Aruzhan Muratbek

The Influence of AI-Generated Imagery on Credibility Perceptions in Digital Media

April 29, 2026

Introduction

Topic of Research: The impact of AI-generated visual content on credibility judgments in digital media.

The impact of artificial intelligence-generated visual content on audience trust and credibility judgments in digital media environments represents an increasingly urgent area of inquiry as synthetic media becomes ubiquitous in online information ecosystems.

As AI image generation becomes increasingly accessible and visually sophisticated, questions arise about how audiences perceive the credibility of media content when accompanied by AI-generated versus human-created imagery. This study examines whether the type of image presented alongside a media claim - human-generated, AI-generated, or no image - affects young people's perception of that claim's credibility. Understanding this relationship has important implications for journalism ethics, visual literacy education, and audience trust in an era where synthetic media is becoming commonplace.

The proliferation of AI-generated images in news and social media presents unprecedented challenges for media credibility. Tools like DALL-E, Midjourney, and Stable Diffusion have democratized image creation, allowing anyone to produce photorealistic content within seconds. While this technology offers creative possibilities, it also raises concerns about misinformation, manipulation, and the erosion of visual evidence as a trusted form of proof. Recent research suggests that audiences struggle to distinguish AI-generated images from authentic photographs, yet little experimental work has examined whether knowledge of an image's AI origin affects credibility judgments. This multi-method study addresses that gap by isolating image type as a variable and measuring its effect on perceived credibility. The study reveals that disclosure significantly affects credibility judgments even when AI-generated images achieve high visual realism. Journalism students rated claims accompanied by labeled AI-generated images as substantially less credible than identical claims with human photographs, though any image

increased credibility compared to text alone. Qualitative findings indicate students set clear ethical boundaries around AI imagery, viewing it as acceptable for illustration but inappropriate as evidence.

Literature Review

Photographs have long served as proof in journalism and media because they capture real moments. AI-generated images challenge this fundamental trust because they are created from scratch by algorithms, with no real-world scene behind them, even though they can look completely realistic.

Velásquez-Salamanca, Martín-Pascual, and Andreu-Sánchez (2025) conducted an experimental study examining whether image origin influences interpretation and trust. Their study involved 161 participants who evaluated 24 AI-generated images and 8 human-made photographs, with measurements focusing on source identification, perceived realism, and perceived credibility. Results demonstrated that participants correctly identified real photographs 78 percent of the time compared to only 61 percent accuracy for AI-generated images, meaning approximately 39 percent of synthetic images were mistaken for authentic content. In terms of ratings, human-made images scored significantly higher on both perceived realism (4.22 versus 3.58 on a five-point scale) and credibility (4.19 versus 3.53). These findings suggest that awareness of synthetic production reduces trust even when images achieve visual realism.

The researchers examined five different AI image generation models and found substantial variation in detection rates. Images from the Kolors model were correctly identified as AI-generated 87 percent of the time, while the advanced FLUX.1-dev model achieved only 29 percent detection accuracy. Visual professionals outperformed non-professionals at identifying AI images, but these modest differences indicate that expertise provides limited advantage in distinguishing synthetic from authentic imagery. Participants frequently demonstrated high confidence in incorrect classifications, particularly with FLUX.1-dev images. The researchers termed this

phenomenon "epistemic vulnerability" - the danger of being confidently wrong without recognizing the need for greater caution. When asked to explain their decision-making processes, participants cited indicators such as texture irregularities, unnatural lighting, and implausible object placement. However, for highly advanced AI models, participants increasingly relied on intuitive assessments, describing images as appearing "too perfect" or eliciting an "uncanny" response without being able to articulate specific technical flaws.

These findings support **Hypothesis 1**: Media claims accompanied by AI-generated images will be perceived as less credible than those with human-generated photographs when image type is explicitly disclosed. This expectation rests on the assumption that audiences associate AI images with disconnection from authentic events. However, given the varying performance across AI models, this research also raises **Research Question 1**: Can sufficiently realistic AI images perform comparably to authentic photographs despite their synthetic origin?

Högemann, Betke, and Thomas (2025) examined 104 media professionals in Germany who evaluated 50 images - 25 real photographs and 25 AI-generated images across five different models. Overall detection accuracy reached 63.7 percent, which exceeds chance but indicates substantial error rates. The advanced FLUX.1-dev model again proved particularly deceptive, with correct identification occurring in only 29 percent of cases. A notable finding concerned age effects. Older participants demonstrated progressively worse detection performance, with approximately 12 percent accuracy reduction per age category. Other demographic factors including education level, gender, and media experience showed no significant effects on detection ability. Interestingly, participants with regular AI tool experience exhibited poorer calibration between confidence and accuracy. Regular AI users demonstrated higher miscalibration scores (0.272) compared to non-users (0.207), suggesting that familiarity with AI technology may generate overconfidence.

This pattern gives rise to **Research Question 2**: How does self-reported familiarity with AI technology influence credibility judgments among students? Will students with greater AI tool experience evaluate AI-generated images differently than those with limited exposure? Will they demonstrate heightened skepticism or, conversely, greater acceptance of synthetic content?

Bohm (2024) recruited 172 German media professionals to evaluate six images, three real and three AI-generated, across multiple quality dimensions including texture and detail, color harmony, composition and structure, creativity and originality, emotional impact, and narrative strength. Real images received higher overall quality ratings, yet this quality perception advantage did not translate to accurate identification. Only one real image was correctly classified by the majority of participants - a highly detailed photograph of a parrot. Participants relied predominantly on technical characteristics such as texture quality and color composition when making classification decisions, rather than subjective dimensions like creativity or emotional resonance. Approximately one-third of participants expressed uncertainty about AI's broader implications for media, and attitudinal positions regarding AI showed no correlation with detection performance. Neither photo editing experience nor AI tool usage improved classification accuracy, suggesting that even professionals may overestimate their capacity to identify AI-generated content.

Osińska et al. (2025) employed eye-tracking methodology to examine how individuals visually process and evaluate images. They presented 36 participants with 24 images - half AI-generated using GPT-4o and half authentic photographs - spanning six content categories including faces, pets, landscapes, architecture, vehicles, and art. Eye-tracking data revealed that participants spent approximately 20 milliseconds longer fixating on AI-generated images, suggesting increased cognitive processing demands. Certain image categories proved easier to classify than others. AI-generated pet images and real architectural photographs were relatively straightforward to identify, while art proved most challenging. The study identified significant gender differences, with male participants exhibiting viewing patterns characteristic of visual experts regardless of professional

training. Participants with digital imaging knowledge demonstrated similar efficient viewing patterns. While participants could articulate specific detection cues including lighting inconsistencies, unnatural color palettes, and texture smoothing, their confidence levels frequently diverged from actual performance, particularly for advanced AI models.

Long et al. (2024) investigated a distinct context - AI-generated profile pictures in social media environments. Their study recruited 817 U.S. adults who evaluated what they believed to be social media profiles, with the researchers manipulating both the actual generation method (AI using the Remini app versus professional photography) and disclosure status (whether participants were informed about AI generation). Results indicated that AI-generated profile pictures received higher quality ratings than professional photographs when generation method remained undisclosed. When participants were informed that an image was AI-generated, quality ratings declined but retained an advantage over professional photography. Critically, neither actual AI generation nor disclosure of AI generation significantly influenced perceptions of the depicted individual's attractiveness or trustworthiness. These findings suggest that while audiences demonstrate some skepticism toward explicitly labeled AI imagery, contemporary generative AI tools achieve sufficient quality that the images themselves do not inherently undermine credibility of the depicted subject. Although this study focused on profile pictures, it demonstrates that disclosure effects operate without fundamentally transforming evaluative judgments.

