

Trí thông minh nhân tạo (AI) và sự tham gia vào lĩnh vực truyền hình

TS. NGUYỄN NGÀ HUYỀN

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

Nhận ngày 3 tháng 3 năm 2024; chấp nhận đăng tháng 4 năm 2024.

Tóm tắt: Trí tuệ nhân tạo (Artificial Intelligence, viết tắt là AI) đã len lỏi vào hầu hết mọi lĩnh vực của đời sống xã hội trong bối cảnh của kỷ nguyên số. Ở những lĩnh vực có sự gắn kết mật thiết với kỹ thuật - công nghệ sẽ càng thấy rõ điều này. Đối với truyền hình, sự tham gia của trí thông minh nhân tạo vào nhiều công đoạn trong quá trình tổ chức sản xuất các tác phẩm, chương trình truyền hình hiện đã không còn xa lạ. Bài viết tổng hợp và giới thiệu một số sản phẩm, chức năng phổ biến từ AI đã và đang được áp dụng trong lĩnh vực này.

Từ khóa: trí tuệ nhân tạo, truyền hình, video.

Abstract: Artificial Intelligence (AI) has infiltrated almost every aspect of life in the digital age. This trend is particularly evident in fields closely linked to technology. In the television industry, AI's involvement in various stages of production for television programs and content is becoming increasingly common. This article synthesizes and introduces some common AI products and functions that are currently being applied in this field.

Keywords: artificial intelligence, television, video.

1. AI tạo video từ văn bản

Đối với truyền hình, hình ảnh là một trong những yếu tố làm nên thành công của tác phẩm. Để hỗ trợ công việc của các phóng viên truyền hình cũng như những người dẫn chương trình truyền hình, các ứng dụng hiện nay đều xây dựng kho dữ liệu về hình ảnh và âm thanh rất lớn để phục vụ người dùng. Giờ đây người dùng có thể không biết về các kỹ thuật biên tập, hậu kỳ hình ảnh và âm thanh thì vẫn có thể sử dụng các ứng dụng để làm việc đó ở cấp độ đơn giản. Những video được sản xuất từ video gốc, từ ảnh, từ văn bản hiện đã trở nên phổ biến và ngày càng được cải thiện, nâng cấp về công nghệ bởi các nhà sản xuất.

Trung tuần tháng 2 năm 2024, Open AI đã ra mắt Sora, một mô hình AI tạo video từ văn

bản. Theo mô tả, Sora có khả năng tạo ra video thời lượng tối đa 1 phút với chất lượng hình ảnh rất cao và bám sát câu lệnh (prompt) của người dùng.

Đây là câu lệnh mà Sora từ đó đã sản xuất ra video mẫu và giới thiệu trên toàn thế giới, tác giả tạm dịch: □ Một người phụ nữ lịch lãm bước đi dọc con đường ở Tokyo đầy ánh đèn neon ấm áp và biển hiệu thành phố nhấp nháy. Cô ấy mặc một chiếc áo khoác da màu đen, một chiếc váy đồ dài, đi đôi ủng đen, và mang theo một chiếc túi đen. Cô ấy đeo kính râm và son môi đỏ. Cô ấy đi bước tự tin và thoải mái. Con đường ẩm ướt và phản chiếu, tạo ra hiệu ứng gương của ánh sáng đầy màu sắc. Nhiều người đi bộ qua lại □

Một câu lệnh đơn giản khác của Sora đã đưa

ra kết quả : □Historical footage of California during the gold rush.□(tác giả tạm dịch: □Đoạn phim lịch sử về California trong thời kì cơn sốt đào vàng□).

Những video này của Sora được nhận định không khác gì video quay cảnh thật bởi các thiết bị sản xuất video chuyên nghiệp đương thời.



Ảnh 1. Ảnh cắt từ video của Sora AI.
Nguồn: <https://openai.com/sora>

Ngoài Sora, nhiều mô hình khác cũng được giới chuyên môn đánh giá cao về tiềm năng như Gen-2 của Runway Research. AI tạo sinh này cho phép sản xuất nhiều mô hình khác nhau: tạo video từ văn bản; tạo video từ văn bản và ảnh; tạo video từ ảnh; tạo video phong cách riêng từ video cho sẵn; chuyển đổi các bản mẫu mô phỏng thành các hình ảnh hoàn toàn được thiết kế và hoạt hình; phân định chủ thể trong video và sửa đổi chúng với các gợi ý văn bản đơn giản; chuyển đổi các đầu ra không có ngữ cảnh thành các kết quả thực tế bằng cách áp dụng một hình ảnh hoặc gợi ý đầu vào v.v..

