Asia Pacific Science Press

Journal of Educational Theory and Practice
ISSN(O):3005-9313
ISSN(P):3005-9305
Vol. 2 No.4 (2025)

Teaching Innovation and Competency Development in the Age of Generative AI: A Case Study of the Course "Online Communication and Public Opinion Supervision"

Zhuo Wang*

School of digital media, Guangzhou University of Software, Guangzhou, 510900, China

Corresponding author: Zhuo Wang

Abstract: Generative Artificial Intelligence (AIGC) has emerged as a disruptive force, fundamentally altering digital communication and information ecosystems. This development poses significant challenges to public opinion supervision, as the lines between authentic and synthetic content are increasingly blurred. Traditional communication curricula are often unprepared to address the threats of deepfakes, automated manipulation, and authenticity crises. This paper presents a pedagogical case study of a redesigned undergraduate course, "Online Communication and Public Opinion Supervision," at a teaching-focused university. To address AIGC-related challenges, a series of teaching innovations were implemented. These included updated curriculum modules on AIGC, adversarial "red team vs. blue team" simulations, AI-augmented project-based assignments, and oral crisis response drills. The aim was to cultivate four core competencies: technical understanding, critical discrimination, ethical judgment, and human-AI collaboration. Empirical outcomes from the course implementation were evaluated. Significant improvements in students' analytical performance, ethical reasoning, and engagement were observed. This case study provides insights into effective pedagogical strategies. It offers a model for adapting communication and media education to prepare students for the complexities of the generative AI era.

Keywords: Generative AI; AIGC; Teaching Innovation; Pedagogy; Public Opinion Supervision; Online Communication; AI Literacy; Competency Development

Published: Nov 14, 2025

DOI: https://doi.org/10.62177/jetp.v2i4.847

1.Introduction

Generative Artificial Intelligence (AI), frequently identified by the acronym AIGC (AI-Generated Content), has rapidly transitioned from a niche technical concept to a disruptive mainstream force in digital communication. Since late 2022, the widespread availability of advanced generative models, such as large language models (LLMs) and diffusion-based image generators, has empowered a global user base to create highly realistic text, images, audio, and video with unprecedented ease[1]. This technological leap is fundamentally blurring the line between authentic, human-created content and synthetic, machine-generated media. As a consequence, traditional cues for establishing credibility and trust in the digital sphere are being systematically undermined, presenting a profound societal challenge. The pervasive integration of AIGC has been noted by scholars as becoming "deeply embedded" within human communication processes. A new "GAI logic" is being introduced, which is actively reshaping established media structures, discourse norms, and the very economics of content

production. In practice, the societal impact is already being witnessed. AI-generated rumors, sophisticated deepfakes, and automated influence campaigns can escalate and influence real-world events at a velocity that outpaces traditional verification mechanisms. Such incidents, ranging from synthetic images of political events to fabricated audio of public figures, highlight an emerging and deepening crisis of authenticity in the public sphere [2].

This disruptive influence of AIGC on the formation and flow of public opinion has drawn significant global concern from academics, policymakers, and industry leaders. Governments worldwide are beginning to respond with regulatory frameworks. For example, China's 2023 "Interim Measures for the Management of Generative AI Services" emphasizes a dual approach: promoting technological innovation while simultaneously mitigating societal risks. These measures explicitly require that AI-generated content must be accurate, reliable, and align with core societal values. Concurrently, international organizations and technology consortiums are stressing the urgent need to develop and implement robust safeguards to protect the global information ecosystem from malicious AIGC applications. Amid this fast-evolving and uncertain landscape, educators, particularly in the fields of media, communication, and journalism, face an urgent and complex challenge. Students—especially future media professionals, policymakers, and corporate communicators—must be adequately prepared to navigate, analyze, and responsibly supervise online public opinion in an age saturated with generative AI. Traditional curricula, often designed for the social media era, must be fundamentally rethought. Educational models must now account for new threat vectors, such as hyper-realistic deepfakes, algorithmically-driven bot networks, and the subtle biases embedded within AI models. Simultaneously, these curricula must also incorporate the new analytical tools and collaborative workflows that AIGC offers[3].

This paper examines these multifaceted challenges and details the pedagogical responses developed within a specific academic context. A comprehensive case study of a redesigned undergraduate course, "Online Communication and Public Opinion Supervision," is presented. This course, a core requirement for New Media majors at a teaching-focused university, was recently and substantially revamped to address the AIGC revolution. The paper is structured as follows: First, a literature review is conducted to analyze the specific disruptions generative AI brings to the public opinion ecosystem, focusing on three key areas. Following this analysis, the core competencies that students must now acquire are identified and defined. Next, the core of the paper details the innovative teaching strategies that were designed and implemented. These strategies range from updated theoretical modules and adversarial "red team vs. blue team" simulations to project-based learning assignments integrated with AI tools. These strategies were explicitly designed to cultivate the previously identified competencies.

Subsequently, the empirical outcomes from the course are presented and analyzed. These observations, drawn from classroom exercises, assignment analysis, and qualitative student feedback, are used to evaluate the impact of the innovations on student performance, AI literacy, and engagement. The goal is to derive actionable insights into effective teaching innovations. These innovations should foster the critical, analytical, and ethical skills necessary for future media professionals. Ultimately, this case study aims to shed light on how university-level curriculum and pedagogy can be effectively adapted, ensuring that students are prepared to become competent and responsible supervisors of public opinion in the complex generative AI era.

