

DEEFAKE VÀ THÔNG TIN KHÔNG KIỂM CHỨNG: HỆ QUẢ TÂM LÝ VÀ HỌC THUẬT ĐỐI VỚI SINH VIÊN KHỐI NGÀNH KINH TẾ TRONG KỶ NGUYÊN SỐ

Nguyễn Đức Quang, Phan Thị Thanh Chi
Khoa Luật, Trường Đại học Công nghệ Thành phố Hồ Chí Minh

Tóm tắt: Trong bối cảnh bùng nổ công nghệ trí tuệ nhân tạo, thông tin không kiểm chứng và các sản phẩm deepfake lan rộng với tốc độ chưa từng có. Sinh viên đại học, đặc biệt là nhóm sinh viên khối ngành kinh tế – nhóm học tập và làm việc trong môi trường đòi hỏi tiếp cận và xử lý lượng lớn thông tin – đang trở thành đối tượng dễ bị tác động mạnh mẽ. Bài báo này phân tích các ảnh hưởng tâm lý – nhận thức – hành vi của hiện tượng deepfake đối với sinh viên kinh tế, bao gồm: lo âu thông tin, khủng hoảng niềm tin, suy giảm khả năng ra quyết định, ảnh hưởng đến hiệu suất học tập và chất lượng đời sống xã hội, cùng nguy cơ bị thao túng trong môi trường số và kinh tế số. Đồng thời, bài báo tổng hợp các mô hình lý thuyết liên quan (thuyết tải nhận thức, thuyết lo âu thông tin, mô hình ra quyết định trong điều kiện bất định), kết quả nghiên cứu gần đây, và đề xuất chương trình can thiệp cho trường đại học nhằm nâng cao năng lực miễn nhiễm với thông tin giả.

Từ khóa: Deepfake; Thông tin không kiểm chứng; Sinh viên khối ngành kinh tế; Lo âu thông tin; Tải nhận thức; Niềm tin truyền thông; Sức khỏe tâm lý sinh viên.

DEEFAKES AND UNVERIFIED INFORMATION: PSYCHOLOGICAL AND ACADEMIC CONSEQUENCES FOR ECONOMICS STUDENTS IN THE DIGITAL ERA

Abstract: In the context of the rapid expansion of artificial intelligence technologies, unverified information and deepfake products are spreading at an unprecedented pace. University students—especially those majoring in economics, who study and work in environments that require constant access to and processing of large volumes of information—are becoming highly vulnerable to these impacts. This article analyzes the psychological, cognitive, and behavioral effects of the deepfake phenomenon on economics students, including information anxiety, crisis of trust, reduced decision-making capacity, negative impacts on academic performance and quality of social life, as well as the risk of manipulation in the digital and digital economy environment. At the same time, the paper synthesizes relevant theoretical models (cognitive load theory, information anxiety theory, decision-making models under uncertainty), recent empirical findings, and proposes a university-level intervention program aimed at enhancing students' resilience and immunity to misinformation.

Keywords: Deepfake; Misinformation; Economics students; Information anxiety; Cognitive load; Media trust; Students' mental health.

Nhận bài: 20/10/2025

Phản biện: 19/11/2025

Duyệt đăng: 24/11/2025

I. ĐẶT VẤN ĐỀ

Trong những năm gần đây, sự phát triển mạnh mẽ của trí tuệ nhân tạo, đặc biệt là học sâu, đã thúc đẩy sự bùng nổ của nội dung số tổng hợp, trong đó deepfake là công nghệ gây lo ngại hàng đầu. Với khả năng tạo hình ảnh, âm thanh, video mô phỏng con người gần như thật, deepfake đe dọa nghiêm trọng tính xác thực thông tin, làm xói mòn niềm tin truyền thông và gia tăng nguy cơ thao túng dư luận, tấn công danh dự, xâm phạm đời tư. Trong bối cảnh đó, sinh viên khối ngành kinh tế là nhóm dễ tổn thương do thường xuyên tiếp cận lượng lớn thông tin về tài chính, doanh nghiệp, kinh tế vĩ mô và phụ thuộc mạnh vào độ tin cậy dữ liệu trong học tập, phân tích và định hướng nghề nghiệp. Deepfake có thể dẫn đến lo âu thông tin, quá tải nhận thức, suy giảm niềm tin vào nguồn tin chính thống và các hệ quả tâm lý như lo âu, trầm cảm, mất an toàn cá nhân. Dựa trên các khung lý thuyết về khủng hoảng niềm tin

