



(REVIEW ARTICLE)



Ethical implications of generative AI in journalism: Balancing innovation, truth, and public communication trust

Samson Olufemi Olanipekun ^{1,*} and Oluwasegun Olakoyenikan ²

¹ *Rave 91.7 FM, Osogbo, Osun State, Nigeria.*

² *Arthur L. Carter Journalism Institute, New York University, USA.*

World Journal of Advanced Research and Reviews, 2022, 16(03), 1293-1311

Publication history: Received on 16 October 2022; revised on 22 December 2022; accepted on 26 December 2022

Article DOI: <https://doi.org/10.30574/wjarr.2022.16.3.1159>

Abstract

The integration of generative artificial intelligence (AI) into journalism marks a transformative era in content creation, newsroom automation, and audience engagement. Tools such as large language models and generative image and video systems now enable rapid production of headlines, articles, infographics, and multimedia packages, dramatically improving editorial efficiency. However, this innovation introduces profound ethical questions that challenge long-standing journalistic values of truth, transparency, accountability, and public trust. This paper explores the ethical implications of generative AI in journalism from both philosophical and operational perspectives, aiming to define a responsible framework for its adoption. At a broad level, the study examines how AI-generated content risks eroding trust when its authorship is ambiguous or when factual inconsistencies go unrecognized. It also addresses the societal impact of automating editorial decisions, which may reflect algorithmic bias, reduce diversity of perspectives, or marginalize human editorial judgment. Narrowing its focus, the paper reviews emerging industry practices in disclosure, AI audit trails, and newsroom policy formation to govern AI usage responsibly. Drawing from real-world case studies involving synthetic anchors, AI-generated news summaries, and deepfake reports, the paper highlights both the potential and perils of AI augmentation in journalism. It also evaluates regulatory proposals and ethical guidelines issued by professional journalism bodies and media councils. Ultimately, the paper advocates for a hybrid model of human-AI collaboration where journalistic oversight, algorithmic transparency, and ethical foresight coalesce to maintain credibility. This model ensures that while innovation is embraced, truth and public communication trust remain the cornerstone of responsible journalism in the digital age.

Keywords: Generative AI; Journalism Ethics; Public Trust; Media Integrity; AI Disclosure; Editorial Accountability

1. Introduction

1.1. Background and Motivation

Artificial Intelligence (AI) is rapidly transforming journalism, offering new tools for content generation, data analysis, audience engagement, and newsroom automation. From algorithmically generated financial reports to AI-assisted transcription and natural language summaries, these technologies promise increased speed, efficiency, and customization in journalistic output [1]. News organizations are now deploying machine learning systems to detect emerging trends, recommend personalized stories, and even write initial drafts for routine topics.

The integration of AI into journalism is not just a technological innovation—it signals a deeper evolution in how information is curated, validated, and disseminated in the public sphere. For example, The Washington Post's "Heliograf" and Bloomberg's "Cyborg" systems demonstrate how AI can support human journalists in producing content

* Corresponding author: Samson Olufemi Olanipekun

at scale, especially during breaking news events or financial reporting cycles [2]. Additionally, AI-driven analytics enable editorial teams to better understand reader preferences, shaping editorial strategies in real time.

However, while the efficiency and scalability offered by AI tools are compelling, they also introduce significant ethical dilemmas. Questions arise concerning editorial control, algorithmic bias, misinformation amplification, and the erosion of traditional journalistic norms such as transparency and accountability [3]. The growing use of generative models such as GPT in content creation raises particular concerns about authorship, credibility, and manipulation, especially when outputs are indistinguishable from human-written narratives [4].

Given the central role of journalism in shaping public discourse and democratic participation, the deployment of AI technologies in newsrooms demands careful scrutiny. It is no longer sufficient to assess these tools based on performance or productivity metrics alone; their societal and ethical impacts must be evaluated alongside their functional capabilities [5].

1.2. Research Problem and Ethical Tensions

Despite AI's potential to enhance journalistic workflows, its unchecked integration into editorial processes raises profound ethical tensions. One central concern is algorithmic opacity—many AI systems function as “black boxes,” offering little insight into how they generate or prioritize content. This undermines journalistic transparency, a key pillar of public trust [6].

Another issue is editorial bias embedded in algorithms. Training data often reflect social inequalities, which can inadvertently perpetuate racial, gender, or political biases in story recommendations or automated coverage [7]. For instance, AI-generated headlines may disproportionately feature male voices or favor dominant political narratives, skewing media representation and discourse.

Moreover, accountability gaps arise when AI-generated stories are disseminated without clear attribution or human oversight. If misinformation is published by an AI system, it is often unclear who bears responsibility—the journalist, the developer, or the organization [8]. These challenges raise critical questions about editorial integrity and the limits of automation in normative journalistic roles.

As newsrooms experiment with AI across content creation, moderation, and distribution, there is an urgent need for frameworks that reconcile innovation with ethical accountability, preserving journalism's democratic function while embracing technological progress [9].

1.3. Objectives and Structure of the Paper

This paper aims to critically explore the intersection of AI innovation and journalistic ethics, with the goal of developing practical frameworks for responsible implementation in digital newsrooms. The primary objectives are to:

- assess the current state of AI adoption across journalistic practices and identify common use cases;
- evaluate the ethical risks associated with algorithmic decision-making in content generation, curation, and audience targeting;
- propose governance frameworks grounded in transparency, fairness, and human oversight; and
- recommend design principles for integrating ethical AI tools into editorial workflows [10].

The paper is organized into five sections. Section 2 reviews the landscape of AI tools in journalism, including generative models, personalization engines, and newsroom automation systems. Section 3 addresses ethical concerns related to bias, transparency, misinformation, and responsibility. Section 4 presents a multi-stakeholder framework for AI governance in journalism, drawing from legal, technical, and ethical literature. Section 5 discusses emerging practices in responsible innovation, highlighting case studies from leading newsrooms. Finally, Section 6 offers policy recommendations and a roadmap for future research.

By advancing a dialogue between innovation and responsibility, the paper contributes to shaping an AI-augmented journalism ecosystem that upholds public trust, democratic values, and professional standards in the digital age.

2. Foundations of generative ai in journalism

2.1. Overview of Generative AI Technologies (LLMs, GANs, Diffusion Models)

Generative Artificial Intelligence (GenAI) refers to a class of machine learning models designed to produce original data outputs, such as text, images, or audio, based on learned patterns in training datasets. These systems have become central to AI integration in journalism due to their capacity to automate content creation while mimicking human-like complexity. Among the most prominent are Large Language Models (LLMs), such as GPT, BERT, and Claude, which are trained on billions of text parameters and can generate coherent narratives, summaries, and question-answer pairs across domains [5].