2.The Disruptive Influence of Generative AI on Public Opinion Ecosystems

The public opinion and information ecosystem is being fundamentally and irreversibly transformed by the proliferation of generative AI. This transformation is not incremental; it represents a paradigm shift in how information is created, disseminated, and consumed. The influence of AIGC can be understood through several critical, interrelated challenges that strike at the heart of public discourse.

2.1 Authenticity and Credibility Crisis

Perhaps the most pervasive and immediate issue is the systemic erosion of epistemic trust in online content. Generative models are now capable of producing text, visuals, and audio that are indistinguishable from genuine, human-created content for the average user, and often for trained experts as well. This capability makes it increasingly difficult, and in some cases impossible, for both the general public and professional fact-checkers to verify authenticity at scale. As a result, highly sophisticated misinformation and disinformation can be injected directly into public discourse, bypassing traditional

journalistic gatekeepers and spreading rapidly through social networks. This phenomenon exacerbates the conditions of a "post-truth" environment, where objective facts are less influential in shaping public opinion than appeals to emotion and personal belief [4]. A "liar's dividend" has been warned of by scholars: a scenario where the mere existence of convincing fake content (like deepfakes) provides malicious actors with plausible deniability, allowing them to dismiss real events, authentic videos, or genuine documents as "fake." This uncertainty further undermines public trust in foundational institutions, including media, government, and science.

Public opinion supervision, which as a discipline relies on the ability to establish a baseline of factual truth and hold actors accountable, faces a profound methodological challenge when factual reality itself becomes a disputed commodity. In one pertinent Chinese analysis, it was observed that AIGC-driven false information is effectively "hijacking social trust" by creating fractured cognitive pictures of reality among the populace, thereby manipulating discourse orientation for political or commercial gain. In short, a deep and persistent crisis of authenticity has been precipitated by generative AI. Citizens are left increasingly unable to trust digital content, and authentic information itself is at constant risk of being dismissed or distorted. Consequently, public opinion supervisors and media professionals must now learn to operate in an environment of chronic uncertainty, requiring a new suite of verification skills and technological tools.

2.2 Deepfakes and Synthetic Media Threats

A particularly acute and high-profile facet of the authenticity crisis is the rise of deepfakes and other forms of synthetic media. Deepfakes are hyper-realistic AI-generated videos or audio files that purport to show real people saying or doing things they never said or did. This technology moves beyond simple photo manipulation into the realm of behavioral and identity fabrication. Unprecedented and asymmetric risks to public discourse are posed by these technologies. These risks range from the targeted defamation of public figures and political candidates to the incitement of social conflict, diplomatic crises, or financial market manipulation.

By 2022–2023, the threat moved from theoretical to practical. Real-world instances began appearing with increasing frequency and sophistication. These included a widely circulated deepfake of Ukraine's president seemingly issuing a surrender order, as well as forged videos of corporate CEOs making false announcements that briefly impacted stock prices. Each incident had the potential to sway public opinion or cause real-world panic before it could be authoritatively debunked. Deepfakes exploit the powerful cognitive bias that "seeing is believing." They cause viewers to form strong, emotionally resonant memories of events that never occurred, making subsequent corrections less effective. Viewers are thus conditioned to doubt the veracity of even verified, authentic information. Beyond political and financial applications, the technology is also used to create non-consensual pornographic imagery, constituting a severe form of digital-age harassment and abuse [5].

The threat to digital authenticity is therefore twofold. On one hand, the public may be comprehensively fooled by increasingly realistic fakes. On the other hand, a "reality apathy" may set in, where even truthful, verifiable documentation of real-world events (such as human rights abuses) is casually questioned or dismissed as a potential AI fabrication. Both outcomes are profoundly damaging to the integrity of public discourse and the function of a democratic society. It has been noted that people are now increasingly unable to reliably distinguish between AI-generated and authentic media. Human perception and traditional media literacy heuristics alone can no longer be relied upon to guard against manipulated content. This escalating technological arms race between synthesis and detection places new and urgent responsibilities on educators. Students must be taught the technical-social dynamics of deepfakes, how to detect them, how to respond to them in a crisis, and how to ethically manage the profound implications of AI-altered information.

2.3 Automated Opinion Manipulation

Generative AI also serves as a powerful accelerant, supercharging the scale, sophistication, and cost-effectiveness of automated information manipulation. In the previous era of social media (dubbed "social media manipulation 2.0"), orchestrated disinformation campaigns, such as those originating from "troll farms" or social bot networks, were often limited by the human labor required to craft plausible messages. These bots were often relatively easy to detect due to repetitive, simplistic, or non-contextual behaviors. Generative AI, however, enables "social media manipulation 3.0." An avalanche of unique, context-aware, and tailored posts, comments, and even interactive conversational messages can now be produced by

AI-driven bots at a massive scale. These AI agents can flood social media platforms with particular narratives—a strategy known as agenda flooding or "computational astroturfing." This activity is designed to create a false impression of grassroots consensus, manufacturing the appearance of widespread public support or opposition to a policy, brand, or idea.