thông tin, tải nhận thức và ra quyết định trong điều kiện bất định, bài báo nhằm phân tích tác động tâm lý – nhận thức của deepfake đối với sinh viên kinh tế, lý giải vì sao họ đặc biệt dễ tổn thương, làm rõ hệ quả đối với học tập và đời sống xã hội, đồng thời đề xuất giải pháp tăng năng lực “miễn nhiễm thông tin” và kỹ năng nhận diện nội dung giả trong bối cảnh chuyển đổi số giáo dục đại học.

II. NỘI DUNG NGHIÊN CỨU

2.1. Tác động tâm lý và học thuật của deepfake đối với sinh viên khối ngành kinh tế

Sự lan rộng của công nghệ deepfake những năm gần đây đã tạo ra những tác động đáng kể lên đời sống tâm lý, nhận thức và hoạt động học thuật của sinh viên, đặc biệt là sinh viên khối ngành kinh tế – nhóm thường xuyên tiếp cận lượng lớn thông tin đa phương tiện về tài chính, doanh nghiệp và kinh tế – xã hội. Do đặc thù phải dựa mạnh vào dữ liệu số, sinh viên kinh tế chịu ảnh hưởng rõ rệt hơn

trước sự nhiễu loạn thông tin và các thách thức liên quan đến độ xác thực của nội dung trực tuyến. Về tâm lý – cảm xúc, việc ý thức rằng hình ảnh, giọng nói hoặc dữ liệu cá nhân có thể bị chiếm dụng để tạo nội dung giả mạo gây bối rối, quấy rối dẫn tới trạng thái bất an kéo dài, đặc biệt ở nữ sinh và các nhóm dễ bị tổn thương. Các trường hợp deepfake mang tính xúc phạm danh dự, gắn với quấy rối tình dục hoặc lừa đảo tài chính được ghi nhận kèm theo các hệ quả như trầm cảm, né tránh xã hội, suy giảm cảm giác an toàn và giảm động lực học tập.

Ở bình diện nhận thức và học thuật, sinh viên kinh tế phải xử lý khối lượng lớn thông tin từ báo chí chuyên ngành, báo cáo thị trường, phân tích tài chính. Sự pha trộn giữa thông tin thật – giả làm gia tăng “lo âu thông tin” và “mệt mỏi thông tin”, buộc người học liên tục kiểm chứng nguồn tin, nhất là các tư liệu đa phương tiện có nguy cơ bị deepfake. Điều này làm tăng tải nhận thức, giảm khả năng tập trung, khiến sinh viên có xu hướng tránh các chủ đề phức tạp, hạn chế chiều sâu tiếp thu tri thức và dần hình thành trạng thái hoài nghi cực đoan, suy giảm niềm tin vào cả nguồn tin chính thống lẫn phi chính thống.

Deepfake cũng ảnh hưởng trực tiếp tới chất lượng ra quyết định – một năng lực trung tâm trong đào tạo kinh tế. Khi thông tin trở nên thiếu ổn định, sinh viên dễ dựa vào cảm tính, heuristic hơn là phân tích lý trí, đặc biệt trong bối cảnh video chuyên gia tài chính, CEO hoặc tín hiệu thị trường có thể bị giả mạo. Điều này dẫn đến phân tích sai lệch, đánh giá không chính xác về doanh nghiệp và quyết định học tập, đầu tư thiếu cơ sở. Ở góc độ hành vi – xã hội, nhiều sinh viên chủ động thu hẹp “dấu chân số”, hạn chế xuất hiện trên mạng xã hội hoặc các sản phẩm học tập ghi hình để tự bảo vệ, nhưng đồng thời đánh mất cơ hội xây dựng thương hiệu cá nhân – vốn rất quan trọng trong lĩnh vực kinh tế, kinh doanh. Deepfake còn có thể trở thành công cụ bắt nạt trực tuyến, làm suy giảm niềm tin lẫn nhau và gây khó khăn cho học tập hợp tác.

