngừng đổi mới, từ chống thực dân phong kiến thời kỳ tiền khởi nghĩa, đến chống chiến tranh tâm lý của đế quốc trong hai cuộc kháng chiến, từ đối phó "diễn biến hòa bình" sau năm 1975 đến ứng phó làn sóng thông tin giả trên mạng xã hội đầu thế kỷ XXI. Mỗi thời kỳ đặt ra những thách thức riêng và Đảng ta luôn chứng tỏ khả năng thích ứng, sáng tạo, kiên định về nguyên tắc trong khi linh hoạt về phương thức.

Tuy nhiên, thách thức mà trí tuệ nhân tạo (AI) đặt ra trong cuộc đấu tranh tư tưởng hiện nay mang tính bước ngoặt, không phải sự thay đổi về mức độ của những thách thức quen thuộc, mà là sự thay đổi về bản chất. Trước đây, muốn tổ chức một chiến dịch tuyên truyền quy mô lớn, kẻ địch cần bộ máy người đông đảo, ngân sách khổng lồ, thời gian chuẩn bị dài và dễ bị phát hiện. Ngày nay, chỉ cần một vài mô hình AI ngôn ngữ lớn (Large Language Models - LLM) và một khoản đầu tư nhỏ, bất kỳ thế lực nào cũng có thể sản xuất hàng nghìn bài viết giả học thuật, hàng trăm video deepfake về lãnh đạo cấp cao, và triển khai hàng nghìn bot để lan truyền nội dung đó trên mạng xã hội, tất cả trong vài giờ, từ bất kỳ đâu trên thế giới.

Nghị quyết số 35-NQ/TW ngày 22/10/2018 của Bộ Chính trị xác định bảo vệ vững chắc nền tảng tư tưởng của Đảng, đấu tranh phản bác có hiệu quả các quan điểm sai trái, thù địch là “một nội dung cơ bản, hệ trọng, sống còn của công tác xây dựng, chỉnh đốn Đảng” (Đảng Cộng sản Việt Nam, 2018). Gần tám năm sau, ngày 08/4/2026, Ban Chấp hành Trung ương Đảng ban hành Quy định số 19-QĐ/TW về công tác chính trị, tư tưởng trong Đảng, lần đầu tiên Đảng ta có một quy định chuyên đề mang tính hệ thống, toàn diện về công tác này, đặt ra yêu cầu chủ động nắm bắt tình hình tư tưởng, tích cực bảo vệ nền tảng tư tưởng, kiên quyết đấu tranh với các quan điểm sai trái, thù địch, không để bị động, bất ngờ (Đảng Cộng sản Việt Nam, 2026a). Yêu cầu “không để bị động, bất ngờ” trong bối cảnh AI phát triển vũ bão vừa là thách thức, vừa là mệnh lệnh hành động cấp bách đối với toàn hệ thống chính trị.

2. Phương pháp nghiên cứu

Bài viết được trình bày thành ba kỳ đăng tải liên tiếp trên Tạp chí Lý luận chính trị và Truyền thông, tiếp cận vấn đề trên cơ sở phương pháp luận của chủ nghĩa duy vật biện chứng và chủ nghĩa duy vật lịch sử. Việc kiên định phương pháp luận mác-xít có ý nghĩa quyết định: nó bảo đảm nhìn nhận hiện tượng tấn công tư tưởng bằng AI không sa vào thuần túy kỹ thuật hay hiện tượng luận, mà luôn đặt trong mối quan hệ biện chứng giữa tồn tại xã hội và ý thức xã hội, giữa hạ tầng số và kiến trúc thượng tầng tư tưởng, tránh nguy cơ chệch hướng khi đánh giá bản chất chính trị, giai cấp của vấn đề.

Trên nền tảng đó, bài viết sử dụng các phương pháp cụ thể: phân tích - tổng hợp tài liệu đối với văn kiện của Đảng, tác phẩm kinh điển và báo cáo khoa học trong, ngoài nước; nghiên cứu trường hợp các vụ việc tấn công tư tưởng điển hình; phân tích hệ thống để nhận diện cơ chế vận hành của từng nhóm thủ đoạn; và tiếp cận liên ngành kết hợp chính trị học - truyền thông học - khoa học máy tính - tâm lý học xã hội. Phạm vi nghiên cứu tập trung vào không gian mạng Việt Nam từ năm 2023 đến nay, thời điểm AI tạo sinh bùng nổ và bắt đầu được sử dụng như một công cụ tấn công tư tưởng có tổ chức.

3. Kết quả và thảo luận

3.1. Kết quả nghiên cứu

3.1.1. Cơ sở lý luận về bảo vệ nền tảng tư tưởng trong kỷ nguyên trí tuệ nhân tạo

Thứ nhất, quan điểm triết học Mác - Lênin về ý thức xã hội, đấu tranh tư tưởng và lý luận khoa học

Triết học Mác - Lênin khẳng định mối quan hệ biện chứng giữa tồn tại xã hội và ý thức xã hội: tồn tại xã hội quyết định ý thức xã hội, nhưng ý thức xã hội có tính độc lập tương đối và có khả năng tác động ngược trở lại mạnh mẽ đối với tồn tại xã hội. Tính độc lập tương đối và khả năng tác động ngược đó chính là lý do tại sao công tác tư tưởng có tầm quan trọng chiến lược trong mọi cuộc cách mạng xã hội.