It is explained that bots armed with generative AI can effectively "blur the line between human and machine." Their ability to mimic human linguistic nuance, emotional expression, and interaction patterns makes it exceptionally challenging to discern automated accounts from real users. Public opinion can thus be swayed by systematically saturating information feeds with compelling, targeted, and seemingly "human" content. Notably, influence operations can be cheaply and efficiently scaled up by malicious actors—be they state-sponsored groups, political campaigns, or commercial entities. Recent analyses from think tanks indicate that massive networks of AI-driven "personas" could be deployed. These personas can appear eerily human, complete with generated profile pictures, fabricated personal histories, and a consistent, AI-generated posting history. Propaganda, falsehoods, or divisive rhetoric are thereby amplified with minimal human oversight. These networks leverage AI's ability to generate high-quality disinformation with extremely low marginal cost and low rates of detection [6].

The implications for the discipline of public opinion supervision are profound. Traditional methods of monitoring (such as simple keyword tracking or sentiment analysis) must be augmented with new technical tools capable of network analysis and bot detection. AI-orchestrated influence campaigns must be identified and understood not just as a collection of false posts, but as a systemic manipulation of the discourse architecture. Additionally, students preparing for this field must grasp complex concepts such as algorithmic amplification, the political economy of data, bot detection methodologies, and the difficult ethical dilemmas surrounding counter-disinformation tactics. In summary, a game-changing ability to automate and mass-produce opinion shaping has been introduced by generative AI. This new reality must be a central component of how educators train future communication professionals.

In light of these interconnected challenges, the responsibilities of public opinion supervision are both heightened and fundamentally transformed. It must be ensured by educators that students understand not only the nature of these AIGC-driven threats but also the sociotechnical context behind them. This includes how AI models are trained, how they can be misused, and what legal or technical safeguards exist. The next section outlines a competency framework developed to address these needs. The core skills and literacies needed to effectively supervise public opinion in the era of generative AI are identified.

3.Core Competencies for Public Opinion Supervision in the AIGC Era

To adequately meet the foregoing challenges, students of communications, journalism, and new media must develop a robust and integrated set of competencies. This framework builds upon foundational principles of digital and media literacy but extends them significantly to address the unique affordances of generative AI. A competency framework for the AIGC era is proposed. It consists of four critical and interrelated skill domains: (1) Technical Understanding, (2) Critical Discrimination, (3) Ethical Judgment, and (4) Human–AI Collaboration. These domains were synthesized from and align with recent recommendations from international literature on AI literacy, adapted specifically to the high-stakes context of public opinion supervision. Each competency is defined and rationalized below.

3.1 Technical Understanding of Generative AI

A foundational functional literacy in how generative AI models operate is first needed by students. This is not to suggest that communication students must become computer scientists, but that they must move beyond treating AI as a "black box." This competency includes knowing the basic principles of machine learning, neural networks, and generative models (e.g., how large language models are trained on vast datasets, and how they produce probabilistic outputs). In the specific context of public opinion, this technical understanding enables students to grasp why AI-generated content can be so convincing. They can learn the technical markers of deepfake algorithms or the "hallucination" tendencies of language models. This understanding also attunes them to the inherent limitations of these systems, such as data biases inherited from training corpora, a fundamental inability to verify factual truth, or a lack of real-world context and common-sense reasoning. For example, a student with this competency should understand that ChatGPT predicts the next most plausible word based on statistical patterns; it does not "know" or "confirm" truth, which is precisely why it may fabricate plausible-sounding

misinformation, complete with fake citations.

Technical know-how also involves a functional familiarity with the ecosystem of tools for detection. This includes an awareness of deepfake forensics, content watermarking, cryptographic origin verification (like C2PA), and AI-output detectors. Crucially, this competency also includes an awareness of the current limitations of these detection tools (e.g., the high false-positive rates of AI text detectors or the ability of adversaries to "fool" detection algorithms). By demystifying how AIGC works "under the hood," future media professionals can more effectively scrutinize AI-generated materials. They can also better explain to the public or organizational decision-makers why a given piece of content may or may not be trustworthy. This competency corresponds directly to what UNESCO's framework terms "AI foundations and applications." In the redesigned course, this was emphasized by teaching the basics of different generative models (GANs, LLMs, Diffusion) and having students conduct hands-on experiments with AI tools to observe and document their behaviors, successes, and failures.

3.2 Critical Discrimination and Analytical Thinking

In an information environment rife with sophisticated AI-generated distortions, students must dramatically hone their ability to critically evaluate information authenticity, quality, and intent. This competency builds directly upon traditional media literacy and critical thinking but places a new and urgent emphasis on detecting AI-specific signs of falsehood and manipulation. It involves a suite of advanced analytical skills. These include traditional methods like verifying sources, cross-checking facts with reliable databases, and using lateral reading strategies to assess the reputation of an information source. However, it also adds new technical heuristics, such as analyzing metadata for signs of manipulation, conducting reverse image searches, and scrutinizing visual or audio content for subtle artifacts (e.g., unnatural blinking, background warping, mismatched audio-video synchronization). Students should learn to ask pointed, hypothesis-driven questions: Does this video exhibit any known artifacts of a specific deepfake generation method? Is this online "persona" posting with superhuman frequency, suggesting automation? Is this quote traceable to a reputable source, or does it have the linguistic hallmarks of being AI-generated? Students must also become adept at recognizing the subtle biases, framing effects, or omissions that AI-generated content might introduce. Generative AI can produce content with a fluent, "human-like" rhetorical style that artfully conceals bias, unreliability, or a lack of source grounding [7].

