

TRÍ TUỆ NHÂN TẠO TRONG TRUYỀN THÔNG SỐ: TỪ CÁ NHÂN HÓA NỘI DUNG ĐẾN NGUY CƠ CỰC ĐOAN HÓA THÔNG TIN QUA HIỆU ỨNG BUỒNG VANG VÀ BONG BÓNG LỘC

Nguyễn Trần Thái Dương
Khoa Kinh tế, Học viện Hành chính và Quản trị công

Tóm tắt: Bài viết phân tích tác động của AI trong truyền thông số, đặc biệt là cơ chế cá nhân hóa nội dung và hệ quả như buồng vang, bong bóng lọc. Những hiện tượng này làm hạn chế tư duy phân biện và gia tăng xu hướng cực đoan hóa. Tác giả đề xuất giải pháp về công nghệ, giáo dục và pháp lý nhằm xây dựng môi trường truyền thông số minh bạch, bền vững.

Từ khóa: AI, truyền thông số, cá nhân hóa nội dung, buồng vang, bong bóng lọc, cực đoan hóa

ARTIFICIAL INTELLIGENCE IN DIGITAL MEDIA: FROM CONTENT PERSONALIZATION TO THE RISK OF INFORMATION EXTREMIZATION THROUGH ECHO CHAMBER AND FILTER BUBBLE EFFECTS

Nguyen Tran Thai Duong
Faculty of Economics, Academy of Administration and Public Governance

Abstract: This paper analyzes the impact of artificial intelligence (AI) on digital media, focusing on content personalization mechanisms and their consequences such as echo chambers and filter bubbles. These phenomena limit critical thinking and increase the tendency toward extremization. The author proposes technological, educational, and legal solutions to build a transparent and sustainable digital media environment.

Keywords: AI, digital media, content personalization, echo chamber, filter bubble, extremization

Nhận bài: 08/05/2025

Phản biện: 25/05/2025

Duyệt đăng: 30/05/2025

I. ĐẶT VẤN ĐỀ

Trong kỷ nguyên số, trí tuệ nhân tạo (AI) đã và đang tạo ra những bước đột phá trong nhiều lĩnh vực, đặc biệt là lĩnh vực truyền thông và báo chí. Các thuật toán AI không chỉ hỗ trợ tạo ra và phân phối nội dung mà còn giúp tối ưu hóa trải nghiệm người dùng thông qua việc cá nhân hóa thông tin. Các nền tảng như Facebook, Twitter, Google, YouTube sử dụng AI để phân tích dữ liệu người dùng và cung cấp các nội dung phù hợp với sở thích, nhu cầu và hành vi trước đó. Mặc dù sự cá nhân hóa này mang lại tiện lợi và cải thiện trải nghiệm người dùng, nhưng nó cũng tiềm ẩn nhiều rủi ro cho xã hội và hệ sinh thái thông tin.

Một trong những tác động đáng lo ngại của việc cá nhân hóa là sự hình thành của các hiệu ứng buồng vang (echo chamber) và bong bóng lọc (filter bubble). Cả hai hiện tượng này không chỉ làm thay đổi cách thức người dùng tiếp cận và tiêu thụ thông tin, mà còn có thể dẫn đến sự cực đoan hóa quan điểm. Khi thông tin nhận được ngày càng đồng nhất và củng cố các niềm tin hiện có, người dùng có thể trở nên cực đoan hơn trong cách nhìn nhận vấn đề. Điều này gây ra sự phân cực xã hội và làm cho các cuộc thảo luận, đối thoại hay việc tìm kiếm sự đồng thuận trở nên khó khăn hơn.

Trong bối cảnh này, việc nghiên cứu tác động của AI đối với truyền thông số, đặc biệt là nguy

cực đoan hóa thông tin do buồng vang và bong bóng lọc, là rất cần thiết. Điều này giúp chúng ta hiểu rõ hơn về vấn đề và đưa ra giải pháp nhằm xây dựng một môi trường truyền thông số lành mạnh trong kỷ nguyên trí tuệ nhân tạo.

II. NỘI DUNG NGHIÊN CỨU

2.1. Hiệu ứng buồng vang thông tin (Echo Chamber)

Hiệu ứng buồng vang thông tin (Echo chamber) là một hiện tượng truyền thông xã hội trong đó các cá nhân chủ yếu tiếp xúc với những thông tin, quan điểm và lập luận tương đồng với niềm tin sẵn có của họ, trong khi các ý kiến trái chiều bị loại bỏ hoặc ít xuất hiện. Trong môi trường như vậy, thông tin được lặp đi lặp lại và củng cố qua tương tác trong nhóm, dẫn đến sự khẳng định lẫn nhau và gia tăng niềm tin cực đoan. Theo Sunstein (2001): “Khi các cá nhân tương tác chủ yếu với những người có cùng quan điểm, họ có xu hướng trở nên cực đoan hóa trong chính quan điểm đó” – đây chính là hệ quả đặc trưng của buồng vang thông tin.