Trong tác phẩm "Hệ tư tưởng Đức" (1845-1846), C.Mác và Ph.Ăngghen khái quát: "Những tư tưởng của giai cấp thống trị, trong mọi thời đại, đều là những tư tưởng thống trị" (Mác & Ăngghen, 1995, t.3, tr.66). Trong bất kỳ xã hội có giai cấp nào, giai cấp nắm quyền lực kinh tế - chính trị cũng nỗ lực biến hệ tư tưởng của mình thành hệ tư tưởng thống trị toàn xã hội. Cuộc đấu tranh tư tưởng, do đó, không bao giờ trung lập, nó là biểu hiện tất yếu của đấu tranh giai cấp. Ai chiến thắng trên trận địa tư tưởng sẽ có lợi thế quyết định trong cuộc đấu tranh chính trị, xã hội rộng lớn hơn.

Đặc biệt quan trọng là luận điểm của C.Mác trong tác phẩm "Góp phần phê phán triết học pháp quyền của Hêghen. Lời nói đầu" (1844): "Lý luận cũng sẽ trở thành lực lượng vật chất, một khi nó thâm nhập vào quần chúng" (Mác & Ăngghen, 1995, t.1,

tr.580). Cần nhận thức thật chính xác: lý luận chỉ trở thành lực lượng vật chất khi hội đủ điều kiện tiên quyết là “thâm nhập vào quần chúng”, sức mạnh cải tạo hiện thực của nó tỷ lệ thuận với mức độ được quần chúng tiếp nhận, thấu hiểu và biến thành niềm tin, thành hành động. Hệ quả là cuộc đấu tranh tư tưởng, xét đến cùng, là cuộc đấu tranh giành quyền “thâm nhập vào quần chúng” giữa lý luận khoa học, cách mạng và các luận điệu sai trái, thù địch. Trong kỷ nguyên AI, chính điều kiện “thâm nhập” đó bị biến đổi sâu sắc: công nghệ cho phép một hệ thống quan điểm dù đúng hay sai thâm nhập vào hàng triệu người chỉ trong vài giờ, với mức độ cá nhân hóa chưa từng có, khiến việc giành và giữ trận địa ấy trở nên cấp bách hơn bao giờ hết.

V.I.Lênin phát triển học thuyết mác-xít về đấu tranh tư tưởng sắc bén hơn bao giờ hết. Trong tác phẩm “Làm gì?” (1902), Lênin chỉ rõ: “Vấn đề chỉ có thể đặt ra như thế này: hệ tư tưởng tư sản hoặc hệ tư tưởng xã hội chủ nghĩa. Không có hệ tư tưởng trung gian” (Lênin, 2005, t.6, tr.49-50). Đây là luận điểm chiến lược cốt lõi: nơi nào hệ tư tưởng mác-xít không chiếm lĩnh thì hệ tư tưởng tư sản sẽ chiếm lĩnh. Trong không gian mạng kỷ nguyên AI, luận điểm đó càng sắc bén: nếu nội dung bảo vệ tư tưởng của Đảng không đủ hấp dẫn, phổ biến và nhanh để tiếp cận người dùng, thì các thuật toán AI sẽ lập tức lấp đầy “khoảng trống” đó bằng nội dung thù địch.

Một chỉ dẫn của Lênin đặc biệt giàu tính thời sự là phương pháp nhận diện kẻ thù tư tưởng nấp dưới những lớp vỏ bọc “khoa học”, “đổi mới”, “dân chủ”. Trong tác phẩm “Làm gì?”, Lênin đã vạch trần bản chất của “phái kinh tế”, những người nhân danh “tính tự phát” của phong trào công nhân để hạ thấp vai trò của lý luận cách mạng. Trong đấu tranh chống chủ nghĩa xét lại và cơ hội, Người nhiều lần cảnh báo: kẻ thù tư tưởng nguy hiểm nhất thường không đối đầu trực diện, mà ngụy trang dưới hình thức “bổ sung”, “phát triển”, “hiện đại hóa” chủ nghĩa Mác để rút bỏ linh hồn cách mạng của nó. Bài học ấy có giá trị trực tiếp trong kỷ nguyên AI: các luận điệu phi mác-xít, cơ hội, xét lại nay được AI tạo sinh “đóng gói” tinh vi dưới vỏ bọc những bài nghiên cứu trông khách quan, những video “phản biện ôn hòa”, những phân tích “đa chiều”, khiến việc nhận diện khó khăn bội phần. Nắm vững phương pháp “bóc vỏ” của các nhà kinh điển, đi từ hiện tượng đến bản chất giai cấp, vì vậy vẫn là kim chỉ nam để nhận diện kẻ thù tư tưởng trong hình hài công nghệ mới.