Một số AI khác có thể kể đến là: PixVerse - AI miễn phí cho phép tạo video đa thể loại và kích cỡ, có thời lượng vài giây từ chữ và ảnh; Fliki - AI có chức năng dựng video từ văn bản có sẵn. Sau khi người dùng đưa văn bản vào, họ có thể lựa chọn một giọng đọc nam hoặc nữ phù hợp. Dựa trên ngôn ngữ của văn bản mà phần mềm sẽ tự động lựa chọn ngôn ngữ của giọng đọc đúng với ngôn ngữ. Sau đó, người dùng có thể lựa chọn các hình ảnh phù hợp (miễn phí và trả phí) để sắp xếp vào từng đoạn văn bản. Như vậy, họ hoàn toàn có thể tự sản xuất những video

đơn giản chỉ với thứ có sẵn là nội dung văn bản.

Giới phân tích đánh giá, nếu các công nghệ này được thương mại hoá, lĩnh vực truyền hình và điện ảnh sẽ lên một tầm cao mới khi mọi ý tưởng sẽ không còn bị giới hạn và đều có thể được hiện thực hoá. Tuy nhiên, vấn đề đặt ra là khung pháp lý khi sử dụng những công nghệ này vào các tác phẩm, sản phẩm báo chí và nghệ thuật. Khoảng cách giữa sáng tạo của con người và sáng tạo từ AI sẽ ra sao? Trong một tác phẩm báo chí, có được sử dụng hình minh họa ảo do AI sản xuất không, và nếu được sử dụng thì đến đâu là đủ? Nếu một nguyên mẫu từ đời thực được sáng tạo thành một nhân vật AI và sử dụng rộng rãi thì đặt ra những vấn đề gì về bản quyền? v.v. Đó là những câu hỏi mà những người làm truyền hình, điện ảnh mong muốn sử dụng công nghệ này cần cân nhắc.

2. Người dẫn chương trình AI

Giờ đây, việc người dẫn chương trình (NDCT) do AI tạo nên (người dẫn chương trình ảo) xuất hiện trên sóng truyền hình, trên các trang báo mạng điện tử, cũng như các nền tảng trực tuyến đã trở nên phổ biến.



Ảnh 2. Người dẫn chương trình ảo (AI) phiên bản nam giới của Tân Hoa Xã (Trung Quốc), lên sóng lần đầu tiên vào tháng 11 năm 2018.

Nguồn: <https://www.youtube.com/watch?v=GAfiATTQufk>

Đối với lĩnh vực truyền hình, quốc gia đầu tiên đưa NDCT ảo lên sóng là Trung Quốc. Tháng 11 năm 2018, Tân Hoa Xã chính thức giới thiệu NDCT ảo của mình với tạo hình và giọng nói của nam giới (ảnh 2). Đây là NDCT AI được dựng kỹ thuật số dựa trên hình ảnh của

một NDCT thật, sau đó sử dụng kỹ xảo để tạo ra cử động và biểu cảm trên gương mặt cũng như khẩu hình. Tân Hoa Xã mô tả đây là thành viên mới của toà soạn, người có khả năng đưa tin không biết mệt mỏi cả ngày và đưa tin từ mọi nơi trên khắp đất nước.

Chỉ sau đó vài tháng, vào tháng 2 năm 2019, Tân Hoa Xã tiếp tục ra mắt NDCT AI phiên bản nữ giới, được đánh giá là có giọng nói hấp dẫn, tự nhiên hơn phiên bản nam giới. Người dẫn chương trình AI nữ này cũng được tạo nên dựa trên nguyên mẫu là một NDCT thật của Tân Hoa Xã. Hai sản phẩm NDCT AI này là kết quả của sự hợp tác giữa Tân Hoa Xã và hãng công nghệ Sogou (công ty con của Tencent).

Năm 2023, Ấn Độ liên tiếp ra mắt các phiên bản NDCT AI. Đầu tiên là Sana, NDCT nữ giới với giao diện da trắng, tóc dài màu đen, trang phục hiện đại, lịch sự của India Today. Sana có thể dẫn chương trình với 75 ngôn ngữ, bao gồm tiếng Anh, Pháp, Hindi v.v.

Tiếp đó, vào tháng 7 năm 2023, đài Odisha TV của Ấn Độ ra mắt NDCT ảo trên chương trình tin tức của mình, đặt tên là Lisa. Khác với Sana, Lisa được tạo hình với trang phục truyền thống sari của Ấn Độ, có khả năng dẫn chương trình bằng tiếng Hindi và tiếng Odia.