Tổng thể, các tác động tâm lý, nhận thức và hành vi mà deepfake gây ra cho sinh viên kinh tế liên kết thành một chuỗi: tiếp xúc nội dung giả mạo → cảm giác thiếu kiểm soát, bất an → tăng lo âu, tải nhận thức → xuất hiện các chiến lược né tránh, phòng vệ → suy giảm hiệu quả học tập và ảnh hưởng đến định hướng nghề nghiệp. Điều này lý giải vì sao sinh viên khối ngành kinh tế –

nhóm tiếp xúc dữ liệu dày đặc và phụ thuộc nhiều vào thông tin số – trở thành một trong những đối tượng chịu tác động mạnh mẽ nhất trước sự bùng nổ của deepfake.

2.2. Các yếu tố khiến sinh viên khối ngành kinh tế dễ tổn thương trước deepfake

Sự phát triển nhanh chóng của công nghệ deepfake không tác động đồng đều lên mọi nhóm người dùng mà khiến một số nhóm trở nên dễ tổn thương hơn do đặc điểm học tập, hành vi trực tuyến hoặc áp lực nghề nghiệp. Sinh viên khối ngành kinh tế là một trong những nhóm này bởi mức độ tiếp xúc mật độ cao với thông tin tài chính, sự phụ thuộc vào dữ liệu trực tuyến và nhu cầu xây dựng hình ảnh chuyên nghiệp ngay từ giai đoạn đại học. Một trong những yếu tố quan trọng nhất góp phần gia tăng tính dễ tổn thương là mức độ phơi nhiễm thông tin cao. Sinh viên kinh tế thường xuyên theo dõi báo cáo thị trường, phân tích tài chính và tin tức doanh nghiệp, vốn ngày càng được trình bày dưới dạng đa phương tiện. Chính dạng thức trực quan này khiến họ dễ bị thuyết phục bởi nội dung thao túng, đặc biệt là các video mang tính thuyết phục cao—dù là deepfake—có khả năng “áp đặt tính xác thực trực quan” mạnh hơn văn bản (Vaccari & Chadwick, 2020). Trong bối cảnh tài chính, nơi phát biểu của CEO hay quan chức ngân hàng có thể tác động tức thì đến nhận thức thị trường, việc sinh viên tin vào video giả mạo dễ dẫn đến đánh giá sai lệch hoặc quyết định học thuật lẫn đầu tư thiếu chính xác.

Bên cạnh phơi nhiễm thông tin, áp lực cạnh tranh nghề nghiệp khiến sinh viên kinh tế duy trì mức độ “hiện diện số” (digital presence) cao, từ hồ sơ LinkedIn đến video thuyết trình học thuật. Điều này làm gia tăng nguy cơ bị chiếm dụng dữ liệu cá nhân, bao gồm hình ảnh và giọng nói, để tạo deepfake (Navarro Martínez et al., 2024). Nhiều sinh viên do lo ngại lạm dụng dữ liệu đã hạn chế xuất hiện trong các bài trình bày hoặc tránh đăng tải nội dung học thuật, ảnh hưởng đến cơ hội rèn luyện kỹ năng và xây dựng thương hiệu cá nhân. Ngoài ra, đặc thù ngành học khiến sinh viên kinh tế thường tự đánh giá cao năng lực phân tích và nhận diện thông tin (Kahneman & Tversky, 1979; Lewandowsky et al., 2012). Tuy nhiên, nghiên cứu của Nightingale và Farid (2022) cho thấy con người thường đánh giá quá cao khả năng phát hiện deepfake, trong khi độ chính xác thực tế chỉ hơn đoán ngẫu nhiên ở mức hạn chế. Sự chênh

lệch giữa nhận thức chủ quan và khả năng thực tế khiến sinh viên dễ rơi vào cảm giác tự trách hoặc giảm tự tin học thuật khi phát hiện mình bị đánh lừa bởi nội dung “trông có vẻ chuyên nghiệp”.