Lênin cũng chỉ rõ tầm quan trọng của tính linh hoạt chiến thuật, khả năng thích ứng với điều kiện lịch sử mới là yếu tố sống còn của mỗi đảng cộng sản, trong khi kiên định về nguyên tắc chiến lược (Lênin, 2005, t.41, tr.27). Trong điều kiện AI, sự kết hợp giữa kiên định nguyên tắc và linh hoạt phương thức ấy được cụ thể hóa thành yêu cầu: giữ vững bản chất khoa học, cách mạng của nền tảng tư tưởng, đồng thời không ngừng đổi mới công cụ và phương thức bảo vệ cho tương thích với môi trường số.

Thứ hai, tư tưởng Hồ Chí Minh về báo chí, tuyên truyền và chiến lược bảo vệ trận địa tư tưởng

Tư duy của Chủ tịch Hồ Chí Minh về trận địa tư tưởng, báo chí, tuyên truyền là sự vận dụng sáng tạo chủ nghĩa Mác - Lênin vào điều kiện cụ thể của Việt Nam, đồng thời là những chỉ dẫn có giá trị vượt thời đại cho công tác bảo vệ tư tưởng ngày nay. Người coi trận địa tư tưởng có vị trí chiến lược quyết định: muốn chiến thắng về chính trị, quân sự, kinh tế, trước hết phải chiến thắng về tư tưởng.

Chủ tịch Hồ Chí Minh khẳng định: "Vì lợi ích mười năm thì phải trồng cây, vì lợi ích trăm năm thì phải trồng người" (Hồ Chí Minh, 2011, t.11, tr.528). Áp dụng vào trận địa tư tưởng, điều đó có nghĩa: bảo vệ tư tưởng không phải công việc ngắn hạn, đối phó từng sự việc, mà phải là chiến lược lâu dài, bắt đầu từ nền tảng giáo dục - đào tạo. Trong kỷ nguyên AI, điều đó đòi hỏi phải đào tạo ngay từ bây giờ những "chiến sĩ tư tưởng số" cho 5 đến 10 năm tới.

Về vị thế, trách nhiệm của lực lượng làm công tác tư tưởng, báo chí, Người khẳng định: "Cán bộ báo chí cũng là chiến sĩ cách mạng. Cây bút, trang giấy là vũ khí sắc bén của họ" (Hồ Chí Minh, 2011, t.13, tr.466). Luận điểm "vũ khí sắc bén" có ý nghĩa trực tiếp trong thời đại AI: "cây bút, trang giấy" hôm nay đã mở rộng thành các nền tảng số, công cụ phân tích dữ liệu và sản phẩm truyền thông đa phương tiện. Lực lượng bảo vệ tư tưởng muốn làm tròn nhiệm vụ "chiến sĩ" phải làm chủ những "vũ khí" mới ấy, vừa vững vàng bản lĩnh chính trị, đạo đức cách mạng, vừa thành thạo kỹ năng số để nội dung chính thống đủ sức cạnh tranh, hấp dẫn, lan tỏa trên thị trường thông tin số.

Người cũng nêu cao nguyên tắc nói đi đôi với làm, nêu gương trước quần chúng. Áp dụng vào cuộc đấu tranh tư tưởng trên không gian mạng, nguyên tắc đó có nghĩa: lực

lượng bảo vệ tư tưởng không thể chỉ hô hào quyết tâm chống thông tin giả, mà phải thực sự có mặt, có tiếng nói mạnh mẽ, thuyết phục trên chính những nền tảng mà thế lực thù địch đang khai thác, phải “chiếm lĩnh” không gian số bằng hành động thực tế. Bản thân Chủ tịch Hồ Chí Minh là tấm gương đấu tranh tư tưởng sắc bén: bằng ngòi bút châm biếm thâm thúy mà giàu sức thuyết phục, Người đã vạch rõ âm mưu, thủ đoạn của kẻ thù, biến báo chí thành vũ khí “phò chính, trừ tà”. Tinh thần chủ động tiến công, vạch trần bản chất luận điệu thù địch ấy chính là điều công tác bảo vệ nền tảng tư tưởng trong kỷ nguyên AI cần kế thừa và phát triển.

Thứ ba, chủ trương của Đảng về đổi mới phương thức bảo vệ tư tưởng trong kỷ nguyên số

Nhận thức sâu sắc tầm quan trọng của không gian mạng trong đấu tranh tư tưởng, Đảng ta đã và đang có những chủ trương đúng đắn, kịp thời và ngày càng cụ thể. Văn kiện Đại hội XIV của Đảng (2026) nhấn mạnh yêu cầu: “Đổi mới mạnh mẽ phương thức, nội dung công tác tư tưởng theo hướng chủ động, sáng tạo, ứng dụng công nghệ số; xây dựng lớp người Việt Nam có trí tuệ, bản lĩnh, tinh thần yêu nước, đủ sức đứng vững trước mọi tấn công tư tưởng” (Đảng Cộng sản Việt Nam, 2026b, tr.78). Đây là sự kế thừa và phát triển tư duy của Đảng, xác định rõ “ứng dụng công nghệ số” là yêu cầu bắt buộc, không thể né tránh.