Biểu cảm gương mặt của những NDCT AI của Trung Quốc và Ấn Độ rất giống người thật, nhưng họ thường không có các động tác tay như người dẫn chương trình thật. Tuy nhiên, sau này Trung Quốc đã phát triển NDCT từ nguyên mẫu người thật có động tác tay và ngôn ngữ cơ thể tự nhiên ngày càng chân thật, sinh động, thể hiện tiềm năng của công nghệ máy học (machine learning) và công nghệ học sâu (deep learning).

Ngoài Trung Quốc và Ấn Độ, một số quốc gia Châu Á khác cũng đã đưa NDCT AI lên sóng là: Hàn Quốc, Indonesia, Malaysia, Đài Loan, Kuwait v.v.. So với các châu lục khác, châu Á dường như có nhiều công bố, ra mắt về người dẫn chương trình AI hơn trong thời gian qua.

Tại Việt Nam, NDCT ảo cũng đã được lên sóng truyền hình ở mức độ giới thiệu và kết hợp dẫn cùng người dẫn chương trình thật như trong chương trình **Hitech - Công nghệ tương lai** (VTV1), chương trình **Nền tảng số** (VTC1). Mặc dù có động tác tay, cử động miệng, nhưng nhìn chung tạo hình gương mặt của NDCT ảo này còn cứng, và chưa được lên sóng định kỳ.

Bên cạnh truyền hình, các tờ báo mạng điện tử cũng đã có những bản tin trực tuyến với người dẫn chương trình AI như Báo Lao động điện tử, Báo Tuổi trẻ điện tử.

Không chỉ trên thế giới mà ngay tại Việt Nam đã có những công ty thiết kế những phần mềm tạo nên NDCT ảo. AiClip.ai (<https://aiclip.ai>) là thương hiệu nghiên cứu, thiết kế người dẫn chương trình AI chuyên nghiệp của Công ty Cổ phần Giải pháp Công nghệ trí thông minh nhân tạo thế hệ mới AiClip của Việt Nam. Đây cũng chính là đơn vị sản xuất người dẫn chương trình ảo cho các bản tin **Hitech - Công nghệ tương lai** (VTV1), **Nền tảng số** (VTC1) đã đề cập ở trên.

AiClip cho phép người dùng sử dụng bản miễn phí với dung lượng dưới 15 giây và giới hạn 10 lần sử dụng tạo video. Nếu người dùng trả phí, thời lượng và số lượng video sẽ tăng lên tùy theo mức phí. Với AiClip, người dùng có thể tạo video với NDCT AI giao diện người Việt, có hỗ trợ chuyển văn bản thành giọng nói.

Người dùng có thể chọn 6 người dẫn chương trình Việt Nam, với 2 giới tính nam và nữ, độ tuổi từ 20-30. Người dùng có thể thay trang phục cho người dẫn chương trình tùy theo sở thích của mình, hoặc tùy theo nội dung cần dẫn.

Một thương hiệu tạo NDCT ảo khác của Việt Nam là Dizim.ai (<https://dizim.ai>). Đây là thương hiệu của Công ty Cổ phần Dizim, có chức năng cho người dùng tạo ra người dẫn ảo cho video. Sau thời gian dùng thử miễn phí là 2 tuần, người dùng sẽ phải trả 2 đôla cho mỗi video. Điểm khác biệt của Dizim là có dịch vụ để những người nổi tiếng và có ảnh hưởng có

thể cho thuê hình đại diện của mình trên sàn giao dịch "virtual influencers" (người gây ảnh hưởng phiên bản ảo), và những ai có nhu cầu thuê một người nổi tiếng cụ thể nào đó xuất hiện trong video của mình có thể lựa chọn tại đây.

Có nét khác biệt với AiClip và Dizim, D-ID (<https://www.d-id.com>) là một ứng dụng sản xuất nhân vật ảo dựa theo mẫu có sẵn hoặc theo trí tưởng tượng của người dùng. Ưu điểm vượt trội của phần mềm này là có thể sáng tạo ra bất kỳ nhân vật nào theo mô tả bằng chữ viết hoặc mô tả bằng âm thanh. Nhân vật có thể được tạo ra dưới dạng video, có hình ảnh động và có thể nói, hoặc dưới dạng ảnh, dưới dạng hoạt hình, tranh vẽ v.v. Do được tích hợp cùng tính năng của ChatGPT, D-ID đã cho phép người dùng sáng tạo không giới hạn những nhân vật mình mong muốn, dựa theo chính sở thích và trí tưởng tượng của họ. Bên cạnh đó, người dùng có thể lựa chọn hàng trăm ngôn ngữ cho nhân vật của mình với hai giọng đọc nam - nữ riêng biệt. Điều này khiến cho tiềm năng của phần mềm này là rất lớn.