Sinh viên kinh tế cũng dễ bị tổn thương do yêu cầu làm việc với thông tin thời sự—vốn thay đổi nhanh và đôi khi thiếu kiểm chứng. Swire-Thompson và Lazer (2020) chỉ ra rằng tin giả thường tận dụng tốc độ lan truyền của tin nóng để thao túng nhận thức người dùng. Sinh viên kinh tế, quen với việc tiêu thụ thông tin thị trường cập nhật nhanh, vì vậy trở thành mục tiêu lý tưởng của các nội dung deepfake được thiết kế để gây ảnh hưởng đến nhận định của người trẻ. Một số trường hợp thực tế minh họa rõ điều này: một sinh viên ngành tài chính tham gia đầu tư sau khi xem video deepfake giả danh CEO và sau đó mất mát tài chính, đồng thời suy giảm niềm tin vào năng lực phân tích của bản thân (Navarro Martínez et al., 2024); hay các trường hợp sinh viên bị ghép mặt vào video khiêu dâm giả mạo dẫn đến trầm

cảm và rút lui khỏi hoạt động học thuật (Norton Rose Fulbright, 2024). Trong môi trường học tập, việc nghi ngờ tính xác thực của các sản phẩm đa phương tiện có thể làm giảm chất lượng hợp tác nhóm, như Tkáčová (2025) ghi nhận trong bối cảnh học tập hậu đại dịch.

Tổng hợp lại, các yếu tố khiến sinh viên kinh tế dễ tổn thương trước deepfake bao gồm: phơi nhiễm thông tin cao, áp lực xây dựng hình ảnh cá nhân, sự phụ thuộc vào tin tức thời sự và xu hướng tự tin quá mức vào khả năng nhận diện thông tin. Những yếu tố này tương tác và tạo thành một môi trường rủi ro, trong đó nội dung giả mạo có thể ảnh hưởng trực tiếp đến tâm lý, nhận thức, hành vi và định hướng nghề nghiệp của sinh viên. Các trường hợp minh họa cho thấy hậu quả deepfake không chỉ tức thời mà còn kéo dài, tác động sâu sắc đến trải nghiệm học tập và sức khỏe tinh thần của nhóm sinh viên này.

2.3. Mô hình, giải pháp nghiên cứu đề xuất

2.3.1. Mô hình nghiên cứu đề xuất



Hình 1. Mô hình nghiên cứu tác động của thông tin và deepfake đến tâm lý sinh viên

Dựa trên các kết quả phân tích ở những phần trước, mô hình nghiên cứu đề xuất về tác động của deepfake đối với sinh viên khối ngành kinh tế được xây dựng theo hướng tiếp cận tâm lý – hành vi. Mô hình này bao gồm bốn nhóm yếu tố chính: (1) Tiếp xúc deepfake, (2) Đánh giá nhận thức, (3) Phản ứng tâm lý – hành vi, và (4) Kết quả học thuật – xã hội. Mô hình phản ánh chuỗi tác động liên tục, trong đó mỗi giai đoạn có thể khuếch đại mức độ ảnh hưởng của giai đoạn tiếp theo. (Hình 1)

Ở giai đoạn đầu, sinh viên tiếp xúc với nội dung deepfake trực tiếp (ví dụ video, âm thanh, hình ảnh) hoặc gián tiếp thông qua thông tin không kiểm chứng lan truyền trong môi trường số. Giai đoạn thứ hai là tiến trình đánh giá nhận thức, nơi sinh viên cân nhắc mức độ xác thực của nội dung, khả năng kiểm soát rủi ro và mức đe dọa cá nhân. Theo thuyết đánh giá nhận thức của Lazarus (1991), nhận thức về mối đe dọa có vai trò quyết định đối với phản ứng cảm xúc.