Across these studies, a consistent pattern emerges: human-made images generally receive higher credibility ratings than AI-generated images when source attribution is known. This body of research demonstrates that knowing something is AI-generated affects trust independently of visual quality. However, experimental work directly comparing identical visual content in journalism contexts with explicit disclosure remains limited. The study addresses this gap through a controlled experiment which isolates image type as the sole variable.

Based on established findings regarding the persuasive effect of visual evidence, this study proposes **Hypothesis 2**: Media claims accompanied by any image, whether human-generated or AI-generated, will be perceived as more credible than text-only claims. Images serve as powerful persuasive tools, and even AI-generated images may benefit from this general enhancement effect. However, a competing possibility exists in **Hypothesis 2a**: AI-generated images may not differ significantly from the no-image condition in credibility ratings if audience skepticism toward synthetic imagery counteracts the typical persuasive advantage that visual content provides.

Methods

This study employed four complementary research methods: a content analysis of news coverage, an experimental survey, a semi-structured interview, and a focus group discussion. Each method addressed different aspects of how AI-generated imagery affects credibility perceptions in digital media.

Experiment

The data source consisted of participants' self-reported credibility judgments collected through structured survey responses. Participants provided ratings of believability, trustworthiness, credibility, and perceived accuracy for media content. Demographic data and AI familiarity information were also collected.

The study used a between-subjects experimental design with random assignment to one of three conditions. Participants were 18 sophomore journalism students at Northwestern University in Qatar. The experiment was conducted online using Google Forms in March 2026. Three separate survey links were distributed in rotating order to ensure approximately equal distribution across conditions. Each participant completed the survey once, receiving only one version. The survey took approximately five to seven minutes to complete.

The final sample included 18 participants: text-only condition (n=8), human photograph condition (n=5), and AI-generated image condition (n=5). This sample is not representative of the general population, limiting generalizability. However, journalism students represent an informed audience with media literacy training.

The independent variable, image type, was operationalized as a categorical variable with three levels. The text-only condition presented a claim about social media and mental health with no image. The human photograph condition used a licensed photograph from Pexels. The AI-generated image condition used content created with DALL-E 3 via Bing Image Creator, depicting a similar scene with matching composition. Images were explicitly labeled, with the AI image marked "AI-Generated Image." The textual claim remained identical across all conditions.

The dependent variable, perceived credibility, was operationalized using a composite credibility scale with four items on 5-point Likert scales: believability, trustworthiness, credibility, and accuracy. Responses ranged from 1 (strongly disagree) to 5 (strongly agree). The four items were averaged to create a composite credibility score. Secondary measures included behavioral intention to share content and AI familiarity, which assessed exposure to AI-generated images, use of AI tools, and confidence in identifying AI imagery.

Analysis focused on descriptive statistics and exploratory comparisons. Average credibility scores were calculated for each condition and compared using bar graphs. Given the small sample size, findings are exploratory and require replication with larger populations.

Content Analysis.

The content analysis examined how news media outlets frame AI-generated imagery and its implications for visual credibility. This method addressed four research questions: what dominant frames characterize AI imagery coverage (RQ1), which expert sources are most frequently cited (RQ2), what specific credibility concerns receive emphasis (RQ3), and what solutions are proposed (RQ4). Based on journalism's tendency to emphasize conflict and negative

consequences, three hypotheses predicted coverage patterns: threat frames would dominate over opportunity frames (H1), misinformation concerns would receive more emphasis than other credibility issues (H2), and fact-checkers would be cited more frequently than technology company representatives (H3).

The population consisted of English-language news articles published between January 2023 and April 2026 that discussed AI-generated imagery, synthetic media, or text-to-image generation technologies. The sample included 20 articles drawn through purposive sampling from different media outlets. Outlets were selected for their substantial digital reach and coverage of technology topics relevant to young adults. Articles were identified through keyword searches for terms including "AI-generated images," "synthetic media," "deepfakes," "DALL-E," "Midjourney," and "Stable Diffusion." Articles were included if they discussed AI imagery in at least three paragraphs and addressed credibility or authenticity concerns.

The unit of analysis was the individual article. A deductive codebook was developed based on framing theory and prior AI imagery research. Variables included dominant frame (threat/risk, opportunity, or balanced), overall tone (alarmist, concerned, neutral, or optimistic), primary risk emphasis (misinformation, trust erosion, copyright, or other), expert sources cited (tech companies, researchers, fact-checkers, artists, journalists), and solutions proposed (disclosure, detection tools, regulation). Each article received a complete read before systematic coding. Source citations were tallied by counting each distinct named individual or organization.

Interview

A semi-structured interview was conducted with a sophomore journalism student at Northwestern University in Qatar, in March 2026. The interview lasted 30 minutes and was conducted remotely via Zoom with audio recording enabled for transcription purposes.

The interview protocol included 11 open-ended questions progressing from general news consumption habits to specific AI imagery evaluation processes. Questions explored how the

participant encounters AI-generated images, what credibility concerns arise, how they distinguish between authentic and synthetic imagery, and what ethical boundaries they perceive around AI use in journalism.

The interview was transcribed verbatim and analyzed thematically. Key themes included the distinction between evidence and illustration, the importance of transparency through labeling, concerns about deception and manipulation of vulnerable populations, skepticism triggered by AI use in hard news contexts, and recognition that detection difficulty increases as technology advances.

Focus Group

A focus group discussion was conducted with five sophomore journalism students at Northwestern University in Qatar in March 2026. The session lasted 30 minutes and was conducted remotely via Zoom with video enabled and audio recorded for transcription.

The discussion protocol included nine questions structured to facilitate peer-to-peer dialogue. The session began with an icebreaker asking participants to describe AI images in news using one word, establishing initial reactions. Opening questions explored frequency of AI image encounters and specific examples participants had seen. The central discussion involved a visual comparison exercise where participants viewed two wildfire images side-by-side and discussed which appeared more credible, which they believed to be AI-generated, what visual cues informed their judgments, and whether explicit labeling would change their perceptions. Closing questions addressed implications for journalism education and remaining concerns about AI imagery in news.

The focus group was transcribed verbatim and analyzed for consensus points, points of debate, and emergent themes. Key themes included immediate skepticism triggered by AI imagery ("fake," "deceiving," "not trustworthy"), difficulty distinguishing realistic AI content from authentic photographs without labels, and the judgment that context determines acceptability.

Findings

Findings: Experiment

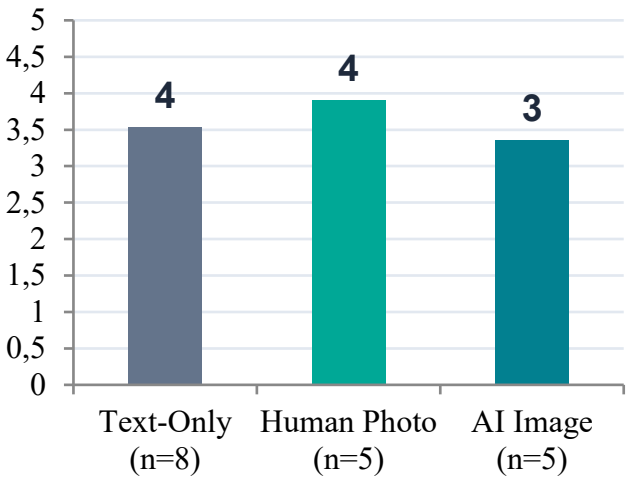

Figure 1.