Đối với gương mặt người thật và 3D có sẵn trên D-ID, họ được thiết kế với đủ giới tính, chủng tộc, màu da, độ tuổi. Người dùng có thể tự sử dụng ảnh của mình để thiết kế theo ý thích. Khi nhập nội dung để nhân vật nói, người dùng chỉ việc gõ văn bản vào và tùy theo số lượng chữ mà phần mềm sẽ tính ra số giây và tính ra chi phí phải trả. Cũng giống như nhiều phần mềm thiết kế nhân vật ảo, D-ID cho người dùng sử dụng miễn phí 20 credit để làm quen. Đây là đơn vị để tính số thời lượng sử dụng cho mỗi video thiết kế, video càng dài thì sẽ tốn nhiều credit. Khi dùng hết credit, người dùng cần mua credit bằng tiền để tiếp tục sử dụng được.

HeyGen (<https://www.heygen.com>) là một ứng dụng cho phép thiết kế hình đại diện, nhân vật đại diện, người dẫn chương trình theo ý thích. Có hơn 100 hình đại diện có sẵn cho người dùng lựa chọn, hoặc người dùng có thể sử dụng hình ảnh khác tùy ý. Trong kho của

HeyGen có các mẫu video có sẵn để tham khảo như các video bán hàng giới thiệu sản phẩm, video chăm sóc khách hàng, video gửi lời chúc, video hướng dẫn v.v. Trong giao diện chỉnh sửa có hai phần: một là nhập chữ trực tiếp, hai là tải lên bản thu âm hoặc thu âm trực tiếp. Trong quá trình chỉnh sửa, người dùng có thể thêm các thành tố hoặc chữ viết vào. Nhìn chung quá trình tạo video với nhân vật lựa chọn rất đơn giản.

Như vậy, công nghệ AI tạo NDCT ảo hiện nay tạo nên một tiềm năng rất lớn đối với việc sản xuất nội dung truyền hình. Những nhà sản xuất nội dung hình ảnh hiện nay không còn cần phải phụ thuộc vào người thật để dẫn chương trình hoặc đọc lời bình minh họa cho video (và cả podcast) nữa. Mặc dù những công nghệ này đòi hỏi chi phí, nhưng nhìn về lâu dài, với sự cải thiện, cập nhật ngày càng nhanh của các phần mềm dựa trên công nghệ máy học và học sâu, thì những lợi ích mang lại từ đây là tiềm năng rất lớn. Các đài truyền hình lớn có thể không cần phụ thuộc vào công nghệ này do có đội ngũ sản xuất hùng hậu, nhưng đây có thể là giải pháp tốt cho các kênh truyền thông vừa và nhỏ, hoặc những cá nhân chuyên sản xuất nội dung trên các nền tảng số.

3. Một số AI khác

Ngoài thiết kế người dẫn chương trình ảo, việc tạo ra giọng đọc ảo cũng là một trong những điều mà trí thông minh nhân tạo đã làm được và ngày càng làm tốt. Có thể kể đến một số cái tên như: Vbee, Resemble, v.v..

Vbee (<https://vbee.vn>) là một phần mềm có chức năng tạo giọng nói cảm xúc từ văn bản. Người dùng chỉ cần nhập văn bản và lựa chọn các giọng đọc phù hợp. Có trên 30 ngôn ngữ khác nhau để người dùng lựa chọn, và mỗi ngôn ngữ đều có nhiều giọng đọc theo địa phương, theo giới tính và theo tính chất của văn bản. Tùy theo chất lượng, độ cảm xúc của giọng đọc mà nó được sử dụng miễn phí hoặc trả phí.

Còn Resemble.ai (<https://www.resemble.ai>) là thương hiệu chuyên về công nghệ giọng nói.

Chỉ cần tải lên giọng nói gốc, AI này sẽ có khả năng phát triển thành giọng AI từ đó, và sử dụng giọng này để thể hiện các văn bản như ý muốn. Như vậy, với công nghệ này, những người dẫn chương trình, các biên tập viên hoàn toàn có thể thực hiện đọc văn bản bằng giọng của mình mà không thực sự đọc nó. Resemble AI cũng có khả năng chỉnh sửa một giọng nói đã có, giúp nó trở nên liền mạch hơn. Ngoài ra, nó cũng có thể thiết kế giọng nói từ ngôn ngữ gốc sang 100 ngôn ngữ khác nhau để mở rộng đối tượng tiếp cận cho nội dung. Công ty sở hữu Resemble cũng phát triển dịch vụ Resemble Detect để giúp khách hàng nhận biết một giọng nói có phải được tạo ra từ công nghệ "deep fake" (giả mạo sâu) hay không.