Tiếp theo là giai đoạn phản ứng tâm lý, bao gồm lo âu thông tin, bất an cá nhân, stress học thuật và giảm niềm tin vào các nguồn tin chính thống. Những phản ứng này dẫn đến các hành vi ứng phó như hạn chế tiếp xúc thông tin, giảm hiện diện số, hoặc ngược lại — tăng mức độ kiểm chứng, tìm kiếm hỗ trợ từ giảng viên và chuyên gia. Cuối cùng, các phản ứng và hành vi này tác động trực tiếp đến kết quả học tập (ví dụ giảm điểm số, giảm chất lượng bài nghiên cứu), đời sống xã hội (giảm tương tác nhóm, tăng xung đột), và định hướng nghề nghiệp của sinh viên.

2.3.2 Các giải pháp đề xuất

Dựa trên mô hình trên và các kết quả phân tích, nghiên cứu đề xuất ba nhóm giải pháp mang tính hệ thống: giải pháp kỹ năng – giải pháp tổ chức – giải pháp chính sách.

2.3.2.1. Nhóm giải pháp kỹ năng: Nâng cao năng lực miễn nhiễm thông tin (Information Resilience)

Sinh viên kinh tế cần được trang bị năng lực nhận diện thông tin trong môi trường số, bao gồm phân tích nguồn tin, nhận diện thao túng, và hiểu cơ bản về cách hoạt động của deepfake. Kietzmann et al. (2020) khẳng định rằng năng lực số (digital literacy) là yếu tố cốt lõi để giảm

tác động của deepfake. Các khóa học bắt buộc hoặc ngoại khóa về phòng chống tin giả, kỹ thuật kiểm chứng nguồn tin, kiến thức căn bản về AI/Deepfake sẽ giúp sinh viên tăng khả năng phát hiện bất thường, từ đó giảm tải nhận thức và lo âu thông tin.

Bên cạnh đó, sinh viên cũng nên được khuyến khích phát triển “thói quen kiểm chứng” như sử dụng công cụ xác thực hình ảnh, video (ví dụ: InVID, Deepware Scanner), so sánh thông tin đa nguồn, và hỏi ý kiến giảng viên/dịch vụ thư viện của nhà trường.

2.3.2.2. Nhóm giải pháp tổ chức: Xây dựng hệ sinh thái học thuật an toàn thông tin

Các trường đại học cần đóng vai trò then chốt trong việc giảm thiểu rủi ro deepfake. Một trong những biện pháp quan trọng là thiết lập phòng hỗ trợ thông tin (Information Support Unit) nhằm hướng dẫn sinh viên cách xác minh nguồn tin. Đồng thời, trường có thể triển khai các workshop định kỳ về an ninh mạng, nhận diện deepfake và đạo đức sử dụng dữ liệu.

Ngoài ra, nhà trường cần mở rộng hệ thống tư vấn tâm lý và hỗ trợ nạn nhân deepfake. Theo Norton Rose Fulbright (2024), các nạn nhân chịu ảnh hưởng sâu sắc nếu không nhận được hỗ trợ kịp thời. Việc xây dựng quy trình báo cáo và xử lý deepfake bồi nhọ trong trường học sẽ giúp sinh viên cảm thấy an toàn và tự tin hơn.

2.3.2.3. Nhóm giải pháp chính sách: Khung bảo vệ người dùng và quản lý nội dung số

Các đơn vị đào tạo khối kinh tế nên phối hợp với doanh nghiệp, tổ chức công nghệ và cơ quan quản lý để xây dựng khung chính sách bảo vệ sinh viên khỏi nguy cơ thao túng thông tin. Chính sách có thể bao gồm quy định về:

- xử lý vi phạm deepfake trong môi trường giáo dục,
- quy chuẩn khi sử dụng nội dung đa phương tiện trong bài tập học thuật,
- hướng dẫn sử dụng dữ liệu và hình ảnh cá nhân,
- tích hợp bộ công cụ chống giả mạo vào hệ thống LMS của nhà trường.