Thus, critical discrimination includes developing a mental model of AI systems' tendencies (e.g., large language models often reflect the dominant biases in their training data or will "confidently" generate incorrect information). Educators should cultivate a "healthy skepticism" in students—a professional reflex to double-check extraordinary claims and to be acutely aware that manipulated or entirely fake content can circulate widely and achieve social validation before verification can occur. Importantly, this competency is not purely technical; it also draws deeply on the social sciences. It requires understanding how misinformation spreads through networks, the psychology of social influence, and the political or economic motives behind coordinated disinformation campaigns. By integrating these technical and social-scientific perspectives, students learn to become true analysts of complex information ecosystems, not just passive consumers or simple fact-checkers.

3.3 Ethical Judgment and Responsibility

The advent of AIGC raises novel and complex ethical questions for those who create, disseminate, or regulate public information. Students must develop a strong, principled sense of ethics and normative judgment regarding AI's use in communication. This competency is essential for maintaining public trust and mitigating societal harm. It involves a clear understanding of established ethical guidelines (e.g., journalistic principles of honesty, accuracy, transparency, and respect for privacy) and emerging legal standards (e.g., evolving laws on defamation, intellectual property, and data rights) as they pertain to AI-generated content. Students must be prepared to grapple with real-world dilemmas. For example, if a journalist uses ChatGPT to help draft an article, what are the ethical duties to fact-check the AI's output and to disclose the use of AI assistance to the audience? Or, if a public relations team deploys AI-powered chatbots to engage with social media users during a crisis, how do they ensure these bots do not mislead or deceive the public?

Students should also be guided to grapple with broader, macro-level societal ethics. This includes debates on the appropriate

balance between free expression and harm prevention in the context of deepfakes. It involves analyzing the risk of reinforcing systemic biases (e.g., racial or gender biases) through the uncritical use of AI content. It also includes the critical question of accountability: who is responsible when an AI model disseminates harmful falsehoods—the developer, the user, or the platform?

In our framework, ethical judgment includes both individual ethical usage of AI (micro-ethics) and the ability to contribute to organizational and public governance discussions on AI in media (macro-ethics). This maps to what some pedagogical frameworks call "ethical AI literacy." Practically, this competency was instilled by presenting students with real-world and hypothetical case-study dilemmas. They were guided through established frameworks of ethical decision-making (e.g., utilitarian, deontological) to analyze the stakes and stakeholders involved. By fostering a principled and reflective mindset, the aim is to produce graduates who can uphold truth and the public interest when they inevitably encounter complex AI-related challenges in their careers.

3.4 Human–AI Collaboration Skills

Finally, beyond simply defending against AI's risks, students must learn to work effectively, efficiently, and safely with AI tools. This capability transforms AIGC from a perceived threat into a potential "co-worker" or analytical assistant. This represents a form of "AI fluency" that will be a critical workplace skill.

In the public opinion supervision context, this means knowing how to strategically leverage AI for appropriate tasks. These might include content analysis (e.g., using an LLM to perform thematic analysis on thousands of social media comments), data mining (e.g., identifying emerging narratives), language translation, drafting initial reports, or generating creative multimedia for public service announcements. The key, however, is maintaining rigorous human oversight, editorial judgment, and ultimate accountability. The notion of human-AI collaboration includes a set of practical skills. "Prompt engineering"—the ability to formulate clear, effective, and nuanced queries to elicit desired outputs from AI systems—is paramount. So is the skill of critically evaluating AI outputs, rather than passively accepting them at face value. This involves integrating AI assistance into professional workflows in a manner that is both transparent and accountable [8].

Students should be trained to recognize where AI can provide a significant advantage—e.g., using natural language generation to summarize large volumes of data quickly—but also, critically, where human expertise is irreplaceable. This includes tasks requiring deep contextual understanding, nuanced cultural interpretation, subjective judgment, and ethical discernment. Developing this competency prepares students for a workplace that increasingly expects proficiency with a wide array of AI tools. It also aligns with calls for a new "rhetorical literacy" in the age of AI—i.e., the skill of using AI-generated language strategically and responsibly to achieve communication goals. In our teaching, human-AI collaboration was incorporated by assigning project work where students were required to use generative AI as a creative aid or research assistant, and then critically evaluate its contributions in a meta-reflective report.

In summary, this comprehensive competency framework for the AIGC era spans technical, critical, ethical, and collaborative skill sets. These domains are designed to be mutually reinforcing: technical knowledge aids critical analysis; ethical principles must guide the use of technical tools; and human-AI collaboration can only flourish when it is underpinned by rigorous critical oversight and high ethical standards.

4.Teaching Innovations in "Online Communication and Public Opinion Supervision"

LeBron To translate the four-domain competency framework from theory into practice, the course "Online Communication and Public Opinion Supervision" underwent a significant pedagogical redesign. A traditional lecture-and-exam format was deemed insufficient. It was replaced with a blended, active-learning model. A variety of teaching innovations were introduced, combining foundational theoretical learning with intensive, hands-on practice. The key innovations are detailed below.