Nghị quyết số 57-NQ/TW ngày 22/12/2024 của Bộ Chính trị về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia (Đảng Cộng sản Việt Nam, 2024) xác định AI là một trong những lĩnh vực ưu tiên hàng đầu, đồng thời nhấn mạnh yêu cầu quản lý, sử dụng AI có trách nhiệm, ngăn chặn các rủi ro an ninh phi truyền thống, trong đó có rủi ro tư tưởng. Sự kết hợp giữa Nghị quyết 57-NQ/TW và Quy định 19-QĐ/TW tạo khuôn khổ chính sách tổng thể: vừa phát triển AI như động lực tăng trưởng, vừa bảo vệ an ninh tư tưởng trong không gian AI.

Cuộc thi chính luận về bảo vệ nền tảng tư tưởng của Đảng lần thứ Sáu, năm 2026 với chủ đề “Bảo vệ nền tảng tư tưởng của Đảng trong kỷ nguyên phát triển mới” (Học viện Chính trị quốc gia Hồ Chí Minh, 2026) là một biểu hiện rõ nét của tinh thần chủ động, dũng cảm đặt thẳng vào trọng tâm thách thức thay vì né tránh: thay vì chờ đến khi các vụ tấn công tư tưởng bằng AI gây ra những hậu quả không thể đảo ngược, Đảng ta

chủ động mở diễn đàn nghiên cứu, tranh luận, tìm giải pháp ngay từ khi thách thức đang hình thành.

3.1.2. Thực trạng không gian mạng Việt Nam và các thủ đoạn tấn công tư tưởng bằng AI

3.1.2.1. Bức tranh không gian mạng Việt Nam - “chiến địa” tư tưởng thế hệ mới

Để nhận diện chính xác thách thức, trước hết cần phác thảo bức tranh không gian mạng Việt Nam, nơi đang diễn ra cuộc đấu tranh tư tưởng thế hệ mới. Theo báo cáo Digital 2025 của We Are Social & Meltwater (2025, tr.18-22), đầu năm 2025 Việt Nam có khoảng 79,8 triệu người dùng Internet (tỷ lệ thâm nhập 78,8%) và 76,2 triệu người dùng mạng xã hội, thuộc nhóm các quốc gia có số người dùng mạng xã hội cao hàng đầu Đông Nam Á. Trung bình mỗi người dùng dành hơn 6 giờ mỗi ngày trên mạng, trong đó khoảng 2 giờ 30 phút cho mạng xã hội. Các nền tảng phổ biến nhất gồm Facebook, YouTube, TikTok, Zalo và Instagram.

Điều đặc biệt quan trọng là cơ cấu nhân khẩu học người dùng: nhóm tuổi 18-34 chiếm tỷ lệ cao nhất trong số người dùng mạng xã hội tích cực, cũng là nhóm đang hình thành thế giới quan, giá trị quan và bản lĩnh chính trị. Theo Sách trắng Công nghệ thông tin và Truyền thông Việt Nam 2024 (Bộ Thông tin và Truyền thông, 2024, tr.45-47), mỗi ngày có hàng chục triệu lượt xem video, hàng trăm triệu lượt tương tác trên các nền tảng. Không gian mạng không còn là không gian ảo, nó đã trở thành không gian sinh tồn, nơi ý thức chính trị, niềm tin vào Đảng và Nhà nước, tâm lý xã hội và nhiều quyết định quan trọng của hàng chục triệu người được hình thành và bị tác động hằng ngày, hằng giờ.

Sự kết hợp giữa mức độ thâm nhập Internet cao và quá trình chuyển đổi số quốc gia đang được đẩy nhanh theo Nghị quyết 57-NQ/TW (Đảng Cộng sản Việt Nam, 2024) tạo ra một nghịch lý: trong khi cơ hội phát triển kinh tế số, xã hội số lớn lên, “bề mặt tấn công” tư tưởng cũng mở rộng theo tỷ lệ tương đương, thậm chí còn nhanh hơn. Đây là mâu thuẫn cốt lõi mà chiến lược bảo vệ nền tảng tư tưởng phải đối mặt trực tiếp.