Alongside LLMs, Generative Adversarial Networks (GANs) have revolutionized image and video synthesis by pitting two neural networks—a generator and a discriminator—against each other to refine synthetic realism. In the media domain, GANs are being used to create compelling visualizations, photo enhancements, and even AI-generated anchors or avatars [6]. Their hyper-realistic outputs have significant potential for news design but also present new risks related to misinformation and visual forgery.

More recently, Diffusion Models have emerged as powerful generative tools that construct content by iteratively refining random noise into structured data. These models are gaining traction for creating photorealistic images and audio that can complement or illustrate journalistic stories [7]. Unlike GANs, diffusion models tend to offer more control and transparency in the generation process, which aligns better with newsroom accountability needs.

Together, these generative technologies are reshaping how stories are conceptualized, produced, and distributed in journalism. However, their technical sophistication necessitates careful editorial governance to ensure their use enhances rather than compromises news integrity [8].

2.2. Applications in Journalism: Text, Visual, and Audio Content Generation

In journalistic practice, generative AI has seen widespread adoption across text, visual, and audio production pipelines. Text generation is perhaps the most visible application. News organizations now use LLMs to draft articles on earnings reports, sports recaps, or weather summaries, thereby accelerating content production for routine topics [9]. AI-generated summaries and headlines are also employed to enhance content discoverability and personalization. These tools allow journalists to focus on investigative or analytical pieces while delegating repetitive writing to machines.

Additionally, generative AI is employed in transcription and translation, automatically converting interviews or foreign-language sources into editable text. This reduces manual labor and improves multilingual reporting, which is essential for covering global issues [10]. Chatbots powered by LLMs are also being used for interactive storytelling, offering readers a conversational exploration of complex topics such as climate change or policy reform.

On the visual side, GANs and diffusion models generate infographics, illustrations, or synthetic faces for privacy-sensitive reporting. For example, when interviewing vulnerable populations or reporting from conflict zones, AI-generated avatars can replace real photos to protect identities while maintaining engagement [11]. AI-powered design assistants are also integrated into tools like Adobe Express, helping journalists produce social media visuals or infographic-based summaries at speed.

Audio content generation has likewise advanced, with text-to-speech models enabling the creation of podcast narrations or audio summaries of articles. Some outlets use AI anchors to read news bulletins in multiple languages, expanding accessibility and audience reach. These tools also facilitate news delivery on smart assistants and wearable devices [12].

Despite the operational advantages, the use of generative AI in journalism raises ethical concerns, especially when synthetic content is not transparently labeled. There is a growing push for AI-content disclosure standards, where audiences are informed when a story or element has been produced or assisted by AI [13].

As generative tools become embedded in editorial workflows, maintaining editorial oversight, disclosure, and source traceability becomes critical to preserving public trust in journalism.

2.3. Editorial Automation vs. Human Judgment: A New Paradigm?

The integration of generative AI into editorial functions introduces a fundamental shift in how newsrooms approach journalistic decision-making. While automation promises efficiency, the growing reliance on AI systems challenges traditional notions of human editorial judgment, especially in tasks involving framing, prioritization, and ethical considerations [14].

AI models, when used for content curation or headline generation, are often optimized for engagement metrics such as click-through rates or reader retention. This optimization can lead to algorithmic sensationalism, where content is shaped to provoke rather than to inform. Without human review, such tendencies may compromise the integrity of editorial values, including fairness, balance, and factual accuracy [15].

Moreover, generative models lack contextual understanding, empathy, and cultural nuance—attributes that human journalists bring to the editorial process. For example, an AI system may accurately summarize a political speech but fail to recognize coded language, historical allusions, or implications for marginalized communities [16]. This limitation underscores the need for a human-in-the-loop framework, where AI tools serve as assistants rather than autonomous agents.

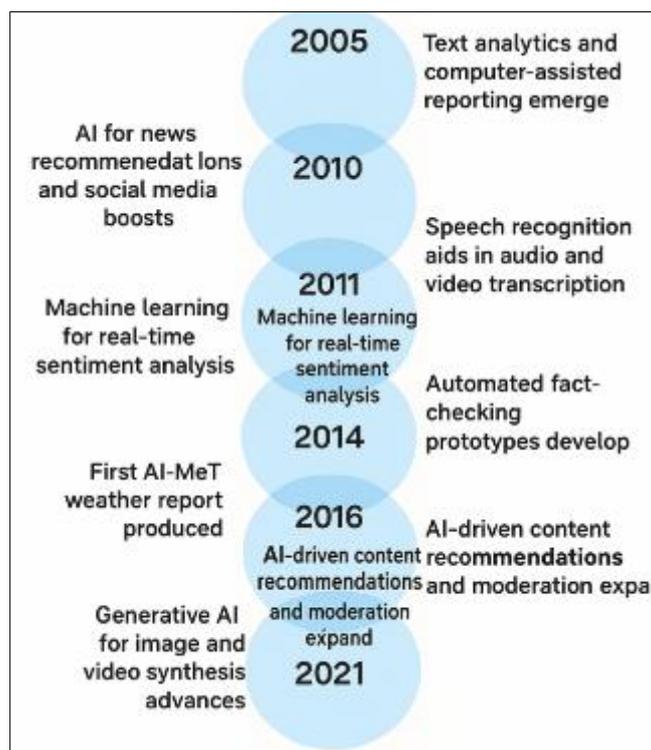


Figure 1 Evolution timeline of AI applications in journalism (2005–2022)

There is also a risk of editorial homogenization, as the same AI models are deployed across news organizations, producing uniform styles, tones, and perspectives. Such uniformity may reduce diversity in storytelling and erode media plurality [17].

Balancing automation with journalistic craftsmanship requires clear editorial policies, training programs, and AI audit mechanisms. Rather than displacing journalists, AI should augment their work—streamlining workflows while safeguarding human editorial oversight.

This new paradigm demands not only technological adaptation but also a cultural redefinition of journalistic roles in an AI-augmented newsroom.