Condition	n	Mean	SD
Text-Only	8	3.53	0.59
Human Photo	5	3.90	0.58
AI Image	5	3.35	1.02

Table 1.

The experiment tested how image type affects perceived credibility when accompanying factual claims. As shown in Figure 1 and Table 1, participants in the human photograph condition assigned the highest credibility ratings (Mean (M)=3.90, Standard Deviation (SD)=0.58), followed by text-only (M=3.53, SD=0.59), and AI-generated images (M=3.35, SD=1.02). The AI condition also showed notably higher variability in responses, indicating less consensus among participants.

H1 predicted that AI images would be perceived as less credible than human photographs when disclosed. The data support this hypothesis, with AI images rated 0.55 points lower than human photographs on the 5-point scale. H2 predicted that any image would be more credible than text-only content. This hypothesis received partial support. Human photographs exceeded the text-only baseline by 0.37 points, confirming the expected visual advantage. However, AI images fell 0.18 points below text-only, contradicting the predicted enhancement effect. The competing hypothesis H2a proposed that AI images might not differ from text-only if skepticism counteracts visual advantages. The data reject this hypothesis, as disclosed AI images performed worse than text-only rather than equivalently.

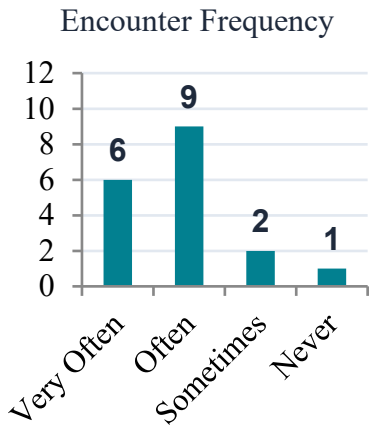

Figure 2.

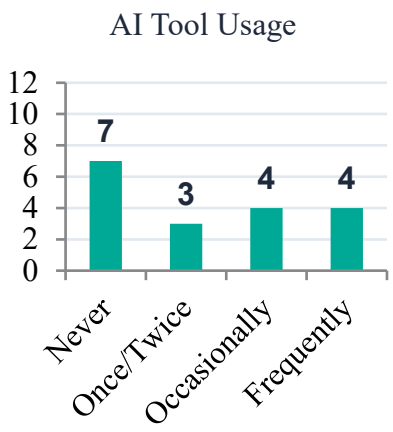

Figure 3.

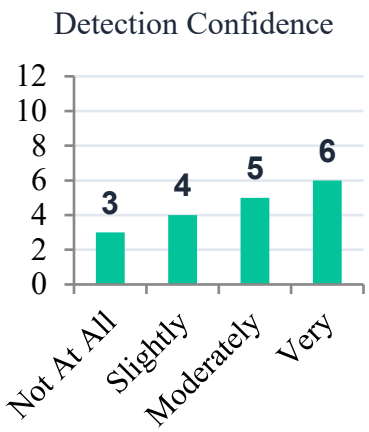

Figure 4.

RQ1 asked whether realistic AI images can perform comparably to authentic photographs. The findings indicate they cannot. Even when AI images achieve visual realism matching human photographs, explicit disclosure produces a credibility penalty. The 0.55-point gap demonstrates that audiences differentiate based on generation method than visual quality alone. RQ2 examined whether AI familiarity moderates credibility judgments. Figures 2, 3, and 4 show that 83% of participants encounter AI images daily or more, 61% have used AI tools, and 61% feel confident detecting AI imagery. Despite this high familiarity, credibility ratings followed consistent patterns regardless of individual AI exposure or detection confidence. Participants with frequent AI tool experience rated AI images similarly to those with no experience, suggesting that disclosure effects operate independently of individual familiarity with AI technology.

Findings: Content Analysis

The content analysis of 20 news articles examined how media outlets frame AI-generated imagery and its credibility implications. The complete content analysis report and codebook with operational definitions for all variables are provided in Appendices C and D. Threat-focused framing dominated coverage, appearing in 70% of articles, while opportunity framing appeared in only 10% and balanced framing in 20%. Overall tone reinforced this pattern, with 65% of articles adopting alarmist or concerned language. The primary risk emphasis shifted from initial predictions, with financial fraud and scams dominating at 45% of articles, surpassing

misinformation concerns at 35%. Trust erosion appeared in 20% of articles, while copyright issues received no coverage. Source citation patterns revealed heavy reliance on technology industry voices, with tech companies cited in 90% of articles and researchers in 70%, while fact-checkers appeared in only 25% of articles. Artists received zero citations across the entire sample. Solutions proposed emphasized technical approaches, with detection tools mentioned in 80% of articles and disclosure requirements in 70%, while regulatory solutions appeared in only 35%. These patterns confirm H1, that threat frames would dominate coverage. H2 received partial support, as misinformation was emphasized but fraud concerns exceeded it. H3 was rejected, as tech companies appeared far more frequently than fact-checkers, indicating industry voices shape the discourse more than verification specialists. The findings suggest news coverage constructs AI imagery primarily as a technological problem requiring technological solutions, marginalizing regulatory approaches and excluding creative stakeholder perspectives entirely.

Findings: Focus Group

The focus group with five journalism students revealed immediate skepticism toward AI-generated imagery in news contexts. When asked for one-word associations, participants responded "fake," "deceiving," "not trustworthy," and "not newsworthy" before any prompting about credibility concerns. This unanimous negative framing suggests journalism students have internalized professional norms distinguishing photographic evidence from synthetic content.

Participants reported frequent encounters with AI images on TikTok, Instagram, and X, but detection proved difficult without explicit labeling. One described spending time evaluating content "thinking that it's true" only to discover images were synthetic, characterizing this as "really disturbing." Most platforms provide no disclosure, though participants noted Instagram occasionally labels AI content and Chinese media enforces strict labeling requirements.

In the visual comparison exercise, participants correctly identified the AI-generated wildfire image based on unnatural color saturation and intact vegetation despite flames. However,

several questioned whether they would detect synthetic origin without the research context priming scrutiny. One stated, "I'm questioning myself if I saw this on social media, would I actually think it was AI-generated?" This reveals detection depends on skepticism rather than automatic recognition.

Disclosure affects trust negatively but remains preferable to independent discovery. Participants explained that AI images "ruin the purpose of journalism" because audiences "want to see the amount of damage, how it actually looks like in real life." However, transparency with explanation might mitigate skepticism. One noted that distrust would be lower "if there was an explanation in the label" about why AI was necessary when no real footage existed.

Discussion revealed clear boundaries based on content type. Participants distinguished explainer videos, where AI aids comprehension, from breaking news coverage requiring authentic documentation. AI imagery becomes unacceptable when "used to create some chaos or panic among the people" or covering war and disaster.

When comparing AI to Photoshop, participants emphasized a fundamental difference. One explained, "We need to have actual image to edit, and yes, it makes it more valid." Participants recognized both can manipulate but maintained that synthetic creation without authentic photographic basis violates journalism's evidentiary standards.