Những biện pháp mang tính chính sách này không chỉ giảm rủi ro tâm lý mà còn tạo ra môi trường học tập công bằng và an toàn hơn cho sinh viên.

Bảng 1. Tóm tắt kết quả chính của nghiên cứu

Khía cạnh	Kết quả chính	Nguồn tham khảo tiêu biểu
Tâm lý – cảm xúc	Sinh viên trải qua lo âu thông tin, bất an cá nhân, giảm niềm tin vào nội dung số	Digital Trauma Working Group (2025); Navarro Martínez et al. (2024)
Nhận thức – học thuật	Tăng tải nhận thức, giảm hiệu quả phân tích, nguy cơ trích dẫn sai	Sweller (1988); Lewandowsky et al. (2012)
Hành vi – xã hội	Giảm hiện diện số, ngại giao tiếp, xung đột nhóm	Nightingale & Farid (2022); Tkáčová (2025)
Quy trình tác động	Deepfake → đánh giá nhận thức → phản ứng cảm xúc → hành vi ứng phó → kết quả học thuật – xã hội	Mô hình đề xuất của nghiên cứu
Đề xuất giải pháp	Nâng cao năng lực miễn nhiễm thông tin, xây dựng môi trường an toàn, hoàn thiện chính sách	Kietzmann et al. (2020); Norton Rose Fulbright (2024)

Các kết quả phân tích cho thấy deepfake không chỉ là vấn đề công nghệ mà còn là vấn đề tâm lý – xã hội sâu sắc, đặc biệt đối với sinh viên khối ngành kinh tế. Việc tiếp xúc thường xuyên với tin tức kinh tế – tài chính khiến nhóm này bị đặt vào môi trường thông tin có tính bất định cao, nơi deepfake có thể nhanh chóng lan truyền và tác động đến nhận thức, cảm xúc và hành vi. Điều này phù hợp với mô hình ảnh hưởng của misinformation được mô tả trong các nghiên cứu trước đây (Lewandowsky et al., 2012; Swire-Thompson & Lazer, 2020).

Một đóng góp quan trọng của nghiên cứu là chỉ ra sự tương tác giữa lo âu thông tin, tải nhận thức, và giảm niềm tin trong môi trường học thuật. Những yếu tố này không xuất hiện rời rạc mà tạo thành chuỗi phản ứng tâm lý – nhận thức, khiến sinh viên ngày càng khó duy trì hiệu quả học tập cũng như sự tự tin trong đánh giá thông tin chuyên môn. Các trường hợp minh họa thực tế cho thấy hậu quả của deepfake có thể kéo dài và nghiêm trọng hơn dự đoán, bao gồm tác động đến định hướng nghề nghiệp và sức khỏe tinh thần lâu dài.

Các giải pháp đề xuất của nghiên cứu gợi ý rằng việc phát triển “năng lực miễn nhiễm thông tin” (information resilience) nên trở thành mục tiêu trọng tâm của giáo dục đại học trong kỷ nguyên AI. Chỉ khi sinh viên được trang bị đúng kỹ năng, môi trường học thuật được bảo vệ và các chính sách quản trị nội dung số được thiết lập, tác động tiêu cực của deepfake mới có thể được giảm thiểu.

III. KẾT LUẬN

Sự phát triển mạnh mẽ của trí tuệ nhân tạo, đặc biệt là deepfake, đang đặt sinh viên khối ngành

kinh tế vào một bối cảnh học tập và xã hội phức tạp hơn bao giờ hết. Với đặc thù phải tiếp xúc liên tục với tin tức kinh tế, báo cáo doanh nghiệp và dữ liệu thị trường, sinh viên kinh tế không chỉ đối mặt với áp lực xử lý thông tin mà còn chịu những tác động sâu sắc về tâm lý, nhận thức và hành vi. Các phân tích trong nghiên cứu cho thấy deepfake làm gia tăng lo âu thông tin, xói mòn niềm tin vào môi trường học thuật và gây ra quá tải nhận thức—những yếu tố trực tiếp làm suy giảm hiệu suất học tập và năng lực ra quyết định, vốn là trọng tâm trong đào tạo kinh tế.