3.1.2.2. Năm nhóm thủ đoạn tấn công tư tưởng mới sử dụng AI - phân tích chi tiết

Trên cơ sở tổng hợp tài liệu và phân tích thực tiễn, bài viết nhận diện năm nhóm thủ đoạn chính mà các thế lực thù địch đang sử dụng AI để tấn công nền tảng tư tưởng của Đảng trên không gian mạng Việt Nam hiện nay:

Thứ nhất, công nghệ deepfake nhắm vào lãnh đạo cấp cao và sự kiện chính trị. Deepfake (ghép từ “deep learning” - học sâu và “fake” - giả mạo) là công nghệ dùng trí tuệ nhân tạo để tạo ra hình ảnh, video, giọng nói giả mạo nhưng trông như thật, chủ yếu dựa trên mạng nơ-ron đối kháng tạo sinh (GAN) và các mô hình khuếch tán. Công nghệ này cho phép gắn khuôn mặt, giọng nói của một người thật vào những phát ngôn, hành động mà người đó chưa từng thực hiện, ở mức độ mắt thường khó nhận ra. Theo Europol (2024, tr.12-18), deepfake nhắm vào lãnh đạo chính trị là một trong những mối đe dọa an ninh hàng đầu hiện nay; công nghệ này lại ngày càng rẻ và dễ sử dụng, từ chỗ đòi hỏi chuyên gia với thiết bị đắt tiền, nay chỉ cần vài giờ và một máy tính thông thường. Ở Việt Nam, đã xuất hiện các video deepfake giả mạo hình ảnh, giọng nói của lãnh đạo Đảng và Nhà nước nhằm tung tin giả về các quyết sách quan trọng, gây hoang mang dư luận trước các sự kiện chính trị trọng đại.

Thứ hai, mạng lưới bot và tài khoản AI tự động hóa chiến dịch tấn công tư tưởng quy mô lớn. Bot là các tài khoản tự động do phần mềm điều khiển; hàng nghìn bot được điều phối tập trung tạo thành một “mạng lưới bot”. Các thế lực thù địch dùng hàng nghìn đến hàng chục nghìn tài khoản bot do AI điều phối để đồng loạt tung tin giả, “thích”, “chia sẻ”, “bình luận” nhằm tạo cảm giác nội dung được cộng đồng đón nhận rộng rãi, đồng thời tấn công các tài khoản chính thống và chi phối xu hướng trên mạng xã hội. Kỹ thuật này, còn gọi là “astroturfing”, tức giả tạo một phong trào quần chúng “tự phát” từ bên dưới, đặc biệt nguy hiểm vì đánh trực tiếp vào tâm lý đám đông và hiệu ứng băng chứng xã hội. Theo báo cáo minh bạch của Meta, nhiều chiến dịch phối hợp có sử dụng AI nhắm vào các quốc gia Đông Nam Á đã bị phát hiện và xử lý (Meta Platforms Inc., 2024).

Thứ ba, AI tạo sinh nội dung sản xuất “thông tin giả chất lượng cao” với tốc độ công nghiệp. Các mô hình ngôn ngữ lớn (LLM) và mô hình tạo sinh nguồn mở cho phép sản xuất nội dung sai lệch với tốc độ và quy mô khó tưởng tượng. Theo báo cáo AI Index 2025 của Đại học Stanford (Stanford University, 2025, tr.23-27), sản lượng nội dung

do AI tạo ra tăng mạnh trong giai đoạn 2023-2024. Điều làm thủ đoạn này đặc biệt nguy hiểm là khả năng tạo “thông tin giả học thuật”, những bài viết có bề ngoài nghiêm túc, đầy đủ cấu trúc luận điểm, số liệu (giả nhưng trông thuyết phục), chú thích và phong cách hàn lâm, dễ đánh lừa cả người đọc không chuyên lẫn một số nhà nghiên cứu. Chúng được dịch tự động sang nhiều ngôn ngữ rồi phân phối toàn cầu trong vài phút, nhằm bác bỏ tính chính đáng của Đảng, xuyên tạc lịch sử dân tộc, phủ nhận thành tựu phát triển và bóp méo chủ nghĩa Mác - Lênin.

Thứ tư, thuật toán gợi ý tạo “buồng vọng” (echo chamber) - phân cực tư tưởng có hệ thống. Đây là thủ đoạn tinh vi nhất vì nó không đến từ bên ngoài mà được tạo ra từ bên trong chính hệ thống người dùng đang dùng hằng ngày. Thuật toán AI của các nền tảng mạng xã hội được thiết kế để tối đa hóa “thời gian sử dụng” và cách hiệu quả nhất là hiển thị nội dung gây cảm xúc mạnh, nhất là phẫn nộ, lo lắng và bất bình. Kết quả là người dùng dần bị “nhốt” trong một không gian thông tin ngày càng hẹp và cực đoan hơn, chỉ tiếp nhận những quan điểm xác nhận và tăng cường niềm tin tiêu cực về Đảng và Nhà nước, không tiếp xúc với thông tin phản biện hay chính thống. Theo Freedom House (2024, tr.15-19), tác động của thuật toán phân cực đang ngày càng bào mòn lòng tin vào các thể chế chính trị tại nhiều quốc gia, trong đó có các quốc gia Đông Nam Á.