The findings demonstrate journalism students apply professional frameworks when evaluating AI imagery, drawing clear distinctions between documentation and illustration. Their emphasis on transparency and context-specific acceptability reflects training in media ethics and credibility standards.

Findings: Interview

The semi-structured interview with a sophomore journalism student explored individual reasoning processes when evaluating AI-generated imagery in news contexts. The participant

described encountering AI images frequently across Telegram channels, The Guardian, and YouTube news coverage, noting that visual content appears in nearly every news post. She recalled specific examples including a viral AI-generated image of Dubai flooding that depicted "massive waves in the streets, very dramatic" but with water that "looked too perfect, like glass" with incorrect reflections. She also mentioned AI visualizations of future cities where disclosure was "buried" in article text.

When asked about her reaction to AI-labeled images, the participant articulated a skepticism that extends beyond the image itself to question the entire article's credibility. She stated, "Why are they using this? Is there no real photo? It makes me skeptical." This response reveals that AI imagery signals potential deficiencies in reporting quality. She added that "AI feels like a shortcut; it doesn't capture real damage or emotion," distinguishing sharply between photography as evidence of witnessing and AI generation as disconnected from reality.

The participant identified acceptable use cases as contexts requiring visualization of abstract concepts. She suggested AI might be appropriate "for things like health articles, where they need to visualize medical concepts" or "economic stories where they want to show abstract ideas like inflation or market trends." However, she maintained absolute opposition to AI imagery "for reporting on actual events. Never for that."

When evaluating visual credibility, the participant focused on technical cues including lighting quality, shadow consistency, and presence of imperfections. She described the AI image as having "this weird, too-perfect quality" with "lighting that is off, everything is too perfect in a way that doesn't happen in nature." In contrast, authentic photography "has shadows, irregular details."

The participant's most significant concern centered on institutional implications and vulnerable populations. She stated, "Deepfakes fooling everyone, panic from fake disasters, vulnerable people like my grandma scammed by AI videos. That's what worries me most." She

warned that "authoritarian governments could use AI to create false evidence of things that never happened, or deny things that did happen by claiming real photos are AI." She articulated anxiety about "the death of visual truth itself," explaining that "for over a century, photographs have been how we document reality. If we lose faith in that medium, we lose something really valuable." This framing positions AI imagery as threatening journalism's evidentiary foundation.

Looking forward, the participant indicated she would adapt by "relying more on text reporting from trusted journalists and less on visual evidence," characterizing this shift as "sad, because photography used to be one of the strongest forms of proof." She suggests that AI imagery's proliferation may fundamentally alter how audiences weight different evidence types in credibility assessment, potentially diminishing photography's historic role in documentation.

Discussion

Transparency about AI-generated imagery works as a credibility problem. The most important finding is that explicit disclosure triggers skepticism strong enough to cancel out the persuasive advantages typically associated with visual content. The outcome challenges basic assumptions about visual evidence in media credibility research. Professional training changes how audiences process synthetic imagery.

The credibility penalty associated with disclosed AI imagery extends previous research in an unexpected direction. While earlier work showed that synthetic content performs worse than authentic photography when both are labeled, the current findings show that labeling alone is enough to make AI imagery worse than providing no image at all. Disclosure does not reduce visual enhancement effects to neutral levels but signals weak reporting. The mechanism appears to work through audiences interpreting AI usage as evidence that journalists could not get real photographic documentation. Journalism students do not accept synthetic imagery as an alternative visualization method. They read it as an admission of inadequate source access or questionable news value. However, this interpretation assumes a particular cause-and-effect direction. An

alternative explanation is that the specific AI image used may have looked obviously artificial even without labeling, triggering skepticism that would not apply to better-quality synthetic imagery. The experimental design cannot tell whether the credibility penalty comes from disclosure itself or from visual characteristics of this particular image.

Professional norms about photographic evidence separate journalism training from general media literacy. The qualitative data showed participants categorizing AI imagery by content function, not visual quality. They accepted synthetic visuals for abstract concept illustration while completely rejecting them for event documentation. The functional distinction fits with journalism's evidentiary standards, where photographs serve to verify claims about reality rather than boost engagement. The basic principle treats photography as physically connected to what it shows, while AI generation represents imagination disconnected from witnessed events. However, students discussing AI imagery with a researcher studying credibility may perform more skeptical positions than they hold when consuming news casually.

The resistance to AI imagery in news contexts appears independent of individual experience with the technology. Despite substantial exposure and reported detection confidence, participants uniformly applied skeptical frameworks when encountering disclosed synthetic content. The null effect shows that professional identity overrides personal familiarity in shaping credibility judgments. Journalism education builds boundaries between legitimate and illegitimate visual evidence, and these boundaries persist regardless of whether students personally use AI tools or encounter synthetic content daily. The finding challenges assumptions that getting used to AI through exposure will reduce skepticism toward AI imagery in professional contexts.

Media discourse appears to strengthen these professional norms. The dominance of threat framing in news coverage positions AI imagery as endangering visual credibility, cited through technology industry voices that emphasize technical solutions over structural responses. The framing presents the problem as needing better detection tools and disclosure requirements instead

of regulatory intervention or professional standards development. Journalism students absorb the discourse while simultaneously learning evidentiary standards that position photography as uniquely suited to documentation. The combination produces skepticism based on both immediate fraud concerns and deeper worries about photography's documentary function being undermined.

The emphasis on individual vulnerability in both media coverage and participant responses needs critical examination. Framing AI imagery mainly as a threat to less sophisticated audiences positions the problem as one of consumer protection. The individualized framing hides structural questions about who controls image generation and verification technology, whose voices shape discourse about appropriate use, and what rules might govern synthetic content in journalism. The complete absence of artist perspectives in news coverage is particularly striking given that visual creators represent stakeholders directly affected by synthetic imagery growth yet excluded from defining the problem or proposing solutions.

Participants did express concerns extending beyond individual deception to institutional implications. The anxiety about authoritarian governments manipulating visual evidence or denying reality by claiming authentic photographs are AI-generated reveals awareness that credibility challenges work at systemic levels. The characterization of AI imagery as potentially causing "the death of visual truth itself" frames the issue as threatening journalism's documentary foundation. Journalism training builds understanding that photographic credibility has collective value beyond its usefulness for personal information assessment.

The theoretical contribution lies in showing that credibility works relationally based on content function and audience expertise. General credibility models treat visual content as either credible or not based on source characteristics and message features. The current findings show that journalism students evaluate credibility through context-based frameworks asking what function the image serves and what professional standards apply. An AI-generated visualization of economic trends may be perfectly credible while an AI-generated image claiming to document a

protest would be suspect, regardless of visual quality. The context-dependence challenges universal models. Credibility assessment requires understanding what kind of claim an image makes.

The finding that disclosure penalties can exceed visual advantages has practical implications for transparency policies. Policies requiring AI imagery labeling assume disclosure enables informed evaluation while keeping visual content's communicative benefits. The current findings show that in professional journalism contexts, disclosure may make synthetic imagery worse than text-only presentation. The pattern shows that labeling alone does not solve credibility concerns. Journalism organizations using AI imagery for legitimate illustration purposes may find that disclosure triggers skepticism undermining their broader credibility.

Conclusion

Labeling AI-generated images in news reduces credibility more than having no image at all, which contradicts typical assumptions about how visual content enhances persuasiveness. Journalism students rate AI images lower than text-only content because they interpret AI use as a sign of poor reporting.