Table 1 Comparison of Generative AI Tools by Output Type and Journalistic Use Case

AI Tool	Output Type	Primary Journalistic Use Case	Strengths	Limitations
ChatGPT (OpenAI)	Text (articles, headlines, summaries)	Drafting news content, summarizing sources, editorial assistance	Fluent, fast content generation; adaptable tone	May hallucinate facts; lacks real-time news awareness
DALL·E (OpenAI)	Images and illustrations	Visual accompaniment for stories; concept visualization	Generates unique visuals from prompts	Contextual relevance to journalism may be weak
Synthesia	AI-generated video avatars with voice	Explainer videos, multilingual news translation	Supports rapid video production in multiple languages	Avatars may appear synthetic; limited expressive nuance
Descript Overdub	Audio (AI voice cloning)	Podcasting, voiceover corrections, emergency narration	Accurate voice replication; useful for corrections	Ethical concerns with voice cloning; risk of misuse
Runway Gen-2	Video (text-to-video synthesis)	B-roll generation, scenario simulation	Innovative for fast, creative news visuals	Still evolving; occasional quality inconsistencies
Midjourney	Artistic-style images	Opinion pieces, abstract editorial illustrations	High aesthetic value; useful for op-eds	Less suitable for realistic or event-based imagery

3. Ethical frameworks and guiding principles

3.1. Key Ethical Pillars: Accuracy, Fairness, Accountability, and Transparency

AI-augmented journalism must rest on an ethical foundation that ensures the protection of public trust and democratic values. The four central pillars—accuracy, fairness, accountability, and transparency—serve as essential benchmarks for evaluating the responsible use of AI technologies in editorial settings. Accuracy refers to the fidelity of AI-generated content to verifiable facts. Without mechanisms to audit AI outputs, there is a risk that misleading or partially correct information could propagate through legitimate channels [9]. This is especially problematic in breaking news environments, where misinformation can spread rapidly before verification protocols catch up.

Fairness, meanwhile, pertains to the representational equity of AI tools in content production. Datasets used to train language models often reflect existing societal biases. If left unchecked, these systems can replicate or amplify discriminatory narratives that marginalize certain communities or perpetuate harmful stereotypes [10]. Ethical journalism must thus include strategies to identify and mitigate algorithmic bias before deployment.

Accountability in AI-assisted reporting demands clear assignment of responsibility when errors occur. It must be evident whether an editor, journalist, developer, or AI system bears the burden for inaccuracies or harm. To that end, many leading newsrooms have begun drafting AI responsibility charters, detailing human review procedures and roles [11].

Transparency requires full disclosure of AI involvement in journalistic work. Audiences should be informed when content has been produced, edited, or influenced by automated systems. This includes labeling AI-generated visuals, metadata trails, and source citations. Without transparency, media consumers cannot accurately assess credibility or intent, especially in an age of hyperrealistic synthetic content [12].

Upholding these four pillars is critical as newsrooms transition from experimental AI use to systemic integration. Ethical foresight will define whether AI augments journalism or undermines its foundational values.

3.2. AI Bias, Hallucination, and the Risks of Misinformation

The deployment of generative AI in journalism introduces new vulnerabilities around algorithmic bias, model hallucinations, and misinformation risks. Large Language Models (LLMs), while powerful, are not immune to embedding and amplifying the prejudices present in their training data. This creates structural inequities in news coverage, such as

reinforcing gendered language, underrepresenting minority perspectives, or mischaracterizing politically sensitive events [13].

Even more concerning are hallucinations, where models generate statements or entire passages that appear coherent but are factually incorrect. These errors are not just technical anomalies; they pose significant reputational and informational risks when embedded in news reporting [14]. For instance, an AI summarizing a legal case might fabricate non-existent precedents or misquote judicial rulings. If published without proper review, such outputs can contribute to public misperception and even legal liability.

These risks are magnified when AI is used to generate real-time or high-volume content, such as live election coverage, crisis updates, or public health alerts. In these contexts, misinformation spreads faster than it can be corrected—especially on digital platforms that prioritize engagement over verification [15].

Mitigation strategies include algorithmic audits, the use of fact-checking APIs, and hybrid pipelines where human editors systematically review AI outputs. Ultimately, the reputational trust of a media institution depends on its ability to detect and prevent AI-induced misinformation before it reaches the public.

3.3. Autonomy and the Role of Human Oversight in AI-Assisted Reporting

The rise of generative AI in journalism has sparked debate around the balance between automation and human editorial autonomy. While AI can streamline workflows, its increasing role in story framing, topic selection, and headline generation may inadvertently reduce the autonomy of journalists and editors in shaping narratives. The risk lies in delegating critical thinking processes to probabilistic systems optimized for scale and engagement rather than context and nuance [16].

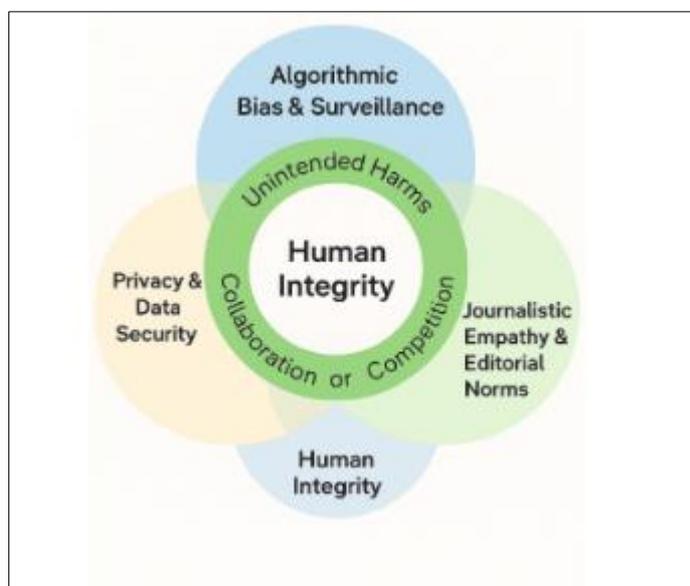


Figure 2 Ethical tension map in AI-generated newsrooms

A core ethical safeguard is the principle of human-in-the-loop oversight—ensuring that editorial decisions, especially those with societal consequences, remain under the control of qualified individuals. Journalists must retain authority to reject, revise, or contextualize AI-generated content. This practice reinforces accountability and ensures content aligns with journalistic standards and institutional values [17].

There is also a psychological dimension to this shift. As AI tools become more embedded in everyday reporting, journalists may develop a form of cognitive offloading—relying on systems to suggest angles, structure arguments, or generate quotes. While efficient, this could erode journalistic intuition and analytical depth, especially among early-career reporters trained in AI-dominated environments [18].

Furthermore, editorial diversity may suffer if AI systems become primary gatekeepers. Algorithms often prioritize topics and tones that mirror dominant media narratives, sidelining unconventional or minority viewpoints. Human

editors play a vital role in recognizing gaps, elevating underreported issues, and applying ethical judgment that machines lack [19].

To preserve autonomy, editorial policies must clearly define the scope and limits of AI in content creation. Training programs should empower journalists to interrogate AI outputs, understand model limitations, and retain editorial voice. Investment in explainable AI and transparency tools can support this balance, fostering a newsroom culture that views AI as a collaborator—not a substitute—for human reasoning and editorial integrity.