4.1 Curriculum Update - AIGC and Public Opinion Module

The most fundamental innovation was the introduction of a dedicated, multi-week module on "Generative AI in the Information Ecosystem." This module was strategically placed early in the semester to provide foundational context for all subsequent activities. This module was explicitly interdisciplinary, bridging communication theory with basic computer science concepts.

The module delivery included:

Technical Primers: Lectures were designed to demystify AIGC. They covered the basic principles of generative models (e.g., Generative Adversarial Networks for images, transformers for language), the role of training data, and the concept of probabilistic generation. This was crucial for building the "Technical Understanding" competency.

Real-World Case Studies: The module was heavily case-driven. Students were required to analyze recent, high-profile incidents where AIGC had a measurable impact on public opinion. Both international and domestic cases were examined, for example, AI-generated fake images of political events, domestic viral rumors traced to AI-generated audio, and the "Pentagon explosion" fake image incident. For each case, students analyzed the full lifecycle: the content's creation, its amplification vectors (e.g., social media algorithms, influential accounts), the public's reaction, and the subsequent debunking and verification efforts.

Countermeasure Landscape: Embedding such up-to-date cases made the learning authentic and urgent. To avoid a purely dystopian framing, this module also covered the landscape of emerging countermeasures. This included discussions on deepfake detection technologies (and their limitations), digital content authentication standards (like C2PA), and regulatory policies from different national contexts (e.g., China, EU, US).

This foundational knowledge, covering the threat, the technology, and the response, fed directly into the subsequent practical and simulation-based activities.

4.2 "Red Team vs. Blue Team" Simulation Exercises

To move from passive knowledge to active skill, inspiration was drawn from cybersecurity training and professional wargaming. Multi-day simulation exercises were implemented where the class was split into opposing teams to role-play a live information contest. This pedagogy mirrors the concept of treating disinformation as an adversarial, dynamic conflict.

The simulation was structured in three phases:

Phase 1: Preparation. The class was divided into a "Red Team" (attackers) and a "Blue Team" (defenders). The Red Team's objective was to create and strategically disseminate a piece of AI-generated disinformation on a pre-approved, low-stakes topic (e.g., a fabricated campus policy, a false local event). They were required to use publicly available AI tools. The Blue Team's objective was to establish a "public opinion monitoring center" and develop a protocol for detecting and responding to falsehoods.

Phase 2: Execution. Over a 48-hour period, the Red Team "released" its content into a closed digital environment (e.g., a private class forum or social media group). The Blue Team was tasked with actively monitoring these channels, aiming to be the first to detect the fake, verify its falsehood, issue a "public clarification," and trace its origin.

Phase 3: Debrief. This was the most critical phase. A mandatory, two-hour reflective debrief session was held. Students stepped out of their roles and presented their strategies. The Red Team detailed their creation process and dissemination tactics. The Blue Team presented their detection workflow, including any false positives or misses.

Such "red vs. blue" scenarios created a competitive yet highly educational dynamic. Students on the Red Team gained visceral insight into how easily and quickly convincing false content can be produced. Students on the Blue Team practiced critical analysis and crisis response under time pressure. This single exercise powerfully reinforced technical skills (using AI, using detection tools) and critical skills (spotting fakes). It also surfaced complex ethical discussions. Many Red Team members reported feeling "uneasy" or "guilty" about crafting convincing lies, which led to a profound class conversation on the ethics of information warfare and the responsibilities of communicators. The simulation proved to be a high-impact pedagogy, making abstract threats concrete and memorable.

4.3 Project-Based Assignments with AI Integration

To foster the "Human–AI Collaboration" competency, several major course assignments were redesigned to require the ethical and critical integration of AI tools. This was structured within a project-based learning (PBL) format. One major project was an "AI-Augmented Media Analysis Report." Student teams selected a recent, real-world public opinion event (e.g., a corporate PR crisis, a viral social movement). Their task was to write an analytical report, but they were mandated to use generative AI tools to assist in at least two phases of their research. For instance, they might prompt a chatbot to:

Summarize hundreds of social media comments to identify key themes.

Draft an initial literature review on the topic.

Analyze a set of images for signs of manipulation.

Generate code for a simple data visualization.

The crucial part of the assignment was not the final report itself, but a mandatory 2-page "AI Collaboration Appendix." In this appendix, students had to document every prompt they used, the raw output the AI provided, and a critical evaluation of that output. They had to "grade" the AI's contribution, identify any inaccuracies or biases it introduced, and describe how they corrected or refined its work. The AI, in essence, became a "team member" subject to human supervision and performance review. Another assignment had students create a public service announcement (PSA) on digital literacy in the AI era. They were encouraged to use text generation tools to brainstorm scripts or slogans and image generation tools for visuals. However, they were graded on how well they curated, edited, and fact-checked the AI outputs to align with factual correctness, ethical guidelines, and an appropriate public-service tone. This project-based use of AI kept engagement high. It moved students from theoretical awareness of AI to practical, hands-on fluency, ensuring they learned with AI in a critical and meaningful way.