Trong môi trường truyền thông số, đặc biệt là trên mạng xã hội, hiệu ứng buồng vang thông tin càng trở nên phổ biến do sự kết hợp giữa cơ chế thuật toán cá nhân hóa nội dung và “thiên kiến xác nhận” - tâm lý lựa chọn thiên lệch (confirmation bias) của người dùng. Các nền tảng như Facebook,

YouTube hay TikTok sử dụng thuật toán đề xuất dựa trên hành vi quá khứ của người dùng, từ đó ưu tiên hiển thị những nội dung đồng thuận với sở thích hoặc niềm tin đã có.

2.2. Hiệu ứng bong bóng lọc (Filter Bubble)

Hiệu ứng bong bóng lọc (Filter bubble) được Eli Pariser giới thiệu vào năm 2011. Theo đó, nó được định nghĩa là một hiện tượng trong truyền thông số, mô tả tình trạng người dùng Internet – đặc biệt trên các nền tảng mạng xã hội và công cụ tìm kiếm – bị giới hạn trong một không gian thông tin cá nhân hóa cao đến mức chỉ còn tiếp cận các nội dung phù hợp với sở thích, niềm tin hoặc hành vi quá khứ, trong khi các thông tin trái chiều bị loại trừ gần như hoàn toàn. Eli Pariser cũng cảnh báo rằng: “Người dùng càng được cá nhân hóa trải nghiệm, họ càng ít nhìn thấy những gì có thể thách thức hoặc mở rộng thế giới quan của mình”.

Khác với hiệu ứng buồng vang – vốn thiên về yếu tố xã hội và sự lựa chọn có chủ đích của người dùng trong việc gắn kết với các nhóm cùng quan điểm – hiệu ứng bong bóng lọc chủ yếu do các thuật toán của nền tảng số tạo ra. Các hệ thống như Google, Facebook, TikTok hay YouTube sử dụng thuật toán học máy để phân tích hành vi người dùng (lịch sử tìm kiếm, lượt thích, thời gian xem...), từ đó cung cấp nội dung cá nhân hóa. Hệ quả là người dùng dần tiếp cận thông tin trong không gian khép kín, không phải do tìm kiếm chủ động mà là sản phẩm được điều chỉnh bởi các tiêu chí tối ưu hóa tương tác – thường vì mục tiêu thương mại.

2.3. Trí tuệ nhân tạo và cơ chế cá nhân hoá nội dung trong truyền thông số

Một trong những ứng dụng phổ biến và có ảnh hưởng sâu rộng nhất của AI trong lĩnh vực truyền thông số là cơ chế cá nhân hóa nội dung – tức là khả năng điều chỉnh thông tin mà người dùng nhìn thấy dựa trên hồ sơ hành vi, sở thích và tương tác cá nhân. Các hệ thống AI sử dụng các thuật toán học máy (machine learning), học sâu (deep learning) và xử lý ngôn ngữ tự nhiên (NLP) để phân tích dữ liệu lớn được thu thập từ hành vi của người dùng: từ lịch sử tìm kiếm, lượt nhấp chuột, thời lượng xem video, đến tương tác với nội dung trên mạng xã hội. Trên cơ sở đó, AI đưa ra các gợi ý nội dung được cho là “phù hợp” nhất với từng người dùng – nhằm tối đa hóa thời gian sử dụng, khả năng tương tác và lợi nhuận từ quảng cáo.

Một mặt, cơ chế cá nhân hóa này đã tạo ra một bước tiến đáng kể trong việc thay đổi hành vi tiếp

nhận thông tin và cải thiện trải nghiệm của công chúng. Đặc biệt, khi môi trường truyền thông ngày càng quá tải thông tin, việc lọc ra các nội dung phù hợp giúp người dùng tiết kiệm thời gian, đồng thời cảm thấy được “phục vụ theo nhu cầu cá nhân”.