4. Impact on public trust and information integrity

4.1. Erosion of Source Credibility in the Age of Synthetics

The rise of generative AI technologies—especially large language models (LLMs), synthetic media engines, and text-to-image diffusion models—has complicated traditional definitions of journalistic credibility. Historically, trust in journalism has hinged on verifiable sourcing, journalistic rigor, and editorial accountability. However, in the AI era, these benchmarks are increasingly obscured by the automated nature of content production and the proliferation of highly convincing synthetic outputs [14].

AI-generated articles, while efficient, are often detached from traditional sourcing norms. These systems may draw from a wide corpus of public data but lack the transparency to trace individual claims back to credible, verified sources. In the absence of disclosure, the line between human and machine authorship becomes blurred, undermining audience confidence in the integrity of the reporting [15].

Moreover, synthetic content such as deepfake videos, audio clones, and image manipulations have accelerated the erosion of visual and auditory credibility. When even seemingly authentic audiovisual materials can be fabricated or altered at scale, audiences are left questioning the legitimacy of everything from breaking news clips to investigative documentaries [16].

Newsrooms that deploy AI without adequate labeling or audit mechanisms risk alienating their audiences. A lack of visible editorial intervention in AI-generated work further exacerbates distrust. Rebuilding credibility in this new era will require not just technological transparency, but an organizational commitment to reaffirming human editorial authority and content authenticity at every stage of production [17].

4.2. Public Reactions to AI-Generated News: Surveys and Perceptions

Understanding how the public perceives AI-generated journalism is critical for evaluating its societal viability. Surveys conducted by Reuters, Pew Research Center, and the Tow Center indicate a nuanced but cautious public sentiment. A growing segment of news consumers express curiosity about AI-enhanced media, appreciating its speed and breadth, but simultaneously harbor concerns over accuracy, transparency, and manipulation [18].

According to a 2024 Reuters Institute Digital News Report, nearly 60% of respondents in the U.S. said they would be less likely to trust a news article if they knew it had been generated or significantly written by AI. Interestingly, the distrust levels were higher among older and politically moderate audiences, whereas younger, tech-savvy users showed greater openness to AI-assisted reporting, provided transparency standards were upheld [19].

Audience perceptions are also strongly influenced by how clearly the AI's role is disclosed. Studies show that articles accompanied by disclaimers or metadata indicating AI involvement received significantly higher trust ratings than those that did not, even if the actual content was identical [20]. This underscores the public's demand for clarity, not just content quality.

Emotional tone plays a role as well. Readers tend to be more skeptical of AI-written opinion pieces or emotionally charged content than straightforward news summaries. There is a prevailing belief that AI lacks the moral and contextual judgment necessary for sensitive or nuanced reporting, such as coverage of war, racial injustice, or political dissent [21].

Public trust is further shaped by media literacy levels. Individuals familiar with the technical limitations of AI—such as its inability to verify real-time facts or avoid hallucinations—are more cautious consumers of AI-generated news. In contrast, less-informed audiences may either over-trust or completely dismiss AI journalism, depending on ideological predispositions [22].

To foster sustainable public engagement, newsrooms must incorporate audience feedback mechanisms, promote transparency by design, and empower journalists to guide AI narratives rather than surrender editorial control entirely.

4.3. Managing Deepfakes, Disinformation, and Satirical Content

AI's capacity to generate synthetic multimedia has blurred the line between journalism, disinformation, and satire. Deepfakes—AI-generated video and audio content mimicking real individuals—pose a severe threat to the credibility of journalistic outputs. In politically sensitive moments, such as elections or international crises, deepfakes have the potential to mislead millions within hours of dissemination, often long before fact-checking mechanisms can respond [23].

Satirical content, while historically a valid form of social commentary, becomes problematic when AI amplifies its realism. Satirical AI outputs, especially when taken out of context on social media platforms, are often misinterpreted as factual reporting. A 2025 study by Stanford's Center for Media and Democracy showed that more than 40% of surveyed users had mistaken AI-generated satire for genuine news at least once in the past year [24].

Managing these risks requires a combination of technological solutions and editorial interventions. From a tech standpoint, blockchain-based content authentication, digital watermarks, and zero-knowledge proofs are emerging as tools to validate original content provenance. Initiatives like the Content Authenticity Initiative (CAI) and Project Origin aim to create traceable records for visual and audio media, enabling news consumers to verify the source and edit history of digital assets [25].

From an editorial perspective, fact-checking workflows must evolve to address synthetic threats in real time. This includes automated detection of manipulated media, AI-literate journalists trained to spot inconsistencies, and cross-verification with original sources before publication. Furthermore, transparency labels—such as “AI-generated,” “satirical,” or “digitally altered”—must become a standardized part of digital publishing.

Ethically, media organizations must also resist the temptation to use AI for sensationalist or emotionally exploitative content, which may attract clicks but degrade trust. The responsibility lies not only in debunking falsehoods but in resisting their viral logic altogether.

As synthetic media becomes increasingly accessible and convincing, the integrity of journalism will hinge on its ability to anticipate and counteract manipulative AI usage while retaining its commitment to truth, clarity, and public service.

Table 2 Survey Results Comparing Trust in AI-Generated vs. Human-Written News

Survey Metric	AI-Generated News	Human-Written News
Perceived Trustworthiness (%)	38.5%	74.2%
Accuracy of Factual Reporting (%)	42.1%	81.7%
Emotional Resonance / Human Touch (%)	29.4%	86.5%
Clarity and Readability (%)	75.8%	77.9%
Bias Perception (Lower is Better) (%)	56.7%	48.2%
Support for Clear AI Disclosure Labels (%)	92.3%	N/A
Willingness to Read AI-Generated News Again (%)	47.6%	85.4%

5. Case studies and institutional responses

5.1. Use Case: AI-Generated Weather Reports and Election Coverage

AI is already playing an operational role in generating time-sensitive news content, with weather forecasts and election coverage emerging as two of the most established use cases. Automated weather reporting, for instance, leverages structured meteorological datasets to produce real-time updates tailored to local regions. Media organizations like the BBC and Germany's ARD deploy natural language generation (NLG) systems to create thousands of hyperlocal reports daily [17].

This application is generally well-received by audiences, as the content is highly factual, repetitive in structure, and largely free from interpretive bias. The limited editorial discretion required in weather reporting makes it a suitable domain for automation without sacrificing credibility or accuracy. Moreover, the integration of AI allows meteorologists to focus on severe weather analysis and emergency communications, while routine content is automatically published across platforms [18].

Election coverage represents a more complex frontier. The Associated Press (AP) uses AI models to automate the reporting of election results by scraping certified data from state electoral boards and converting it into region-specific stories [19]. These systems update in real time and are particularly effective during live coverage, reducing human fatigue and speeding delivery across multiple platforms.