4.4 Oral "Deepfake Crisis" Response Drills

Public opinion supervision and crisis communication often involve high-pressure, real-time decision-making. To build competency in such scenarios, the course introduced a series of oral crisis response drills. These were, in effect, simulations of a press briefing or internal executive Q&A in the immediate aftermath of a viral AI-driven rumor or deepfake. A typical drill scenario was presented as follows: "It is 8:00 AM. Overnight, a deepfake video of your organization's CEO making discriminatory remarks has gone viral on all major platforms. The media is calling, and employees are panicking. As the head of communications, you must deliver an impromptu 3-minute statement to the press and then answer 5 minutes of tough questions."

Students took turns in the "hot seat" at the front of the class, role-playing as the spokesperson. Their classmates role-played as aggressive journalists or concerned citizens, asking difficult questions. Students had minimal time to prepare, mirroring the reality of how fast-breaking digital crises unfold.

These drills served multiple educational purposes:

Cognitive: They forced students to synthesize information quickly (What do we know? What can we not say? What is the core message?). This reinforced critical discrimination skills under severe time constraints.

Performative: They provided practice in strategic crisis communication—learning to calmly convey facts, authoritatively refute the fake without over-amplifying it, and pivot to messages of public trust and procedural integrity.

Procedural: The exercise highlighted the absolute importance of preparation and protocols. Students quickly realized the value of having pre-established crisis response plans specifically for dealing with AI-driven forgeries, echoing emerging industry best practices.

Feedback indicated these drills were initially intimidating but ultimately the most confidence-building part of the course. They viscerally connected theory (e.g., "deepfake threats") to practice (e.g., "What do I do?"). Collectively, these four innovations—the updated module, the adversarial simulation, the AI-integrated projects, and the live crisis drills—restructured the course into an active, experiential learning environment explicitly geared towards the challenges of the AIGC era.

5.Empirical Outcomes: Classroom Observations and Student Feedback

Player To gauge the effectiveness of these pedagogical innovations, a mixed-methods approach was used. Data were collected throughout the semester, including observational data during exercises, analysis of final project quality, and formal qualitative feedback from students via end-of-term surveys and reflective essays. Although this case study was not a controlled experimental study, the before-and-after comparisons (against previous iterations of the course) and the richness of the qualitative insights provide strong evidence of notable improvements in student competencies and engagement. The key findings from the case implementation are summarized below.

5.1 Enhanced Analytical Performance and AI Literacy

By the conclusion of the course, students demonstrated a marked and measurable improvement in their ability to identify, analyze, and deconstruct AI-generated content. One key metric came from a practical final exam. Students were given a "dossier" containing a set of mixed-media news items (a mix of genuine news articles, AI-fabricated articles, a lightly edited deepfake image, and an AI-generated audio clip). They were tasked to determine which items were suspect and provide a detailed justification for their reasoning.

In the initial weeks, a similar diagnostic quiz showed low accuracy; students primarily relied on "gut feeling." On the final exam, a substantial gain in critical discrimination skills was evident. A large majority of students correctly identified the deepfake image and the AI-generated news story. More importantly, their justifications were no longer vague. They used specific terminology learned in the course, citing "unnatural linguistic patterns and lack of verifiable sources" for the AI text, and "background warping and unnatural eye movement" for the deepfake.

Observationally, during the "Red vs. Blue Team" simulations, the Blue Teams became progressively more adept at using systematic verification tools and cross-referencing information. This demonstrated an internalization of the "critical discrimination" competency. Furthermore, in an AI literacy survey administered at the semester's end, a large majority of students agreed with the statement "I understand how generative AI produces content and its potential inaccuracies," a significant increase from the start of the semester. This outcome supports the hypothesis that active, problem-based learning interventions can significantly boost functional and critical AI literacy among communication students [9].

5.2 Growth in Ethical Reasoning and Judgment

Another significant observed change was in the sophistication and nuance of students' discussions surrounding media ethics. In final reflective essays, many students described a fundamental transformation in how they view their future role as media professionals. Initially, some expressed a narrow, technical view of public opinion work (e.g., "posting news quickly"). By the course end, there was a widespread and articulated acknowledgement of a profound duty to verify information, consider societal impact, and act as a guardian of public trust before disseminating content. This reflects a clear ethical maturation. During in-class debates on hypothetical policies (e.g., "Should all AI-generated content be legally required to carry a watermark?"), students engaged with complex nuances. They moved beyond simple "yes" or "no" answers to weigh concepts like free speech against harm prevention, demonstrating familiarity with ideas like the "liar's dividend" and the importance of maintaining public trust.

One student's reflection was particularly telling: "I realize now that rushing to break a story that might be a deepfake could do real damage, even if it brings clicks. As communicators, we must act as guardians of truth, even if that means delaying publication until verification is complete." This statement reflects a deep internalization of the "ethical judgment" principles. It is noteworthy that students who served on the "Red Teams" (and actively created fake content) had some of the most poignant ethical reflections. They often noted how uncomfortably easy it was to fool their peers and thus how great the responsibility is to combat such deception. These insights suggest that the course's experiential methods effectively sensitized students to the high ethical stakes of AIGC.