The pattern reflects professional training that teaches students to distinguish between documentation and illustration. They accept AI images for explaining abstract ideas like economic trends but reject them completely for documenting news events because real photos serve as evidence of what happened while AI images show what someone imagined. That difference matters more than how realistic the image looks, and it is rooted in journalism's evidentiary standards that position photography as uniquely suited to documenting reality.

AI experience did not change how students evaluated the images, which was surprising given that 83% encounter AI images daily and 61% have used AI tools themselves. Their professional identity impacted their judgments more than their personal familiarity with the technology. The boundaries they learn in journalism classes persist regardless of how much AI

content they see or create, which challenges assumptions that exposure normalizes synthetic imagery.

The study has clear limitations that constrain how broadly the findings apply. The sample of 18 students from one institution means results may not generalize to audiences without journalism training or to different cultural contexts. The experiment tested one claim topic with one AI-generated image, leaving open whether results reflect broader patterns or characteristics specific to these materials. The uneven distribution across experimental conditions and small overall sample size limit confidence in effect magnitude. The content analysis examined a narrow slice of English-language coverage that may not represent global discourse patterns. The qualitative methods relied on articulated reasoning that may differ from actual decision-making in natural contexts.

Future research should examine whether similar patterns emerge in other expert communities with specialized visual literacy, whether disclosure effects vary across content domains and cultural contexts, and whether behavioral measures confirm self-reported reasoning. Longitudinal studies tracking how attitudes evolve as AI imagery becomes more prevalent would clarify whether current skepticism represents stable professional norms or transitional resistance to emerging technology.

The multi-method approach proved valuable because each method addressed blind spots in the others. The experiment established that disclosure reduces credibility but could not explain why, while interviews and focus groups revealed the reasoning process behind students' judgments. Content analysis showed how media discourse shapes the frameworks students apply, and focus groups captured how they deliberate about these issues collectively. The key methodological lesson is that quantitative and qualitative approaches answer different questions rather than producing the same knowledge through different means. The experiment measured effects while qualitative methods explained mechanisms, and neither can replace the other.

Working across four methods revealed that each approach answers different questions. The experiment established credibility differences but could not explain why. The interview uncovered unanticipated concerns, which revealed anxieties extending beyond the original research questions. The focus group captured collective recognition when participants discussed detection difficulty without research context. Content analysis faced sampling constraints from excluding paywalled outlets. The key challenge involved reconciling contradictions: the experiment showed no AI familiarity effect while qualitative methods revealed detection depends on context and peer feedback.

The findings have practical implications for journalism. Transparency about AI images does not solve credibility problems and can complicate them by signaling poor reporting to trained audiences. For journalism education, the results indicate that students need more than detection skills, they need frameworks for analyzing when and why AI images might be appropriate. As AI technology advances, professional communities will need explicit standards governing synthetic content.

The main contribution is demonstrating that professional socialization shapes how audiences evaluate synthetic imagery in ways that make disclosed AI content particularly problematic for journalism contexts. Understanding how professional training creates distinctive evaluation frameworks matters for developing policies around AI use in news and for thinking about how different communities with specialized expertise might respond differently than general audiences.

Sources

Bohm, S. (2024). Human perception and classification of AI-generated images: A pre-study based on a sample from the media sector in Germany. In *Proceedings of GPTMB 2024: The First International Conference on Generative Pre-trained Transformer Models and Beyond* (pp. 15-23). Retrieved from: https://www.thinkmind.org/library/GPTMB/GPTMB_2024/gptmb_2024_1_20_38004.html

Högemann, M., Betke, J., & Thomas, O. (2025). What you see is not what you get anymore: A mixed-methods approach on human perception of AI-generated images. *Frontiers in Artificial Intelligence*, *8*, Article 1707336. Retrieved from: https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2025.1707336/full

Long, J. A., Xiao, J., McCray, S., Ağaoğlu, E., Alajmi, A. M., Akalonu, C., & Xu, Y. (2024). Artificial impressions: Trust and credibility in AI-enhanced profile pictures [Conference presentation]. *AEJMC 2024*. Retrieved from: https://jacob-long.com/talk/artificial-impressions-trust-and-credibility-in-ai-enhanced-profile-pictures/

Osińska, V., Kortas, W., Szalach, A., & Welter, M. (2025). AI images vs. real photographs: Investigating visual recognition and perception. *Journal of Eye Movement Research*, *18*(6), Article 61. Retrieved from: https://www.mdpi.com/1995-8692/18/6/61

Velásquez-Salamanca, D., Martín-Pascual, M.Á., & Andreu-Sánchez, C. (2025). Interpretation of AI-generated vs. human-made images. *Journal of Imaging*, *11*(7), Article 227. Retrieved from: https://www.mdpi.com/2313-433X/11/7/227

Content Analysis Sources

AI Video Detector. (2025). How journalists use AI video detectors to verify news in 2025: Complete guide. AI Video Detector Blog. https://aidetector.video/blog/journalists-ai-video-verification

CBS News. (2023). Fake photos of Pope Francis in a puffer jacket go viral, highlighting the power and peril of AI. CBS News. https://www.cbsnews.com/news/pope-francis-puffer-jacket-fake-photos-deepfake-power-peril-of-ai/

Chipeta, C. (2025). Deepfake statistics 2025: 25 new facts for CFOs. Eftsure. https://www.eftsure.com/blog/deepfake-statistics-2025

Clark, S. (2026). The continued influence of AI-generated deepfake videos despite transparency warnings. Nature Communications Psychology. https://www.nature.com/articles/[article-doi]

Edwards, G. (2025). AI is undermining OSINT's core assumptions. Here's how journalists should adapt. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/news/ai-undermining-osints-core-assumptions-heres-how-journalists-should-adapt

Hamiel, N. (2025). Deepfakes proved a different threat than expected. Here's how to defend against them. World Economic Forum. https://www.weforum.org/stories/2025/01/deepfakes-different-threat-than-expected/

Hernandez, J. (2025). Deepfake detection and digital authenticity in the AI era. Status Labs. https://statuslabs.com/blog/deepfake-detection-digital-authenticity

Hetzner, C. (2023). What is Midjourney? The A.I. image generator used to create the viral image of the Pope in a puffer jacket. Fortune. https://fortune.com/2023/03/28/pope-francis-balenciaga-puffer-jacket-midjourney-ai-deepfake-image-generator-openai-dalle/

Khalil, M. (2025). Deepfake statistics 2025: The data behind the AI fraud wave. DeepStrike. https://deepstrike.io/blog/deepfake-statistics-2025

Lyu, S. (2026). Deepfakes leveled up in 2025 – here's what's coming next. The Conversation. https://theconversation.com/deepfakes-leveled-up-in-2025

Naffi, N. (2025). Deepfakes and the crisis of knowing. UNESCO. https://www.unesco.org/en/articles/deepfakes-and-crisis-knowing

National Security Agency, Federal Bureau of Investigation, & Cybersecurity and Infrastructure Security Agency. (2025). Strengthening multimedia integrity in the generative AI era. https://media.defense.gov/2025/Jan/29/2003634788/-1/-1/0/CSI-CONTENT-CREDENTIALS.PDF

PMC. (2025). AI-driven disinformation: Policy recommendations for democratic resilience. PubMed Central. https://pmc.ncbi.nlm.nih.gov/articles/PMC[article-number]