However, election reporting introduces greater ethical stakes. Errors in real-time vote tallies or candidate projections—if derived from flawed AI outputs—can undermine democratic processes or fuel public distrust. To mitigate this, newsrooms often pair AI reporting with human verification layers, ensuring outputs are cross-checked before release. These practices demonstrate that while automation adds speed and scale, it must be counterbalanced with editorial oversight to maintain credibility.

5.2. Media Houses' Policies on Disclosure and Labeling (BBC, AP, NYT)

In response to growing concerns around AI transparency, major media organizations have introduced explicit policies on the disclosure and labeling of AI-generated content. The British Broadcasting Corporation (BBC), for example, publicly states that all AI-assisted outputs must be approved by an editor and clearly marked to inform audiences of automated involvement. This includes attribution statements like “generated with AI assistance,” prominently placed within the article or metadata [20].

The Associated Press (AP), one of the early adopters of automated journalism, similarly enforces editorial review protocols before AI-generated reports are published. AP maintains that AI should only be used to augment human reporting, not replace it, and mandates that its use be disclosed within the article body or accompanying credits [21]. These rules reflect the outlet's broader principles of trustworthiness, neutrality, and transparency in journalism.

The New York Times (NYT) has taken a more cautious approach. While acknowledging the potential of AI for data visualization and content generation, the NYT prohibits the unsupervised use of generative models in news production. The organization's editorial guidelines require all AI-generated suggestions—such as summaries or headline drafts—to be reviewed, edited, and explicitly acknowledged by human journalists before they are made public [22].

These media policies not only promote audience trust but also serve as institutional safeguards against liability and reputational damage. By establishing clear internal frameworks, these organizations create accountability mechanisms that balance technological innovation with ethical responsibility. Importantly, such transparency also empowers consumers to critically evaluate the origins and reliability of the content they consume, reducing confusion in an increasingly synthetic media landscape.

5.3. Role of Fact-Checking and Human-AI Collaboration Models

The exponential growth of AI-generated content has intensified the need for fact-checking mechanisms and collaborative oversight models in digital journalism. Traditional fact-checking, once reactive and manual, is now evolving into a hybrid system where human expertise is augmented by machine efficiency. Organizations like Full Fact, PolitiFact, and Snopes are increasingly incorporating AI to identify claim patterns, verify sources, and flag content inconsistencies across large volumes of information [23].

AI-powered fact-checkers use techniques such as Named Entity Recognition (NER) and knowledge graph mapping to match claims against structured databases like Wikidata, ProPublica, and legislative records. These systems can rapidly detect anomalies in quotes, numerical data, or public policy references. However, machine fact-checking is not infallible—it often fails to grasp contextual nuances, sarcasm, or evolving narratives in fast-moving news cycles [24].

To address this, several newsrooms are implementing human-AI collaboration frameworks, where journalists use AI tools to accelerate verification tasks but retain final editorial control. For example, The Washington Post's “ModBot” identifies statistical errors in political reporting, but all flagged outputs are reviewed by data journalists before corrections are issued [25].

Another innovation is the integration of AI into content moderation dashboards, where editorial teams can visualize algorithmic recommendations, confidence scores, and source trail information. This approach supports transparency and trains human editors to understand the inner workings of machine-generated outputs.

Importantly, human-AI collaboration also extends to audience interaction models. Platforms like The Guardian and El País are piloting reader feedback loops where users can challenge or comment on AI-generated facts, triggering real-time editorial reviews. This participatory layer strengthens democratic accountability while ensuring that AI tools do not function as black boxes.

Ultimately, the synergy between machine efficiency and human judgment remains essential for building resilient, transparent, and trustworthy AI-driven news ecosystems.

5.4. Regulatory Efforts: EU AI Act, FTC Guidelines, and Media Councils

Governments and regulatory bodies are now grappling with how best to govern the use of AI in journalism, aiming to strike a balance between innovation, public interest, and ethical safeguards. The European Union’s AI Act, passed in 2024, is the most comprehensive legal framework to date. It categorizes AI systems by risk levels and imposes transparency, auditability, and human oversight requirements for high-impact use cases, including media and content generation [26].

Under the Act, AI-generated news must carry visible disclosures, undergo routine algorithmic audits, and include human-in-the-loop mechanisms where decisions affect democratic discourse. Failure to comply may result in fines and license revocations, a signal that media applications are no longer exempt from broader AI governance concerns [27].

In the United States, the Federal Trade Commission (FTC) has issued advisory guidelines on the use of AI in commercial and informational contexts. While not yet law, these guidelines caution media organizations against misleading consumers with unlabeled AI content or false endorsements generated via automated systems. The FTC also emphasizes data protection, urging newsrooms to ensure that AI tools respect privacy and avoid unauthorized scraping of personal information [28].

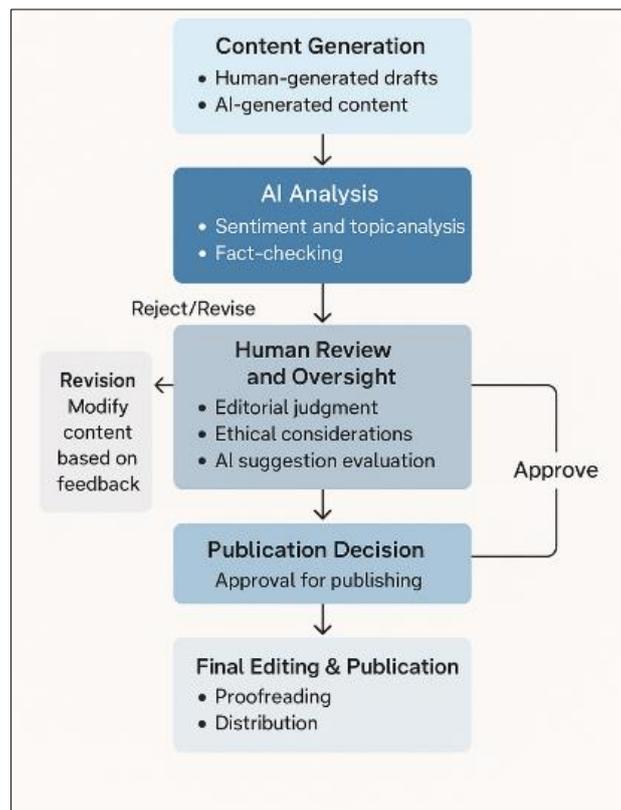


Figure 3 Workflow of a hybrid (human + AI) editorial process

In parallel, national press councils and professional associations—such as the Society of Professional Journalists and the Online News Association—are publishing voluntary ethical codes for AI use. These frameworks recommend best practices for source attribution, editorial transparency, and algorithmic fairness in newsroom settings.

Together, these regulatory and editorial interventions reflect a growing recognition of the need for structured oversight. As AI continues to permeate journalistic ecosystems, harmonizing legal compliance with institutional self-governance will be crucial for sustaining public trust and safeguarding the role of journalism in democratic societies.