Tuy nhiên, mặt trái của quá trình này chính là việc thuật toán dần tạo ra các “không gian thông tin khép kín”, nơi người dùng chỉ tiếp xúc với các nội dung đồng thuận với quan điểm, sở thích hoặc định kiến sẵn có của họ. Đây chính là điều kiện hình thành của hai hiện tượng đáng quan ngại: hiệu ứng buồng vang (echo chamber) và bong bóng lọc (filter bubble). Hệ quả của việc cá nhân hóa nội dung một cách cực đoan là sự suy giảm khả năng tiếp cận thông tin đa chiều, làm yếu đi tư duy phản biện và góp phần thúc đẩy phân cực xã hội. Khi thế giới quan của người dùng bị định hình và củng cố bởi các thuật toán, nguy cơ xuất hiện các xu hướng cực đoan hóa thông tin trở nên rõ ràng hơn bao giờ hết.

2.4. Hiệu ứng buồng vang và bong bóng lọc – từ cá nhân hoá nội dung đến cực đoan hoá thông tin

Khi AI can thiệp sâu vào quá trình thu thập, xử lý và phân phối thông tin, các thuật toán cá nhân hóa trở thành “người định hướng” vô hình, quyết định người dùng sẽ tiếp cận thông tin gì, theo cách nào và trong bao lâu. Dù mang lại nhiều lợi ích như tiết kiệm thời gian tìm kiếm, tăng mức độ tương tác và hiệu quả tiếp thị, cá nhân hóa nội dung cũng đặt ra nhiều thách thức, trong đó nổi bật là nguy cơ tạo ra các hiệu ứng: buồng vang thông tin (echo chamber) và bong bóng lọc (filter bubble).

Cơ chế vận hành của buồng vang và bong bóng lọc đều bắt nguồn từ việc thuật toán được xây dựng để “học” và “ghi nhớ” thói quen của người dùng thông qua các hành vi tương tác: từ nội dung đã xem, các bài viết đã thích... đến mạng lưới bạn bè hay chủ đề thường xuyên tìm kiếm... Trên cơ sở đó, các nền tảng truyền thông như Facebook, YouTube hay TikTok sẽ ưu tiên hiển thị những thông tin tương đồng với các khuôn mẫu hành vi đã ghi nhận. Mục tiêu của việc này không nằm ở giáo dục hay đa dạng hóa nhận thức, mà nhằm giữ chân người dùng lâu hơn – điều vốn mang lại giá trị thương mại to lớn cho các công ty công nghệ. Hệ quả là người dùng ngày càng ít tiếp cận với thông tin trái chiều, các quan điểm phản biện, hay những nội dung có thể khiến cho họ cảm thấy khó chịu.

Ở cấp độ cá nhân, điều này tạo cảm giác thoải mái và xác nhận niềm tin sẵn có – một dạng của

tâm lý lựa chọn thiên lệch (confirmation bias). Tuy vậy, khi được nhân rộng trên quy mô xã hội, những không gian thông tin khép kín này hình thành các cộng đồng có cùng tư tưởng, lặp đi lặp lại các thông tin và quan điểm tương đồng, tạo thành hiệu ứng buồng vang. Tương tự, bong bóng lọc hoạt động như một bộ lọc vô hình, mà ở đó các thông tin trái ngược bị thuật toán chủ động loại bỏ, khiến người dùng không nhận thức được rằng mình đang bị giới hạn trong một hệ sinh thái thông tin có chọn lọc.

Từ cơ chế cá nhân hóa này, một nguy cơ khác xuất hiện: sự cực đoan hóa thông tin. Khi người dùng liên tục được cung cấp các nội dung củng cố quan điểm đã có, niềm tin của họ không những không được kiểm chứng mà còn bị đẩy đến mức độ cực đoan hơn qua thời gian. Các nội dung mang tính kích động, giật gân, hay gây tranh cãi – vốn thường thu hút sự chú ý và tương tác cao – dễ dàng được thuật toán ưu tiên hiển thị, từ đó dẫn đến việc thông tin sai lệch, thiên lệch hoặc cực đoan lan truyền nhanh chóng. Nghiên cứu của Bakshy, Messing và Adamic (2015) trên Facebook cho thấy rằng người dùng có xu hướng tiêu thụ thông tin thiên lệch theo hướng phù hợp với lập trường chính trị của họ, và thuật toán đóng vai trò quan trọng trong việc “nuôi dưỡng” xu hướng này.