Salát, M. (2025). Top 5 ways scammers have used AI and deepfakes in 2025. Norton. https://us.norton.com/blog/online-scams/top-5-ai-and-deepfakes-2025

Scire, S. (2025). Trusted news sites may benefit in an internet full of AI-generated fakes, a new study finds. NiemanLab. https://www.niemanlab.org/2025/08/trusted-news-sites-may-benefit-in-an-internet-full-of-ai-generated-fakes-a-new-study-finds/

Simon, F. (2025). Generative AI and news report 2025: How people think about AI's role in journalism and society. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/generative-ai-and-news-report-2025-how-people-think-about-ais-role-journalism-and-society

Strickland, E. (2024). Content Credentials will fight deepfakes in the 2024 elections. IEEE Spectrum. https://spectrum.ieee.org/deepfakes-election

UNRIC. (2025). Artificial intelligence and the future of journalism: Risks and opportunities. https://unric.org/en/artificial-intelligence-and-the-future-of-journalism-risks-and-opportunities/

Wasi, A. (2025). Seeing isn't believing: Addressing the societal impact of deepfakes in low-tech environments. arXiv. https://arxiv.org/abs/[paper-number]

Zero Threat. (2026). Deepfake Statistics 2026: AI Phishing Attacks, Data, and Cybersecurity Trends. Zero Threat. https://www.keepnet.com/blog/deepfake-statistics-trends-2026

Appendix A: Semi-structured interview transcript

Date: March 23, 2026

Interviewer: Thank you so much for taking the time to do this interview with me today.

Student: Of course!

Interviewer: So let's start with an easy question. Can you tell me about how you typically consume news? What platforms or sources do you use most often?

Student: I mostly use Telegram channels for news, especially Russian channels. I also check The Guardian website sometimes, and I follow a few journalists on Instagram who post about Central Asia. Oh, and YouTube - I watch DW News and Al Jazeera videos sometimes.

Interviewer: How often would you say you encounter images in the news content you consume?

Student: Almost always. Even on Telegram, every post has at least one image or video.

Interviewer: Have you noticed AI-generated images appearing in news articles, social media posts, or other media content recently? Can you describe any specific examples you remember?

Student: Yeah, I've seen quite a few. There was this image going around about flooding in Dubai. It showed these massive waves in the streets, very dramatic. Turned out it was AI-generated, not from the actual flooding. And I've seen AI images used in articles about future technology, like what cities might look like in 2050. Those are obviously not real, but sometimes it's hard to tell if the outlet is using them as illustration or trying to pass them off as predictions from experts.

Interviewer: What made you realize or suspect those images were AI-generated?

Student: With the Dubai one, people in the comments were pointing it out. The water looked too perfect, like glass. And the reflections weren't right. For the future city images, usually the article mentions somewhere that these are "AI-generated visualizations," but they hid that information. You have to read carefully.

Interviewer: Interesting. So walk me through what goes through your mind when you see an image labeled as "AI-generated" in a news post or article. How does that change how you think about the content?

Student: Okay, why are they using this? Is there no real photo? It makes me skeptical. Like, is this story even real if they need to fake the visual? AI feels like a shortcut, because it doesn't capture real damage or emotion. So when I see that label, I start questioning the whole article's quality.

Interviewer: So when you're evaluating whether to believe a news claim, what factors do you typically consider?

Student: I check who wrote it first, like is it a journalist I recognize or trust? Then I look at the publication. Is this a serious outlet or some random site? I also see if they cite sources properly, like quotes from officials or links to documents. And if the claim seems really dramatic or surprising, I try to find it reported somewhere else too.

Interviewer: How important are images or visual elements in that evaluation?

Student: Pretty important. A real photo from the scene makes me feel like the journalist was there, doing real reporting. It's proof that something happened. Without it, the story feels less solid.

Interviewer: Okay, so imagine you see the same news story presented two ways: one with a photograph and one with an AI-generated image labeled as such. How would you approach evaluating each version?

Student: With the photograph, I'd look at it more carefully to see if it actually shows what the article claims. Like, does this photo really support the story? With the AI version, I'd be more skeptical of the entire article. I'd think, "If they can't show me real evidence, maybe this didn't happen the way they're saying."

Interviewer: Would you trust one more than the other? Why?

Student: Definitely trust the photo version more. Because a photograph means someone was there. They witnessed it. With AI, anyone could generate that sitting at a computer anywhere in the world. So there's no connection to reality.

Interviewer: That makes sense. So, do you think news organizations should be allowed to use AI-generated images in their reporting? Under what circumstances, if any, would that be acceptable to you?

Student: Maybe for things like health articles, where they need to visualize medical concepts? Or economic stories where they want to show abstract ideas like inflation. Things that don't have a specific visual already. But not for reporting on actual events. Never for that.

Interviewer: Should there be any restrictions or guidelines?

Student: Yes, absolutely. There should be a very visible labe. And journalists should have to explain why they're using AI instead of a real photo. Like, what's the justification? If they can't give a good reason, they shouldn't use it.

Interviewer: How important is it to you that AI-generated images are clearly labeled in news content?

Student: I would say very important. Without labeling, it's basically deception. The reader has a right to know what they're looking at.

Interviewer: Okay, so now I'm going to show you two images. Can you tell me which of these images you find more credible for illustrating a news story, and why?

[Interviewer shares screen showing human photograph and AI-generated image]

Student: This one on the right looks more credible. The left one has a weird quality. The lighting is off, everything is too perfect in a way that doesn't happen in nature. The right one has shadow and tiny irregular details.

Interviewer: What specific visual cues or elements influenced your judgment?

Student: The shadows and light quality, mainly. Also the right one has little imperfections and random details. The left one looks almost rendered, like from a video game. Too clean.

Interviewer: That's helpful. So, do you think your concerns about AI-generated images differ depending on the context? For example, in hard news reporting versus feature stories, or in social media posts versus traditional journalism?

Student: Yes, definitely. For hard news I have zero tolerance. That needs to be real documentation. For feature stories about culture or lifestyle, I'm slightly more flexible if it's clearly labeled. Social media is different because I already expect most content there to be exaggerated. I'm more skeptical by default.

Interviewer: As AI-generated imagery becomes more realistic and widespread, what concerns do you have about being able to trust visual content in news and social media?

Student: Deepfakes fooling everyone, panic from fake disasters, vulnerable people like my grandma scammed by AI videos. My grandma could believe scammers on the phone that show any pictures, send any pictures or videos. So she would think that it's trustworthy and believe them to do anything that they are asking her to do. I'm worried we'll reach a point where nothing can be trusted visually anymore. Like, every image will need to be verified through some technical process, and regular people won't be able to do that. It puts too much power in the hands of whoever controls the verification technology. And in the meantime, authoritarian governments could use AI to create false evidence of things that never happened, or deny things that did happen.

Interviewer: Do you think you'll change how you evaluate news content in the future?

Student: I'll probably rely more on text reporting from trusted journalists and less on visual evidence. Which is sad, because photography used to be one of the strongest forms of proof. I might also need to learn how to use verification tools myself, if those become available to regular people.

Interviewer: That's a really important point. So, my last question: Is there anything else about AI-generated images, news credibility, or visual media that you think is important for me to understand?