Table 3 Institutional AI Editorial Policies and Risk Mitigation Actions

Institution	AI Use Policy Summary	Risk Mitigation Actions
BBC	Uses AI for metadata tagging, speech-to-text, and archive search. AI-generated content must be labeled.	Human editorial oversight is mandatory. Fact-checking remains manual. AI output is not used for news reporting.
Associated Press (AP)	Employs AI for earnings reports, sports recaps, and transcription assistance.	Provides internal training on AI ethics; all AI-generated text is reviewed by human editors.
Reuters	Uses AI for content recommendation, sentiment analysis, and topic tagging.	Deploys an AI audit framework; includes data ethics reviews and transparency scorecards.
The New York Times (NYT)	Pilots AI in newsroom tools but restricts generative use for reporting or narrative text.	Developed internal AI committee; requires explicit disclosures when AI tools are used.
The Guardian	Uses AI cautiously for automation and reader personalization. Editorial narrative generation is avoided.	Incorporates AI impact assessments; public-facing editorial AI principles published online.
Le Monde	Uses AI for multilingual article summaries and archive curation.	Requires attribution of all AI-generated summaries; uses human translators for sensitive topics.

6. Design and deployment considerations

6.1. Building Explainable and Auditable AI in Newsrooms

As artificial intelligence becomes increasingly embedded in newsroom operations, the need for explainability and auditability becomes paramount. AI models—especially those involved in generating or classifying news content—must offer transparency not only to technologists but to editorial staff, regulators, and the public [21]. In practice, explainability refers to the model's ability to present understandable reasoning for its outputs, such as why a headline was generated or why a news item was categorized in a specific tone or topic.

One of the primary challenges in journalism is the use of black-box models such as large language models (LLMs), which offer minimal interpretability by default. To address this, many newsrooms are now turning to post-hoc explanation tools such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations). These tools help editors interpret model decisions, enabling more informed oversight [22].

However, explanation alone is insufficient without a structured audit trail. AI systems in journalism must be auditable—capable of producing logs of data sources, intermediate model states, and editorial interventions. These trails enable retrospective analysis when content errors, ethical breaches, or misinformation occur. An example is the audit system used by Reuters AI Lab, which tracks metadata and model confidence scores for all generated content [23].

Establishing AI explainability and auditability ensures that journalistic organizations can uphold transparency and accountability—core tenets of media ethics—even as they adopt cutting-edge automation. These principles not only reinforce internal editorial discipline but also enable media outlets to respond effectively to external scrutiny, public inquiries, and regulatory evaluations.

6.2. Importance of Diverse Training Data and Cultural Sensitivity

The ethical design of AI in journalism is deeply influenced by the quality and diversity of training data. When AI models are trained predominantly on data from Western, English-speaking sources, they risk reproducing cultural biases, underrepresenting marginalized voices, and reinforcing global power asymmetries [24]. For a profession that relies on impartiality and social responsibility, such imbalances in training data are not merely technical shortcomings—they are editorial failures with real-world consequences.

For example, AI summarization models have been found to omit key nuances in stories from the Global South or misinterpret idiomatic expressions in minority languages. These limitations affect both representation and accuracy. Addressing this issue requires intentional diversification of training datasets to include multilingual corpora, regionally grounded narratives, and alternative media sources. News organizations such as Al Jazeera and Quartz Africa have made strides in this area by contributing localized data to open-source AI ecosystems [25].

Beyond linguistic and geographic representation, cultural sensitivity must be embedded in model design and application. Journalistic AI must recognize context-specific themes such as religion, gender identity, political ideologies, and indigenous histories. These elements shape how news is perceived and affect whether coverage fosters understanding or alienation. AI that fails to account for this complexity risks amplifying stereotypes or misrepresenting community narratives [26].

Integrating ethical review boards and community feedback loops into the AI development process can help newsrooms align machine outputs with diverse cultural expectations. Responsible AI design, therefore, is not only about reducing bias but also about ensuring epistemic justice—the fair recognition of different knowledge systems, identities, and perspectives in the media landscape.

6.3. Consent, Attribution, and Rights in AI Content Creation

One of the most pressing ethical concerns in AI-augmented journalism involves questions of consent, intellectual attribution, and content ownership. As AI systems harvest vast amounts of web data—news archives, blog posts, social media, podcasts, and multimedia content—to generate new material, the boundaries of legal and ethical reuse become blurred. Most generative AI models today lack mechanisms to distinguish between licensed, open-source, and copyrighted material in their training data, raising concerns over unauthorized content replication [27].

This challenge is particularly acute in journalism, where originality, source acknowledgment, and licensing agreements are foundational. When AI systems paraphrase or remix existing articles without attribution, they may inadvertently violate copyright laws or journalistic norms. Several lawsuits in the U.S. and EU have been launched by news organizations and individual authors against AI firms for unauthorized dataset usage. These cases underscore the urgency of defining data usage rights and reinforcing digital content protection frameworks [28].

Consent is another critical dimension, especially when AI systems extract content from private or semi-public domains such as comment threads, email leaks, or user-submitted news tips. If AI models process these inputs without explicit user consent, they may breach both journalistic ethics and privacy laws. Integrating informed consent protocols and data minimization strategies is essential for maintaining trust between media organizations and their audiences [29].

Furthermore, AI-generated content introduces a novel question: who owns the creative output? If a headline, report, or visual illustration is produced by an AI tool trained on someone else's work, how should credit and remuneration be allocated? Some media outlets are now experimenting with digital watermarking and smart attribution tokens, which track the lineage of content and automate revenue sharing among contributors, model developers, and publishers [30].

Moving forward, policy frameworks must balance innovation with the rights of authors, communities, and platforms whose work sustains AI systems. In journalism, where credibility is tied to responsible sourcing and ethical integrity, failing to address these concerns risks undermining both public trust and institutional legitimacy.