5.3 Increased Engagement and Confidence (Affective Outcomes)

The introduction of interactive, relevant, and challenging activities had a highly positive and measurable effect on student engagement. Class attendance and participation in discussions were significantly higher than in previous, lecture-based offerings of the course. The "Red vs. Blue" team games, in particular, injected a palpable energy into the classroom. Students who were usually quiet in traditional lectures became highly involved in strategizing, sleuthing, and debating during the simulations. In end-of-term feedback surveys, the course was rated as "highly engaging" by a significant majority of students. They frequently mentioned that the "realism" and "novelty" of the tasks (e.g., tackling deepfakes, using cutting-edge AI tools) made them eager to come to class and participate. Many commented that the simulations and drills were "challenging but fun" and that they appreciated the chance to apply theory to "feel like a real investigator or spokesperson."

This engagement is valuable because research has shown that students who feel involved and interested tend to achieve deeper learning outcomes. Additionally, student confidence—or self-efficacy—improved dramatically. In the initial class,

a majority reported feeling “unprepared” or “very unprepared” to deal with AI in their future jobs. By the final class, a vast majority of students stated they felt “prepared” or “very prepared” to handle scenarios involving AI-generated content. This boost in self-efficacy is a critical affective outcome. One student noted that after successfully writing a press release to debunk a fake video in a drill, she felt “much more confident that I could do this in a real job.” Another said the course changed their mindset from fearing AI as a job-destroying threat to viewing it as “an issue I know how to manage—and even a tool I can use.”

5.4 Feedback on Specific Pedagogical Methods

Qualitative feedback from students underscored which innovations they found most effective.

The “Red Team vs. Blue Team” simulations received nearly unanimous enthusiastic praise. Students felt these exercises gave them “practical skills in a memorable way” that a lecture never could. The “AI-Augmented Project” also received positive remarks. Students stated it taught them how to use tools like ChatGPT productively (e.g., for brainstorming or summarizing) while also—and more importantly—teaching them not to trust the tool blindly. Several admitted that in their first attempts, they leaned too heavily on the AI and their analysis was shallow; the requirement to “show their work” in the appendix forced them to become more critical [10].

The “Crisis Response Drills” were consistently cited as the most “challenging” or “nerve-wracking” component, but many acknowledged that this was precisely why it was valuable. They learned the importance of preparation and “thinking on your feet.” One constructive critique received was to incorporate more collaborative oral drills, so students could work as a team to handle a crisis, rather than always being in the “hot seat” alone. This is a suggestion being considered for future iterations.

6.Discussion and Limitations

The aggregate empirical outcomes from this case study are highly encouraging. Students demonstrated objective improvements in skill (analysis accuracy) and subjective growth in confidence, engagement, and ethical awareness. These findings support broader educational claims that innovative, experiential, and problem-based pedagogies are particularly effective for teaching complex socio-technical topics. By simulating the pressures and complexities of real-world public opinion supervision in the AI age, the course helped students build both the “cognitive muscle memory” and the reflective “ethical mindset” needed to act competently and responsibly.

The success of the “Red Team vs. Blue Team” model, in particular, suggests that adversarial learning can be a powerful tool for media education, just as it is in cybersecurity. It forces students to move beyond a passive “receiver” role and become an active participant in the information contest, leading to deeper, more lasting understanding. The “Human-AI Collaboration” projects also proved effective, striking a balance between embracing new tools and instilling critical oversight. This addresses a key anxiety for many students: how to use AI without “cheating” or becoming over-reliant. By making the use of AI explicit, mandatory, and reflective, the course normalized it as a professional tool that, like any tool, requires skill and ethics to wield.

However, the limitations of this study must be acknowledged [11].

Scope and Scale: The findings are from a single case study, implemented in one course with a relatively small cohort (approximately 30 students). The results may not be generalizable to all university contexts, student populations, or different course subjects.

Methodology: The “empirical” data rely heavily on self-reported feedback, instructor observation, and qualitative analysis rather than rigorous, quantitative experimental evaluation. There was no control group (e.g., a class taking a traditional lecture-based version of the course) for direct comparison.

Long-Term Retention: This study measured outcomes at the end of the semester. It did not (and could not) track the long-term retention of these skills or how students will apply them in their professional careers [12].

Rapidly Evolving Tech: The course was designed around AIGC tools and threats from 2023-2024. The technology is evolving at an exponential rate. The specific technical skills taught (e.g., how to spot a certain type of artifact) may become obsolete quickly.

Future work is needed to address these limitations. More formal assessment (e.g., pre/post standardized tests of AI literacy)

and comparative studies with control groups would strengthen the evidence base. Nonetheless, as a practical, pedagogical exploration, this case study offers tangible and useful insights for educators seeking to modernize their teaching approaches in a rapidly changing world.

Conclusion and Future Directions

The rapid, disruptive rise of generative AI is fundamentally reshaping the landscape of public communication. It poses new and profound challenges to the supervision of public opinion, the integrity of information, and the stability of public trust. This paper examined those challenges and presented a detailed case study of pedagogical innovation designed in response. Core competencies for media professionals in the AIGC era were identified: technical understanding, critical discrimination, ethical judgment, and human-AI collaboration [13].