Thứ năm, nhắm mục tiêu vi mô - AI cá nhân hóa chiến dịch tuyên truyền nhắm đúng đối tượng dễ bị tổn thương nhất. Đây là hình thức tấn công tư tưởng tinh vi và khó đối phó nhất từ trước đến nay. Bằng cách thu thập, phân tích dữ liệu người dùng (độ tuổi, nghề nghiệp, thu nhập ước tính, lịch sử tìm kiếm, lịch sử xem, mối quan hệ xã hội, trạng thái cảm xúc suy luận từ ngôn ngữ bài đăng), AI dựng “hồ sơ tâm lý” chi tiết của từng cá nhân. Trên cơ sở đó, kẻ tấn công gửi “thông điệp phù hợp” đến đúng người vào đúng thời điểm: người lo lắng về kinh tế gia đình nhận thông điệp về “sự thất bại” của chính sách kinh tế; thanh niên tìm kiếm định hướng nghề nghiệp nhận nội dung về “sự bất công” của hệ thống; trí thức đang phản biện một chính sách nhận “phân tích học thuật” đẩy họ từ phản biện xây dựng sang đối lập toàn diện. Theo Microsoft - Digital Crimes Unit (2024, tr.8-11), các chiến dịch tinh vi nhất hiện nay không phát tán thông tin giả đồng loạt mà thực hiện “phẫu thuật tư tưởng” nhắm vào từng cá nhân.

3.1.3. Đánh giá hạn chế trong công tác ứng phó hiện nay trên ba phương diện: tính chủ động, tính kịp thời và tính sắc bén

Mặc dù công tác bảo vệ nền tảng tư tưởng đã đạt nhiều kết quả tích cực sau hơn bảy năm triển khai Nghị quyết 35-NQ/TW, thực tiễn cho thấy còn ba hạn chế cốt lõi cần nhìn thẳng vào để khắc phục, thể hiện trên ba phương diện: tính chủ động, tính kịp thời và tính sắc bén.

Hạn chế thứ nhất, về tính chủ động: năng lực phát hiện và phản bác nội dung do AI tạo ra của các lực lượng còn yếu và bị động. Phần lớn thông tin giả chỉ được phát hiện qua phản ánh, báo cáo của người dùng mạng xã hội, thay vì được nhận diện chủ động bởi hệ thống giám sát có công nghệ. Theo Báo cáo tổng kết công tác tuyên giáo năm 2024, nhiều vụ thông tin giả chỉ được xử lý sau khi đã lan truyền đến hàng trăm nghìn người dùng, tức là khi “vết thương” tư tưởng đã hình thành và ăn sâu vào nhận thức của một bộ phận người dân (Ban Tuyên giáo và Dân vận Trung ương, 2025, tr.34). Đây là hạn chế nguy hiểm nhất, bởi trong môi trường mạng xã hội, thông tin giả lan truyền theo hàm số mũ; thế bị động ngay từ khâu phát hiện đồng nghĩa với việc luôn đi sau, chạy theo và khắc phục hậu quả thay vì chủ động ngăn chặn từ sớm, từ xa.

Hạn chế thứ hai, về tính kịp thời: tốc độ phản ứng của thông tin chính thống còn quá chậm so với tốc độ lan truyền của thông tin giả. Nghiên cứu của Trần Văn Phòng (2023) và nhiều nhà nghiên cứu quốc tế đều xác nhận: trên mạng xã hội, thông tin giả lan truyền nhanh gấp nhiều lần thông tin thật, đặc biệt là tin giả về chính trị gây cảm xúc mạnh (phẫn nộ, bất bình, nghi ngờ). Trong khi đó, quy trình phê duyệt và phản hồi hiện nay thường mất nhiều ngày, thậm chí nhiều tuần. “Cửa sổ vàng” để phản bác hiệu quả, thường chỉ vài giờ đầu sau khi tin giả xuất hiện, vì thế bị bỏ lỡ, khiến thông tin chính thống dù đúng vẫn rơi vào thế “đánh chính muộn”, hiệu quả định hướng dư luận suy giảm đáng kể.

Hạn chế thứ ba, về tính sắc bén: thiếu đội ngũ cán bộ kết hợp được năng lực lý luận chính trị vững chắc với năng lực kỹ thuật số, đồng thời sản phẩm phản bác nhiều khi chưa đủ sắc bén, thuyết phục. Như Phạm Thị Thịnh (2024, tr.32) và Nguyễn Thị Ngọc Hoa (2024, tr.18) đã nhận định: hầu hết cán bộ hiện nay được đào tạo theo mô hình truyền thống, giỏi lý luận nhưng chưa đủ năng lực nhận diện deepfake, phân tích

mạng lưới bot hay vận hành hệ thống kiểm chứng thông tin dựa trên AI. Hệ quả là không ít sản phẩm phản bác còn nặng tính phòng thủ, lập luận chung chung, thiếu bằng chứng cụ thể và sức nặng học thuật để "đánh trúng" bản chất luận điệu thù địch. Tính sắc bén vì vậy phải được nâng cao đồng thời ở cả hai phương diện: sắc bén trong lập luận, bằng chứng và sắc bén trong hình thức thể hiện, công nghệ chuyển tải. Khoảng cách giữa năng lực hiện có và yêu cầu thực tiễn là điểm yếu cần khắc phục khẩn cấp.