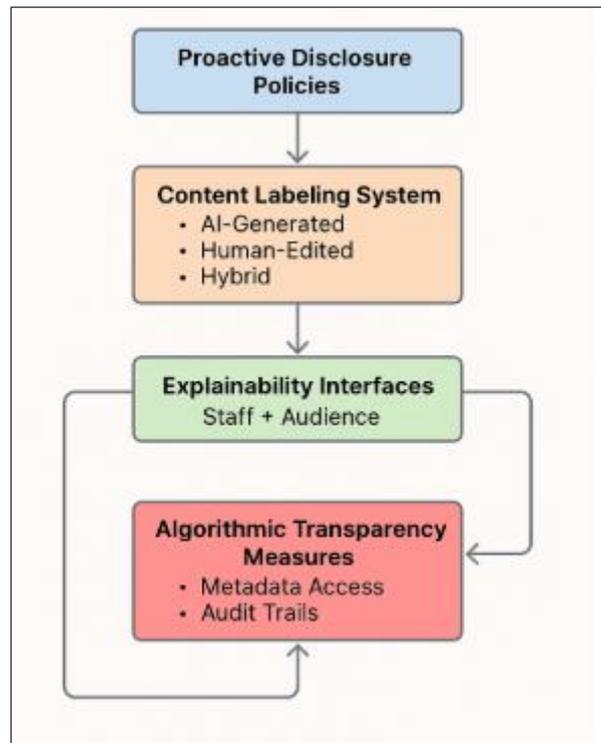


Figure 4 Transparency-by-design model for ethical AI use in journalism

7. Challenges, limitations, and unintended consequences

7.1. Over-Reliance on AI and Deskilling of Journalists

While artificial intelligence has offered efficiency gains in digital newsrooms, one of the most pressing concerns is the progressive deskilling of journalists. As machine-generated content becomes more prevalent, editors, reporters, and even investigative writers may become overly reliant on AI for functions such as data scraping, story drafting, or headline generation [24]. The result is a gradual erosion of core journalistic competencies like fact-checking, sourcing, and interpretive narrative crafting.

AI systems are currently being used to draft earnings reports, local crime summaries, and sports updates with limited human input. Although these domains are structured and data-heavy, their automation risks shifting editorial labor into mere oversight roles rather than creative or analytical journalism [25]. Younger journalists entering the profession may find themselves trained to operate and validate models rather than develop reporting skills. This raises serious concerns about the future of journalistic identity and professionalism.

Moreover, AI often fails in edge cases, satire detection, or emotionally nuanced topics, demanding human judgment to contextualize and validate stories. If such editorial intuition is not consistently exercised, media organizations may lose the depth and integrity traditionally associated with investigative journalism [26].

Some newsrooms have attempted to mitigate this trend by instituting hybrid roles—such as AI editors or human-in-the-loop content engineers—tasked with pairing journalistic training with machine fluency. However, unless these roles are protected and expanded, there remains a long-term risk of diminishing human craftsmanship in news creation. In essence, responsible AI adoption must avoid the unintended consequence of automating away the very skills that define journalism.

7.2. Deepfake Proliferation and Legal Uncertainty

The rise of deepfake technologies presents a particularly challenging gray zone for AI ethics in journalism. While manipulated media is not a new phenomenon, the realism achieved by modern generative adversarial networks (GANs) and diffusion models has reached a point where detection lags behind synthesis [27]. This imbalance leaves newsrooms, legal institutions, and platforms scrambling to define boundaries, ownership, and accountability.

Currently, few countries have enacted comprehensive legislation addressing the unauthorized use of synthetic media for impersonation or misinformation. In the U.S., legal frameworks like Section 230 of the Communications Decency Act provide immunity to platforms hosting third-party content, creating loopholes for malicious deepfake dissemination [28]. Journalistic institutions therefore bear the burden of verification without consistent legal recourse or enforcement mechanisms.

Deepfakes also introduce epistemic uncertainty: if audiences become skeptical of all visual content, genuine reporting suffers collateral damage. In politically sensitive contexts—such as elections, conflict zones, or diplomatic crises—deepfakes can be weaponized to erode public trust in legitimate journalism. The mere possibility that a video “could be fake” undermines journalistic authority, even in the absence of manipulation [29].

AI-assisted tools such as deepfake detectors, forensic watermarking, and blockchain-based provenance solutions are emerging to counter this trend. However, these technologies are reactive, often arriving after public perception has been influenced. Furthermore, they rely on computational power and model generalization that may not keep pace with the rate of synthetic content innovation [30].

Without a clear legal framework and standardized authentication systems, the journalistic ecosystem remains exposed. The gap between technological sophistication and regulatory adaptation leaves both newsmakers and the public vulnerable to disinformation with unprecedented realism and reach.

7.3. Algorithmic Manipulation, Censorship, and Editorial Bias

AI systems embedded in editorial workflows or content delivery platforms may inadvertently reinforce bias, censorship, or ideological filtering, either through flawed training data or through opaque algorithmic priorities. Recommendation engines—used by platforms like Facebook, YouTube, and Google News—play an increasingly significant role in shaping audience access to information. However, these systems optimize for engagement, not balance, often amplifying polarization or suppressing minority viewpoints [31].

In many cases, algorithmic sorting mechanisms prioritize content with high click-through potential, such as sensationalist or emotionally charged narratives, while deprioritizing nuanced or critical journalism. This commercial logic can subtly shape editorial direction, as newsroom metrics become tied to algorithmic performance. Over time, this fosters content homogeneity, editorial conformity, and self-censorship [32].

More concerning is the potential for algorithmic censorship, where AI moderation tools inadvertently suppress stories involving politically sensitive or controversial topics. Even when unintended, these decisions are often irreversible due to lack of transparency in model logic. When AI-driven content filters flag material as “unsafe,” “offensive,” or “low quality,” they may do so based on statistical patterns that reinforce dominant ideologies or suppress dissent [33].

Moreover, generative tools like ChatGPT or image synthesis platforms may exhibit ideological biases encoded in their training sets. This raises ethical questions about who defines objectivity in automated journalism and what safeguards exist to ensure pluralism in AI-generated narratives [34].

Editorial teams must audit not only what content AI produces, but also what content it excludes. Establishing independent algorithmic ethics boards, training models on diverse datasets, and offering editorial override mechanisms are steps toward ensuring that AI augments rather than distorts democratic discourse and journalistic plurality [35].

8. Future prospects and recommendations

8.1. Towards Human-AI Synergy: Augmentation, Not Replacement

The future of journalism lies not in the substitution of human editorial judgment with artificial intelligence, but in a synergistic relationship where AI amplifies journalistic capacity while preserving core human values. Instead of positioning AI as a replacement, industry leaders are beginning to conceptualize it as an augmentation tool—empowering journalists to handle large data volumes, uncover hidden patterns, and produce richer, more contextualized narratives [36].

To operationalize this vision, AI models must be explicitly designed to complement journalistic intuition. For example, sentiment analysis tools can assist in gauging public mood during election cycles, but the interpretation and framing of

these insights should remain firmly in human hands. Similarly, AI-driven trend detectors may flag emerging issues, but reporters must decide which angles warrant deeper inquiry and how to navigate the surrounding ethical terrain [28].

This symbiotic model requires structural shifts: cross-functional teams combining data scientists, editors, ethicists, and engineers must become standard in major newsrooms. Furthermore, feedback loops between machine-generated content and human oversight should be formalized through version control, metadata logging, and transparent correction mechanisms [29].