Bên cạnh ảnh hưởng đến nhận thức chuyên môn, deepfake còn đặt ra nhiều rủi ro đối với đời sống xã hội và cảm xúc của sinh viên. Thực tiễn cho thấy các nội dung giả mạo có thể dẫn đến xung đột trong học tập, tổn hại danh dự cá nhân, thậm chí gây sang chấn tâm lý. Điều này khẳng định rằng deepfake không chỉ là vấn đề công nghệ mà là thách thức toàn diện đối với hệ sinh thái giáo dục đại học.

Mô hình nghiên cứu đề xuất cho thấy tác động của deepfake diễn ra theo chuỗi: từ tiếp xúc thông tin giả → đánh giá nhận thức → phản ứng cảm xúc → hậu quả học thuật và xã hội. Cơ chế này mở rộng hiểu biết về ảnh hưởng của deepfake trong lĩnh vực tâm lý học truyền thông, đồng thời cho thấy tầm quan trọng của việc tăng cường khả năng miễn nhiễm thông tin và kỹ năng kiểm chứng. Các giải pháp được đề xuất—bao gồm nâng cao năng lực nhận diện nội dung giả, xây dựng môi trường học thuật an toàn và cải thiện chính sách quản trị nội dung số—đóng vai trò thiết thực trong việc giảm thiểu rủi ro.

TÀI LIỆU THAM KHẢO

- Chesney, R., & Citron, D. (2019). *Deepfakes and the new disinformation war*. California Law Review, 107(6), 1753–1820.
- Digital Trauma Working Group. (2025). *Digital trauma: Deepfake victimisation and AI-generated violence*. The Lancet Psychiatry.
- Diel, A., Lalgı, T., Schröter, I. C., MacDorman, K. F., Teufel, M., & Bäuerle, A. (2024). Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. *Computers in Human Behavior Reports*, 16, 100538.
- Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature. *The Information Society*, 20(5), 325–344.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Kietzmann, J., Lee, L., McCarthy, I., & Kietzmann, T. C. (2020). *Deepfakes: Trick or treat?* *Business Horizons*, 63(2), 135–146.
- Koopman, C., & Thomas, A. (2023). The risks of synthetic media: Deepfakes and the erosion of trust. *Journal of Cyber Policy*, 8(1), 44–62.
- Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131.
- Mackintosh, A., Watson, G., & McGuire, M. (2023). Synthetic influence: AI-generated media and online manipulation. *Journal of Digital Forensics, Security and Law*, 18(2), 1–21.
- Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), 1–41.
- Navarro Martínez, O., Fernández-García, D., Cuartero Monteagudo, N., & Forero-Rincón, O. (2024). Possible health benefits and risks of deepfake videos: A qualitative study in nursing students. *Nursing Reports*, 14(4), 2746–2757.
- Nightingale, S. J., & Farid, H. (2022). Testing human ability to detect ‘deepfake’ images of human faces. *Cybersecurity*, 9(1), 1–17.
- Norton Rose Fulbright. (2024). Deepfake misuse and implications for victims. *Norton Rose Fulbright Insights*.
- Riba, P., Martos, A., Pérez, J., & Fernandez, C. (2024). Deepfake generation and detection: A comprehensive review. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 20(1), 1–32.
- Salim, N., Malik, M., & Abdullah, S. (2023). Deepfake and media trust: A systematic review. *Information Processing & Management*, 60(2), 103186.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- Swire-Thompson, B., & Lazer, D. (2020). Public health and online misinformation: Challenges and recommendations. *Annual Review of Public Health*, 41, 433–451.
- Tkáčová, H. (2025). Challenges of misinformation in online learning: A post-pandemic perspective. *Encyclopedia*, 5(1), 25.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact on trust in news. *Social Media + Society*, 6(1), 1–13.
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52.
- Wurman, R. S. (1989). Information anxiety. *Doubleday*.