To foster these competencies, the course "Online Communication and Public Opinion Supervision" was comprehensively redesigned. The new curriculum featured a mix of theoretical updates and high-impact experiential learning activities, including modules on AI and misinformation, "red/blue team" disinformation wargames, AI-augmented project-based learning, and simulated crisis response drills. The reported outcomes, based on observational and qualitative data, show that such interventions can significantly improve students' analytical skills, AI literacy, ethical awareness, and professional confidence. Students left the course better equipped to distinguish truth from sophisticated fabrication, to use powerful AI tools judiciously, and to act as responsible, ethical stewards of the information ecosystem.

Several key insights emerge from this case study for educators in related fields.

Active Learning is Essential: Simply warning students about deepfakes in a lecture is insufficient. Engaging them in the process of creating and debunking fake content, as in the adversarial simulation, proved far more impactful. This aligns with constructivist learning theory.

Integrate Ethics Everywhere: Ethics cannot be relegated to a single "Week 10: Ethics" module. In our course, every hands-on activity was paired with a mandatory ethical debrief, helping students connect technical actions with moral principles.

Embrace, Don't Ban, AI Tools: The "AI-Augmented Project" model provides a framework for teaching with AI, rather than attempting to ban it. It cultivates the critical collaboration skills that employers will demand.

Agility is Key: The success of the course was partly due to its timeliness. This suggests that faculty development and curriculum design processes need to become more agile to keep pace with technological change.

Looking ahead, several future directions can be pursued. From a teaching perspective, the interdisciplinary scope of such courses could be expanded. Collaborative modules involving computer science students (demonstrating algorithm design) and communication students (analyzing impact) could be mutually enriching. International case comparisons, analyzing how different media ecosystems and regulatory environments (e.g., China, EU, US) are responding to AIGC, can broaden students' global perspectives.

In terms of research, this pedagogical case study could be extended with more rigorous assessment. Researchers could formally measure how much a simulation improves detection skills versus a control lecture, or track the long-term retention of these skills into graduates' careers. There is also room to investigate the affective dimension further: how do exercises dealing with AI-driven deception affect student attitudes like cynicism or vigilance months later?

In conclusion, generative AI undeniably poses serious challenges to the authenticity and management of public opinion. But with thoughtful curriculum design and a steadfast commitment to competency-based, experiential education, we can prepare the next generation of professionals to meet these challenges. The experience from this course demonstrates that students, when given the right tools, training, and learning opportunities, can indeed rise to the occasion. They can become savvy analysts, ethical communicators, and innovative collaborators, qualities that will be indispensable for safeguarding the information ecosystem.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Hagos, D. H., Battle, R., & Rawat, D. B. (2024). Recent advances in generative ai and large language models: Current status, challenges, and perspectives. IEEE Transactions on Artificial Intelligence.

[2] Brennan, K. M. (2019). “Believe me”: Authenticity, federal social media use, and the problematized record in the American digital public sphere. Journal of Critical Library and Information Studies, 2(2).

[3] Ling, J., Huang, N., Ma, Z., et al. (2025). Investigating student behaviors in multimodal AIGC-enhanced design education: An HCI-based innovative learning model. In Proceedings of the International Conference on Human-Computer Interaction (pp. 240 - 255). Cham: Springer Nature Switzerland.

[4] Van Dyk, S. (2022). Post-truth, the future of democracy and the public sphere. Theory, Culture & Society, 39(4), 37 - 50.

[5] Maddocks, S. (2018). From non-consensual pornography to image-based sexual abuse: Charting the course of a problem with many names. Australian Feminist Studies, 33(97), 345 - 361.

[6] Olakoyenikan, O., & Olanipekun, S. O. Leveraging artificial intelligence to combat disinformation and deepfakes in broadcast journalism and global media communications. [Journal Name]. (Note: The original reference lacks key information such as publication year, journal volume, issue and page numbers; it is recommended to supplement it according to the official publication source.)

[7] Carpenter, P. (2024). FAIK: A practical guide to living in a world of deepfakes, disinformation, and AI-generated deceptions. John Wiley & Sons.

[8] Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. Public Administration Review, 81(5), 825 - 836.

[9] Huang, S. H., & Su, T. C. (2024). Learning effectiveness and reflections on AI literacy in junior high school students with game-based learning and problem-based learning. In Proceedings of the 32nd International Conference on Computers in Education. Quezon City, Philippines: Ateneo de Manila University.

[10] Schrepel, T. (2025). Generative AI in legal education: A two-year experiment with ChatGPT. Legal Education, [Volume(Issue)], [Pages]. (Note: The original reference has a redundant title repetition and lacks specific volume, issue and page information; it needs to be supplemented according to the official published version.)

[11] Clear, T., Cajander, Å., Clear, A., et al. (2025). AI integration in the IT professional workplace: A scoping review and interview study with implications for education and professional competencies. In 2024 Working Group Reports on Innovation and Technology in Computer Science Education (pp. 34 - 67). ACM.

[12] Bernold, L. E., Spurlin, J. E., & Anson, C. M. (2007). Understanding our students: A longitudinal study of success and failure in engineering with implications for increased retention. Journal of Engineering Education, 96(3), 263 - 274.

[13] Hu, Z. Y. (2024). AIGC related context: A new communication culture for human. Journal of Literature and Art Studies, 14(10), 921 - 931.