3.2. Thảo luận

Các kết quả trên cho thấy một vấn đề cốt lõi cần lý giải: vì sao AI tạo ra bước ngoặt về chất chứ không chỉ về lượng trong đấu tranh tư tưởng? Lời giải nằm ở chính luận điểm của C.Mác về điều kiện để lý luận "thâm nhập vào quần chúng" (Mác & Ăngghen, 1995, t.1, tr.580). Trong suốt lịch sử, khâu khó khăn và quyết định nhất để một hệ tư tưởng trở thành "lực lượng vật chất" luôn là quá trình thâm nhập ấy, vốn đòi hỏi tổ chức, thời gian, con người và nguồn lực lớn. Đặc trưng của AI là xóa bỏ gần như toàn bộ rào cản đó: một hệ thống quan điểm, dù đúng hay sai, có thể tiếp cận hàng triệu người trong vài giờ và được cá nhân hóa tới từng người. Như vậy, AI không phủ định luận điểm kinh điển, trái lại làm cho nó thời sự và cấp bách hơn bao giờ hết. Đây cũng là cơ sở để khẳng định phương pháp luận mác-xít không hề lỗi thời trước công nghệ mới, bởi chính nó cho phép nhìn xuyên lớp vỏ kỹ thuật để nắm bắt bản chất chính trị, giai cấp của hiện tượng.

So với các nghiên cứu đi trước, kết quả bài viết vừa kế thừa vừa có điểm phát triển. Các báo cáo quốc tế (Stanford University, 2025; Europol, 2024; Freedom House, 2024) giàu bằng chứng thực nghiệm nhưng chủ yếu tiếp cận dưới góc độ "thông tin sai lệch" hay "xói mòn lòng tin" nói chung, chưa đặt vấn đề trong bản chất chính trị, giai cấp. Ở trong nước, các công trình như Trần Văn Phòng (2023), Phạm Thị Thịnh (2024) và Nguyễn Thị Ngọc Hoa (2024) đã bàn về bảo vệ nền tảng tư tưởng trong bối cảnh chuyển đổi số, song phần lớn xem chuyển đổi số như bối cảnh chung, chưa tách AI thành một bước ngoặt riêng về chất và thường nghiêng về một phía, hoặc lý luận hoặc kỹ thuật. Đóng góp mới của bài viết là: (i) định vị AI như sự thay đổi về bản chất chứ không chỉ về mức độ; (ii) hệ thống hóa năm nhóm thủ đoạn cùng cơ chế vận hành; (iii) đề xuất khung đánh giá ba chiều về tính chủ động, kịp thời và sắc bén, vừa

giúp chẩn đoán đúng "căn bệnh", vừa trực tiếp gợi mở định hướng giải pháp.

Về lý luận, kết quả nghiên cứu củng cố một nhận định có ý nghĩa phương pháp luận: giá trị của chủ nghĩa Mác - Lênin và tư tưởng Hồ Chí Minh không suy giảm mà càng được khẳng định trong kỷ nguyên số, khi luận điểm "Không có hệ tư tưởng trung gian" (Lênin, 2005, t.6, tr.49-50) giải thích chính xác vì sao mọi khoảng trống của thông tin chính thống lập tức bị thuật toán lấp đầy bằng nội dung thù địch; đồng thời quan hệ biện chứng giữa tồn tại xã hội và ý thức xã hội nay có thêm một tầng trung gian mới là thuật toán, khiến việc làm chủ tầng trung gian ấy trở thành địa hạt trọng yếu của đấu tranh tư tưởng. Cách tiếp cận này đòi hỏi tránh đồng thời hai khuynh hướng: tuyệt đối hóa kỹ thuật mà quên bản chất chính trị, và xem nhẹ đặc thù công nghệ nên phản bác chung chung, thiếu sắc bén. Về thực tiễn, toàn bộ phân tích dẫn tới hệ quả tất yếu: công tác bảo vệ nền tảng tư tưởng phải chuyển từ phòng thủ, phản ứng sang chủ động, kiến tạo, từ nhịp độ con người sang mô hình "con người + AI". Tuy nhiên, nhận diện thủ đoạn (Kỳ 1) là cần nhưng chưa đủ; phải lý giải sâu cơ chế tâm lý, xã hội mà AI khai thác và tham khảo có chọn lọc kinh nghiệm quốc tế (Kỳ 2), rồi mới đề xuất hệ thống giải pháp tổng thể, có lộ trình (Kỳ 3). Theo nghĩa đó, kết quả Kỳ 1 vừa là chẩn đoán, vừa là tiền đề lý luận, thực tiễn cho toàn bộ thiết kế "lá chắn tư tưởng số" ở các kỳ tiếp theo.