Không chỉ dừng lại ở cá nhân, hệ quả xã hội của hiện tượng này là sự phân mảnh thông tin và phân cực ý thức hệ trong cộng đồng. Các nhóm xã hội khác biệt, do tiếp cận thông tin khác nhau, phát triển các “thế giới quan” tách biệt – dẫn đến sự đứt gãy trong đối thoại xã hội và suy giảm khả năng hình thành đồng thuận chung. Ví dụ điển hình là trong các cuộc bầu cử hoặc khủng hoảng y tế công cộng (như đại dịch COVID-19), khi người dân sống trong các buồng vang thông tin mâu thuẫn, dẫn đến bất đồng dữ dội trong hành vi ứng xử và lựa chọn chính sách. Điều này đặc biệt nguy hiểm khi các thế lực cực đoan hoặc các chiến dịch truyền thông thao túng lợi dụng các thuật toán này để lan truyền thông tin sai lệch, thao túng dư luận hoặc gây bất ổn xã hội.

Bên cạnh đó, người dùng thường không ý thức được sự hiện diện của các buồng vang và bong bóng lọc trong hành vi tiếp nhận thông tin của mình. Do đó, họ tin rằng mình đang tiếp cận thông tin một cách khách quan, trong khi thực tế lại bị bao bọc bởi những thuật toán vô hình, vận hành theo các tiêu chí thương mại hoặc tư tưởng định hướng. Điều này làm xói mòn năng lực tư duy

phản biện, khiến công chúng khó nhận ra sự thiên lệch của hệ sinh thái thông tin họ đang tiêu thụ.

Ngoài ra, không thể bỏ qua trách nhiệm của các nền tảng truyền thông số và các công ty công nghệ lớn. Việc thiết kế hệ thống đề xuất nội dung dựa trên tối đa hóa tương tác thay vì tối đa hóa hiểu biết xã hội hoặc lợi ích cộng đồng đã góp phần đẩy mạnh các hiệu ứng lọc thông tin. Trong một số trường hợp, các nền tảng này còn có thể lợi dụng hệ sinh thái thông tin phân mảnh để thu hút lợi nhuận từ quảng cáo chính trị hoặc nội dung gây tranh cãi.

Từ đó, có thể thấy rằng quá trình cá nhân hóa nội dung, dù mang tính kỹ thuật nhưng lại dẫn đến các tác động xã hội sâu rộng. Việc nội dung được tùy biến quá mức cho từng người dùng khiến cho không gian truyền thông trở nên thiếu tính đối thoại và dễ bị lợi dụng bởi các chiến dịch tuyên truyền có chủ đích. Chính vì vậy, nhận diện đúng mối liên hệ giữa cá nhân hóa nội dung, buồng vang, bong bóng lọc và cực đoan hóa thông tin là bước khởi đầu quan trọng trong nỗ lực hướng tới một môi trường truyền thông số công bằng, đa chiều và lành mạnh hơn.

2.5. Giải pháp và khuyến nghị nhằm giảm thiểu hệ quả thông tin cực đoan, hướng tới xây dựng môi trường truyền thông số lành mạnh trong kỷ nguyên trí tuệ nhân tạo

Thứ nhất, điều chỉnh các thuật toán cá nhân hóa theo hướng minh bạch và vì lợi ích cộng đồng. Các nền tảng số cần công khai tiêu chí phân phối nội dung, cho phép người dùng kiểm soát mức độ cá nhân hóa. Việc tích hợp yếu tố “đa dạng thông tin” và “khuyến khích phản biện” vào thuật toán là cần thiết để đảm bảo người dùng không bị giới hạn trong không gian thông tin khép kín. Công nghệ AI có thể giải thích (Explainable AI) cũng nên được áp dụng nhằm giúp người dùng hiểu cách nội dung được lựa chọn và đề xuất.

Thứ hai, tăng cường sự can thiệp của các cơ quan quản lý và các tổ chức quốc tế nhằm định hình khuôn khổ đạo đức và pháp lý cho AI trong truyền thông. AI không thể vận hành hoàn toàn tự do khi có nguy cơ lan truyền thiên kiến và thao túng dư luận. Việc xây dựng các bộ quy tắc đạo đức, chuẩn mực vận hành AI trong truyền thông và các cơ chế kiểm tra độc lập là cần thiết để bảo vệ quyền tiếp cận thông tin đa chiều của công chúng.

Thứ ba, nâng cao năng lực tiếp nhận thông tin của công chúng thông qua giáo dục truyền thông số (digital and media literacy). Giáo dục truyền thông số cần được đẩy mạnh nhằm nâng cao khả

năng nhận diện thông tin sai lệch và thiên kiến cho người dùng. Các chương trình giáo dục chính thống nên tích hợp nội dung về tư duy phản biện và kỹ năng đánh giá thông tin. Đồng thời, các chiến dịch nâng cao nhận thức cộng đồng về thuật toán cá nhân hóa cũng rất cần thiết.