Những ứng dụng này giúp cho việc đọc các lời bình trong tác phẩm truyền hình trở nên dễ dàng hơn, đồng thời, nó cũng giúp cho việc sáng tạo các tác phẩm truyền thông khác thuận tiện hơn như: các hệ thống trung tâm cuộc gọi tự động, hệ thống thông báo công khai, trợ lý ảo, tin tức âm thanh, sách âm thanh, podcast, tường thuật phim, đọc truyện v.v..

Bên cạnh giọng đọc ảo, giờ đây, việc người dẫn chương trình muốn xóa phông nền của các đoạn dẫn và thay thế bằng các phông nền tự chọn, hay muốn loại bỏ một vật thể trong khuôn hình đã là khả năng trong tầm tay. Các ứng dụng như Capcut, hay các phần mềm như Cutout (<https://cutout.pro>) đều có tính năng này ở các mức độ miễn phí và trả phí khác nhau. Người dùng chỉ cần có câu lệnh mô tả chính xác những hình ảnh, màu sắc, vật thể mà mình muốn có trong phông nền, phần mềm sẽ thiết kế và tạo ra phông nền theo ý muốn của họ. Người dùng muốn sản xuất, sáng tạo càng nhiều ảnh, video theo cách này thì sẽ càng cần trả nhiều phí.

Ngoài ra, không thể không kể đến việc chỉnh sửa, hậu kỳ video giờ đây đã có sự tham gia đắc lực của AI. Các công nghệ máy học đã hỗ trợ cho các kỹ thuật viên hậu kỳ video rất nhiều trong việc biên tập, chỉnh sửa video. Một số ví

dụ: IBM Watson đã hỗ trợ chỉnh sửa video cho giải golf The Masters, tiết kiệm thời gian lựa chọn và phân phối video nổi bật cho công chúng; Benjamin AI đã viết kịch bản và sáng tạo nội dung phim truyện; DaVinci Neural Engine sử dụng thuật toán máy học để chỉnh sửa màu sắc, tông màu, độ nhiễu của video, giúp công việc được thực hiện chính xác, tiết kiệm thời gian và hiệu quả hơn; Color.io có khả năng đề xuất chính xác chỉnh sửa màu của cảnh quay mới cho đạo diễn v.v.

Tạo nhạc bằng trí thông minh nhân tạo cũng là một trong những lợi ích mang lại cho lĩnh vực truyền hình vốn luôn cần sáng tạo về hình ảnh và âm thanh. Những phần mềm tạo nhạc AI hiện nay đã mở ra những cơ hội lớn đối với những người mong muốn được sử dụng nhạc theo hướng thương mại hoá với chi phí phải chăng so với việc mua bản quyền từ các nhạc sĩ. Các phần mềm phổ biến trong lĩnh vực tạo nhạc AI có thể kể đến là Aiva, Amper Music, Soundful v.v..

Với những ví dụ kể trên, không khó để nhìn thấy những gì mà trí thông minh nhân tạo có thể mang đến cho lĩnh vực truyền hình. Tuy nhiên, sự thú vị ban đầu có thể trở thành thách thức sau đó, nếu như những người đang làm việc trong lĩnh vực này thiếu đi sự thích nghi để hiểu cả những ích lợi và rào cản mà AI tạo nên. Sự xuất hiện của AI, cho đến nay, cũng vẫn đang đặt ra những câu hỏi về mặt đạo đức và pháp luật tại nhiều quốc gia trên thế giới, không riêng Việt Nam. Vì vậy, việc nhận thức về "sức mạnh" của AI là cần thiết, nhưng việc ứng dụng AI vào thực tiễn hoạt động truyền hình nói chung và báo chí truyền hình nói riêng, vẫn đòi hỏi ở mỗi người dùng một sự thận trọng nhất định không thừa./.

TÀI LIỆU THAM KHẢO

1. <https://openai.com/sora>.
2. <https://www.youtube.com/watch?v=GAfiATTQufk>.
3. <https://twitter.com/otvnews>.
4. <https://www.youtube.com/watch?v=4J1Hj3LzsIo>.
5. <https://www.youtube.com/watch?v=VexnwdlITQJM>.