Student: I think the biggest issue is that AI imagery undermines the entire foundation of visual journalism. For over a century, photographs have been how we document reality. If we lose faith in that medium, we lose something valuable. And I'm not sure we can get it back once it's gone.

Interviewer: Thank you so much, Student. This has been incredibly helpful for my research.

Student: No problem!

Appendix B: Focus group transcript

Date: March 31, 2026

Moderator: First, thank you so much for being here, for taking the time to participate in this group. My research is how people perceive the credibility of news content when it includes different types of images, specifically human-generated photos versus AI-generated images. Before we get started, I really want you to remember that there are no right or wrong answers. I want to hear your honest perspectives and opinions. So let's start with something quick and easy. I'm going to go around this Zoom room and I want each of you to share just one word that comes to your mind when you hear the phrase AI-generated images in news content. Just the first thing that pops into your head.

Student 1: Fake.

Student 2: Fake.

Student 3: Yes, for me it's also fake and it's like a sign for me that if the image is AI-created, like the first option is that there may be no videos or photos available, so that's why they used it. Or the second thought is like the news also can be AI-generated.

Moderator: Yeah, true. Thank you so much.

Student 4: Maybe “deceiving.”

Student 4: And like not trustworthy enough.

Moderator: Agree. Student 5, do you have something to share?

Student 5: I have the same words as Student 4. Also, I think it's not newsworthy.

Moderator: Yes, true. Thank you so much, guys. So my first question is, how often do you all encounter AI-generated images in your daily news, and where are you seeing them mostly?

Student 3: Sometimes I see them on Instagram or TikTok, which is really disturbing for me because I keep paying attention to them, thinking that it's true. Sometimes they look very trustworthy. I cannot tell that they are fake, so that's why I keep paying my attention and wasting time just to find out that they are fake.

Moderator: Interesting.

Student 4: Yeah, and recently when I started scrolling on X, like Twitter, I noticed that it's full of deep fakes and AI-generated content that at some point you cannot even differentiate.

Student 3: I agree with Student 4. You can't differentiate between what's real and what's fake. For example, there was this rumor coming out that Netanyahu was killed, and then there was a recently released video of him ordering coffee, and people are questioning that it's AI. And we don't know anymore if it's AI or not. It's kind of scary.

Student 3: And after that, they started making AI videos of other presidents in the same coffee shop getting coffee, and they also looked very realistic at some point.

Moderator: Are these contents usually labeled as AI-generated, or are you detecting them yourselves? How do you know that the content is AI-generated in cases that you can’t differentiate?

Student 3: Sometimes they look like cartoons, or if, for example, my friend posted a picture of a very famous person with a president or someone like this, I could tell that it's AI-generated. But most of the cases, I think Instagram has this label, like AI-generated post, so I can differentiate, but most of the time, I cannot. And most of the time, they do not label them as AI.

Student 2: Yeah. It's getting more and more natural now, and it's getting very scary.

Student 1: I don't think they label much, especially if the purpose of that content is to be shared and create something among the population. I don't think they will label that as AI-generated.

Student 5: I would say in Chinese media, there's very strict policies, and I think most of the AI-generated photos that will appear on Chinese newspaper and media, they will label as AI-generated. Otherwise, someone's going to report. So I think it's very strict, at least in Chinese media.

Student 2: In Kyrgyzstan, there's no restriction on using AI, but I think we need it. We need it definitely.

Moderator: Let me conduct a short exercise. I want to share two pictures. Can you see it, guys?

Students: Yes.

Moderator: Okay, thank you. You should all be seeing these two images on your screen. These pictures show wildfire smoke in the forest landscape. My question is, just imagine both of these images were used to represent the breaking news story about, let's say, recent wildfires and climate change. Which would you find more credible, and why?

Student 2: The right one looks more real, because I've seen fire in real life, and I don't think it looks that orange.

Moderator: Other thoughts?

Student 1: I also agree the second one is more realistic. The first one is more exaggerated. It also depends: a fire can be like the first one, but the second one looks more realistic, as Student 2 said.

Student 3: Yeah, the first one is like more lava, or I don't know, it's pink.

Student 4: The first one seems very neat. I don't see any trees harmed in the picture, they all seem fine.

Student 5: True, I agree.

Moderator: Thank you. You all agree that the first picture is AI-generated, right?

Student 4: Yes, but I feel like, that's because I know we are talking about AI-generated images, and one of them is going to be AI-generated. But now I'm questioning myself if I saw this on social media, would I actually think it was AI-generated, you know?

Moderator: Yes.

Student 4: Maybe I wouldn't go that deep and be like, oh, this is AI-generated.

Student 1: Yes, it depends on the context, because I don't think I would think that it was AI, if I wasn't asked about the AI. Just in general, I would be surprised if none of them AI, if we consider that one of them should be AI.

Student 5: It's the first one.

Moderator: You're right, guys! The first picture is AI generated. It looks very natural, I would say. Especially the lower part of the picture, trees look like they are photographed. But this orange smoke is giving cues that this is AI, definitely.

Moderator: But what if this image was clearly labeled as AI-generated? Does that change your perception of the new story? Would you trust the story more or less, or the same?

Student 2: I would trust less, because it doesn't really depends on the outlet. Why they're using AI images in the first place, like they could just publish something without an image. And I wouldn't know the scales of the disaster if they used AI image.

Student 3: I agree. I would also trust it less, because it ruins the purpose of journalism. As people, we also want to see the amount of damage, how it actually looks like in real life. So I feel like the AI image has no purpose.

Moderator: Yeah, definitely.

Student 5: For me, if they labeled it as AI, and then also gave an explanation, why it's AI, why they chose to use AI…Maybe there's no footage from that or something… I wouldn't distrust if there was an explanation in the label.

Student 5: I would distrust more if there was nothing written, and then I would find out myself that it was AI.

Moderator: Absolutely! Thank you guys for participating in this exercise. It was very interesting. Let's stay on the topic of disclosure. My next question is: should newsrooms be required to label AI-generated images, and what happens if they don't? What do you all think?

Student 5: As I said, I would prefer if newsrooms add labeling, that's important. Because if later on the audience finds out themselves that it's AI, a big drama comes out of it, everyone is going to feel deceived. And after that, I'm 100% sure I will lose my trust towards that news outlet. But if it's labeled AI from the beginning, at least they're not lying to you.

Student 3: I agree with Student 5. I would also trust more if they had this warning that this AI-generated.

Student 4: Yeah, I agree.

Moderator: The follow-up question came to my mind here. Where should we draw the line between what is acceptable and not? Like, are there specific situations or cases where using AI images is totally fine versus not fine? What do you think?

Student 1: I feel like it should be fine if it weren't breaking news. If it was a video explaining something, it would be acceptable. Because I see that a lot, and it's actually very helpful. Entertaining journalism, for example, explaining some complicated concepts through AI, things like that. But I feel like for breaking news, we have to see on the ground actual images. Because that's what journalism is.

Student 2: Also, if it's not used to create chaos or panic among people. If it's not about a war where they're showing AI-generated destruction just to have the news go viral or for some other purposes. And if there is, again, as Student 1 said, video explainers, something that won't harm anyone.

Moderator: Yes.

Student 4: Yeah, I also agree. And I think that if it's for entertainment purposes, it's totally okay, as long as it's also labeled as AI-generated content.

Moderator: For sure! Labeling is important. Some people argue that AI images are not really different from Photoshop, or color correction, or cropping, because journalists use these things to edit their photographs. And some people believe that they're completely different, because nothing real was being photographed. What do you think about that?