4. Kết luận kỳ 1

Từ cơ sở lý luận của chủ nghĩa Mác - Lênin, tư tưởng Hồ Chí Minh và chủ trương của Đảng, có thể khẳng định: bảo vệ nền tảng tư tưởng là nhiệm vụ sinh tử của Đảng, không thể nhượng bộ, thỏa hiệp trong bất kỳ hoàn cảnh nào. Cuộc đấu tranh tư tưởng không bao giờ kết thúc, nó chỉ thay đổi hình thức, phương tiện và chiến trường. Trong kỷ nguyên AI, chiến trường đó là không gian số; phương tiện của kẻ địch là AI; và hình thức tấn công là sự kết hợp tinh vi giữa deepfake, mạng lưới bot, thông tin giả học thuật, buông vọng và nhắm mục tiêu vi mô.

Năm nhóm thủ đoạn được phân tích trong bài viết này không phải là dự báo tương lai, chúng đang diễn ra ngay hôm nay, trên điện thoại thông minh của hàng chục triệu người Việt Nam. Việc nhận diện chính xác, toàn diện bản chất, cơ chế và mức độ nguy hiểm của từng nhóm thủ đoạn, cùng việc đánh giá thẳng thắn những hạn chế hiện nay

về tính chủ động, tính kịp thời và tính sắc bén, chính là tiền đề lý luận và thực tiễn không thể thiếu để kiến tạo một “lá chắn tư tưởng số” đủ mạnh trong kỷ nguyên phát triển mới. Bảo vệ nền tảng tư tưởng của Đảng trên không gian mạng, suy cho cùng, là cuộc đấu tranh giữ vững vai trò chủ đạo của hệ tư tưởng khoa học, cách mạng trong đời sống tinh thần xã hội, một cuộc đấu tranh không khoan nhượng, lâu dài và ngày càng phức tạp, đòi hỏi sự vào cuộc đồng bộ của cả hệ thống chính trị và toàn xã hội./.

Tài liệu tham khảo

1. Ban Tuyên giáo và Dân vận Trung ương. (2025). Báo cáo tổng kết công tác tuyên giáo năm 2024 và phương hướng, nhiệm vụ năm 2025.
2. Bộ Thông tin và Truyền thông. (2024). Sách trắng Công nghệ thông tin và Truyền thông Việt Nam năm 2024.
3. Đảng Cộng sản Việt Nam. (2018). Nghị quyết số 35-NQ/TW ngày 22/10/2018 của Bộ Chính trị về tăng cường bảo vệ nền tảng tư tưởng của Đảng, đấu tranh phản bác các quan điểm sai trái, thù địch trong tình hình mới.
4. Đảng Cộng sản Việt Nam. (2024). Nghị quyết số 57-NQ/TW ngày 22/12/2024 của Bộ Chính trị về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia.
5. Đảng Cộng sản Việt Nam. (2026a). Quy định số 19-QĐ/TW ngày 08/4/2026 của Ban Chấp hành Trung ương Đảng về công tác chính trị, tư tưởng trong Đảng.
6. Đảng Cộng sản Việt Nam. (2026b). Văn kiện Đại hội đại biểu toàn quốc lần thứ XIV. Nhà xuất bản Chính trị quốc gia Sự thật.
7. Europol. (2024). Internet Organised Crime Threat Assessment (IOCTA) 2024: Artificial Intelligence in Crime. European Union Agency for Law Enforcement Cooperation.
8. Freedom House. (2024). Freedom on the Net 2024 - The Struggle for Trust Online.
9. Học viện Chính trị quốc gia Hồ Chí Minh. (2026). Kế hoạch số 99-KH/HVCTQG ngày 03/02/2026 tổ chức Cuộc thi chính luận về bảo vệ nền tảng tư tưởng của Đảng lần thứ Sáu, năm 2026.
10. Hồ Chí Minh. (2011). Toàn tập (t.11, t.13). Nhà xuất bản Chính trị quốc gia Sự thật.
11. Lênin, V.I. (2005). Toàn tập (t.6, t.41). Nhà xuất bản Chính trị quốc gia Sự thật.
12. Mác, C. & Ăngghen, Ph. (1995). Toàn tập (t.1, t.3). Nhà xuất bản Chính trị quốc gia Sự thật.
13. Meta Platforms Inc. (2024). Transparency Report - Adversarial Threat Report, Q3 2024.
14. Microsoft - Digital Crimes Unit. (2024). Cyber Signals Report, Issue 7: AI-Powered Influence Operations.
15. Nguyễn Thị Ngọc Hoa. (2024). Trí tuệ nhân tạo và thách thức đối với công tác tư tưởng của Đảng trong kỷ nguyên số. Tạp chí Lý luận chính trị và Truyền thông, 12, 18-25.
16. Phạm Thị Thịnh. (2024). Công tác bảo vệ nền tảng tư tưởng của Đảng trong bối cảnh chuyển đổi số: Thực trạng và giải pháp. Tạp chí Lý luận chính trị và Truyền thông, 9, 32-38.

17. Stanford University. (2025). Artificial Intelligence Index Report 2025. Stanford HAI.
18. Trần Văn Phòng. (2023). Về đấu tranh phản bác các quan điểm sai trái, thù địch trên không gian mạng trong tình hình mới. Tạp chí Cộng sản điện tử, 10.
19. We Are Social & Meltwater. (2025). Digital 2025 - Vietnam Country Report.