Thứ tư, thúc đẩy phát triển các mô hình nền tảng truyền thông minh bạch, độc lập và phi lợi nhuận. Trước sự chi phối của các nền tảng thương mại, cần có sự đầu tư vào các nền tảng báo chí chất lượng, nền tảng chia sẻ kiến thức cộng đồng, hoặc các hệ sinh thái thông tin phi lợi nhuận nhằm cung cấp nội dung trung lập, được kiểm chứng và mang giá trị giáo dục. Những nền tảng này có thể đóng vai trò như “trạm trung chuyển” giúp người dùng thoát khỏi tình trạng tiêu thụ thông tin lệch chiều và tạo điều kiện để tái thiết lập đối thoại giữa các nhóm xã hội có lập trường khác biệt.

Thứ năm, tăng cường các nghiên cứu liên ngành và sự phối hợp giữa nhà khoa học, nhà công nghệ và nhà hoạch định chính sách. Để giải quyết những hệ quả phức tạp của AI trong truyền thông, cần có sự hợp tác giữa nhiều lĩnh vực – từ khoa học dữ liệu, tâm lý học, truyền thông, chính trị học đến đạo đức công nghệ. Những nghiên cứu này sẽ cung cấp cơ sở lý luận và bằng chứng thực nghiệm nhằm thiết kế các mô hình AI nhân văn, cân bằng giữa lợi ích cá nhân và lợi ích tập thể, đồng thời làm cơ sở cho việc xây dựng các chính sách truyền thông công bằng, bền vững hơn.

Thứ sáu, có cơ chế khuyến khích, hỗ trợ các sáng kiến cộng đồng chống tin giả và lan tỏa thông

tin tích cực. Bên cạnh các cơ chế kiểm soát từ phía nền tảng và nhà nước, cần thúc đẩy các sáng kiến từ cộng đồng – đặc biệt là sự tham gia của báo chí độc lập, tổ chức xã hội dân sự, và các nhóm kiểm chứng thông tin (fact-checking). Những tổ chức này không chỉ giúp xác minh và phân bác thông tin sai lệch, mà còn đóng vai trò quan trọng trong việc xây dựng văn hóa truyền thông có trách nhiệm, qua đó tạo nên lực cản xã hội đối với xu hướng cực đoan hóa.

Thứ bảy, phát triển cơ chế hợp tác quốc tế trong lĩnh vực AI và quản trị truyền thông số. Vì truyền thông số là xuyên biên giới, nên các giải pháp chỉ mang tính nội địa sẽ không đủ để giải quyết triệt để vấn đề. Do đó, cần có sự phối hợp quốc tế trong việc thiết lập các bộ tiêu chuẩn đạo đức AI toàn cầu, chia sẻ dữ liệu nghiên cứu về ảnh hưởng của thuật toán đến dư luận, và tạo ra cơ chế giám sát xuyên quốc gia đối với các nền tảng lớn có ảnh hưởng toàn cầu. Liên minh châu Âu (EU) với Luật Dịch vụ Kỹ thuật số (DSA) là một ví dụ tiêu biểu trong việc thiết lập khung pháp lý chung để kiểm soát nội dung độc hại, có thể là tiền đề cho các sáng kiến tương tự tại khu vực châu Á – Thái Bình Dương, trong đó có Việt Nam.

III. KẾT LUẬN

Để giảm thiểu các tác động tiêu cực của AI và cơ chế cá nhân hóa trong truyền thông số, cần có cái nhìn toàn diện và các giải pháp mang tính hệ thống. Chỉ khi các yếu tố công nghệ, chính sách và xã hội được kết nối hiệu quả, chúng ta mới có thể xây dựng một môi trường truyền thông số công bằng, minh bạch và bền vững.

TÀI LIỆU THAM KHẢO

1. Nguyễn Văn Dũng, Đỗ Thị Thu Hằng (2018), *Truyền thông - Lý thuyết và kỹ năng cơ bản*, NXB Thông tin và Truyền thông, Hà Nội.
2. Castells, M. (2009), *Communication Power (Sức mạnh truyền thông)*, Oxford University Press.
3. McQuail, D. (2010), *McQuail's Mass Communication Theory (6th edition) (Lý thuyết truyền thông đại chúng của McQuail)*, SAGE Publications.
4. Pariser, E. (2011), *The Filter Bubble: What the Internet Is Hiding from You (Bong bóng lọc: Góc khuất của Internet mà bạn chưa biết)*, Penguin Press.
5. Russell, S., and Norvig, P. (2021), *Artificial Intelligence: A Modern Approach (4th edition) (Trí Tuệ Nhân Tạo: Một Cách Tiếp Cận Hiện Đại)*, Pearson.