Student 3: I agree with the second comment. It's completely different, because we need to have actual image to edit, and yes, it makes it more valid.

Student 5: I feel like in the first case, that's also done with a purpose to create some visuals for the audience. That's another type of deceiving content. I've noticed, during some protests, they will crop to show more people, as if more people are joining, or to crop to show less people, to create this idea that no one is joining, or some other things. But I feel like they are different in terms of the way this deceiving content is created. They're not the same. But maybe the purpose can be the same.

Moderator: Looking ahead five years, AI-generated images will be even more common, realistic, and natural. What do you think, what should journalism schools be teaching students like us about assessing and evaluating visual content?

Student 2: Maybe as journalists, we need to learn how to be ethical, and how to differentiate between. I think maybe we need some tools, because it's hard to differentiate with your own eye. If technology gets even more advanced than how AI detectors exist now for text, then definitely should be for images as well. Journalists will use those tools, at least to verify some information. We also need to have very strong fact-checking skills to verify information.

Moderator: Thank you. Is there anything about AI images in journalism that we haven't discussed today? Anything that you want to add to our discussion?

Student 1: I would agree on our last conversation about the journalists, that they should detect the AI-generated content, because I totally agree with what you and Student 2 said. We should find some ways to analyze if the content is AI-generated or not. Maybe create applications that do it, or something like that. Because it's really important nowadays, because we have people that are not really informed about AI. I think we should draw the line between what is real and what is not.

Moderator: Thank you so much, guys! This has been very helpful. Thank you for your insights!
Students: Thank you. Thank you for the interesting focus group conversation. Bye!
Moderator: Have a good day, guys. Bye-bye!

Appendix C: Content analysis report

My content analysis of 20 articles about AI-generated imagery and credibility revealed clear patterns in how news media frame this emerging technology. I coded each article across five dimensions: dominant frame, overall tone, primary risk emphasis, expert sources cited, and solutions proposed.

The most striking finding was the overwhelming dominance of threat-focused framing. 14 of 20 articles (70%) emphasized risks and dangers of AI-generated imagery, compared to only 2 articles (10%) that framed it as an opportunity. The tone was similarly concerning, with 13 articles (65%) coded as either alarmist or concerned, while only 1 article took an optimistic stance.

The primary risk emphasis shifted over time. Early 2023 articles focused on misinformation (e.g., viral Pope deepfake), but by late 2024-2025, fraud and financial scams emerged as the dominant concern. 9 articles (45%) emphasized fraud as the primary risk, compared to 7 articles (35%) emphasizing misinformation. Notably, copyright concerns, despite being prominent in public discourse, did not appear as the primary focus in any article.

The source citation patterns revealed heavy reliance on technology industry voices. Tech company representatives appeared in 18 of 20 articles (90%), averaging 2.3 citations per article. In contrast, fact-checkers were cited in only 5 articles (25%), averaging just 0.4 citations per article. This disparity suggests a potentially industry-driven narrative. Artists received zero citations across all 20 articles, indicating a significant gap in stakeholder representation.

For proposed solutions, detection tools dominated the discourse, mentioned in 16 articles (80%), followed by disclosure/labeling in 14 articles (70%), and regulation in only 7 articles (35%). This pattern suggests a preference for technological solutions over policy interventions.

These findings have important implications for public perception. The heavy emphasis on threats combined with industry-dominated sourcing may shape audiences to view AI imagery primarily as a risk while reinforcing industry-preferred technical solutions over regulatory approaches.

Future improvements would include: (1) expanding the sample to include paywalled outlets like The New York Times and Washington Post for comparison; (2) conducting inter-coder reliability checks with a second researcher; (3) analyzing temporal patterns more systematically across quarterly periods; (4) comparing coverage across different types of outlets (traditional news vs. tech media vs. cybersecurity industry).

Appendix D: Content analysis codebook

Code	Variable	Values	Description / Coding Rules	Example
FRAME	Dominant Frame	1 = Threat/Risk 2 = Opportunity 3 = Balanced	The primary interpretive lens organizing the article. Code based on headline, lead, and overall emphasis.	1: "AI Images Threaten Truth" 2: "AI Revolutionizes Creativity" 3: Equal weight to risks and benefits
TONE	Overall Tone	1 = Alarmist 2 = Concerned 3 = Neutral 4 = Optimistic	Affective stance toward AI imagery, assessed through language and framing.	1: "Existential threat" 2: "Experts urge caution" 3: "Study examines rates" 4: "Exciting possibilities"
RISK	Primary Risk	1 = Misinformation 2 = Trust Erosion	The specific credibility concern receiving most article attention by space and prominence.	1: Fake political images 2: Public losing faith in photos

Code	Variable	Values	Description / Coding Rules	Example
		3 = Copyright 4 = Other		3: Artists' work stolen
SOURCE	Expert Sources (count each type)	Tech companies: ___ Researchers: ___ Fact-checkers: ___ Artists: ___ Journalists: ___	Count each distinct named individual or organization quoted or cited, tallied by category.	Tally each distinct named source by category
SOLUTION	Solutions Mentioned (1/0 for each)	Disclosure: ___ Detection tools: ___ Regulation: ___	Whether article proposes interventions. Mark 1/0 for each solution type, even if mentioned briefly.	Mark 1 if mentioned, even briefly

Coding spreadsheet:

Article	Source	Date	FRAME	TONE	RISK	Tech Co	Research	Fact-ck	Artists	Journal	Discl	Detect	Regul
A01	CBS News	3/28/25	1	2	1	2	1	0	0	2	1	1	0
A02	Fortune	6/19/23	3	3	1	3	1	0	0	1	1	0	1
A03	IEEE Spectrum	3/18/25	2	4	1	6	0	0	0	4	1	1	1
A04	UNESCO	2/5/2026	1	1	2	1	12	0	0	0	0	0	0
A05	WEF	1/10/25	3	3	4	1	0	1	0	1	0	1	1
A06	UNRIC	10/13/25	1	2	4	2	0	1	0	1	1	1	0
A07	Norton	1/7/25	3	3	4	1	0	1	0	1	0	1	1
A08	AI Video Detector	7/31/23	2	3	1	7	0	3	0	3	0	1	0
A09	Status Labs	12/19/25	1	1	4	1	0	0	0	0	1	1	0
A10	NSA/FBI	1/2025	2	3	1	4	2	0	0	1	1	1	1
A11	The Conversation	2/2/26	1	2	2	2	1	0	0	0	1	1	0
A12	NiemanLab	10/28/25	1	2	4	2	0	1	0	1	1	1	0
A13	Eftsure	5/29/25	1	1	4	0	2	0	0	0	0	1	0
A14	PMC Journal	2025	1	2	1	3	5	0	0	0	1	1	1
A15	DeepStrike	9/8/25	1	1	4	1	1	0	0	0	0	1	0
A16	Zero Threat	1/2026	1	1	4	2	2	0	0	0	0	1	0
A17	arXiv	8/13/25	1	2	2	0	8	0	0	0	1	1	1
A18	Nature	1/2/26	1	3	2	0	6	0	0	0	1	0	0
A19	Reuters Inst	12/12/25	1	2	1	3	1	1	0	2	0	1	0
A20	Reuters Inst	10/7/25	3	3	2	0	2	0	0	2	1	0	0