Ultimately, journalism's enduring mission—to inform, challenge power, and foster civic discourse—can only be realized when AI serves as an enabler of excellence rather than a substitute for human conscience. True human-AI synergy must therefore embed ethics by design, ensuring that automation enhances agency rather than erodes it.

8.2. Policy Recommendations for Media Houses, Regulators, and Tech Firms

Establishing responsible AI journalism practices requires multi-stakeholder coordination between news organizations, regulatory bodies, and technology companies. At the newsroom level, media houses must institute editorial AI charters—formal documents outlining acceptable uses, transparency standards, and human-in-the-loop governance policies. These frameworks should be publicly accessible and revised periodically to reflect new technological and ethical challenges [30].

A second recommendation for media institutions is to mandate source transparency for AI-generated content. Readers should be able to clearly distinguish between human-written and machine-assisted content through standardized disclosures, iconography, or metadata tags. Several outlets, including BBC and The Guardian, have begun experimenting with such practices, though no global consensus yet exists [31].

Regulators must also step in to provide legal guardrails that prevent misuse without stifling innovation. This includes developing statutes on synthetic media labeling, clarifying liability for AI-generated misinformation, and establishing privacy standards for training data derived from public and semi-private sources. The European Union's AI Act and the U.S. Federal Trade Commission's guidance are early efforts in this domain, but sector-specific regulations tailored to journalism are still lacking [32].

For technology companies, the responsibility lies in building value-aligned AI tools that offer explainability, allow human overrides, and comply with editorial contexts. Tech firms should also participate in industry-wide partnerships with journalists to co-design tools that reflect the profession's unique demands. Shared development ecosystems—such as the Partnership on AI or the JournalismAI initiative—serve as critical conduits for this collaboration [33].

Only through this triangulated effort can we construct a digital news ecosystem that is transparent, accountable, and resilient, upholding both innovation and democratic values.

8.3. Research Gaps and Future Study Directions

Despite significant progress, major research gaps persist in understanding and governing AI in journalism. First, empirical studies assessing audience trust in AI-mediated news across different demographics, cultures, and political contexts are sparse. Longitudinal data on how machine-assisted journalism affects public knowledge formation would help bridge this gap [34].

Second, more work is needed to quantify AI-induced editorial drift—the subtle influence algorithms may have on story framing, source prioritization, or tone. Ethnographic studies inside AI-equipped newsrooms could illuminate these patterns.

Third, interdisciplinary inquiry must explore how AI ethics frameworks align (or clash) with journalistic codes of conduct, especially in global South contexts where media systems face distinct political and infrastructural challenges.

Lastly, future research should test scalable models of collaborative AI design, involving journalists from inception to deployment. Doing so can ensure tools remain grounded in real-world editorial needs, not abstract technical ambitions.



Figure 5 Recommended ethical AI governance model for journalism organizations

9. Conclusion

9.1. Safeguarding Truth in the Age of Algorithmic Journalism

The accelerated integration of artificial intelligence into journalism has reshaped the foundations of news production, distribution, and consumption. From generative models drafting articles to recommendation systems curating headlines, AI technologies are now embedded in every phase of the editorial process. While these systems have delivered impressive efficiencies—enabling rapid content generation, improved audience targeting, and scalable investigative reporting—they also introduce complex ethical, operational, and societal challenges that cannot be ignored.

One of the most pressing conclusions from this study is the necessity of maintaining editorial integrity in the face of algorithmic mediation. Human oversight must remain a non-negotiable element in AI-assisted reporting. While machines can process vast quantities of information, identify patterns, and generate preliminary content, the final responsibility for truth-telling, contextual framing, and ethical decision-making must lie with trained journalists. Automation should never become an excuse to offload moral judgment or reduce the quality of information delivered to the public.

Equally critical is the role of transparency in rebuilding and preserving public trust. As audiences become increasingly exposed to AI-generated or AI-curated content, clear disclosure practices must be adopted. Readers have the right to know whether the article they are reading was drafted by a machine, edited by a human, or generated through collaborative processes. Transparency extends not only to content origins but also to editorial algorithms, moderation policies, and data collection practices. If the architecture of digital journalism remains opaque, it will only deepen skepticism, polarization, and misinformation.

Furthermore, the research underscores that bias and exclusion can be unintentionally embedded into AI systems through skewed datasets or uncritical design practices. If these systems are trained on historically biased media or lack input from diverse linguistic, cultural, and ideological sources, they will inevitably reflect and reinforce these distortions. Thus, diversity must be embedded not just in newsrooms, but in model architecture, training data, and evaluative frameworks.

Another major insight is the danger of editorial deskilling and over-reliance on automation. While AI can enhance productivity, it should not supplant the very skills that make journalism a craft—curiosity, empathy, skepticism, and ethical reasoning. Newsrooms must invest in continuous learning, equipping journalists with AI literacy while ensuring that these tools empower rather than diminish their professional agency.

The study also revealed regulatory fragmentation and lag in addressing the ethical complexities of AI in journalism. Current legal frameworks are insufficient to address the challenges posed by synthetic media, misinformation, and algorithmic filtering. Therefore, multi-sector policy coordination is essential. Regulators, technologists, journalists, and civil society must co-create standards that reflect the unique public interest role of journalism.

In looking ahead, the future of journalism demands a hybrid human-AI paradigm—one that balances innovation with responsibility, and scalability with accountability. The ethical path forward lies not in resisting technology but in shaping its trajectory through values-aligned design, inclusive collaboration, and informed regulation. Journalism's mission in a democratic society—to inform, scrutinize, and elevate truth—must not be compromised in pursuit of speed or novelty.

Ultimately, the true test of AI in journalism will not be measured by how efficiently it produces content, but by how faithfully it upholds the principles of accuracy, fairness, transparency, and human dignity. This requires an unwavering commitment to ethical integrity at every level—from the developers building AI systems to the editors deploying them in the newsroom. In this critical moment, the future of journalism hinges not on what AI can do, but on what humanity chooses to preserve.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

- [1] Leiser MR. Bias, journalistic endeavours, and the risks of artificial intelligence. In *Artificial intelligence and the media* 2022 Feb 15 (pp. 8-32). Edward Elgar Publishing.
- [2] Adeoluwa Abraham Olasehinde, Anthony Osi Blessing, Adedeji Adebola Adelagun, Somadina Obiora Chukwuemeka. Multi-layered modeling of photosynthetic efficiency under spectral light regimes in AI-optimized indoor agronomic systems. *International Journal of Science and Research Archive*. 2022;6(1):367–385. doi: 10.30574/ijrsra.2022.6.1.0267
- [3] Jones B, Jones R, Luger E. AI 'Everywhere and Nowhere': addressing the AI intelligibility problem in public service journalism. *Digital Journalism*. 2022 Nov 26;10(10):1731-55.
- [4] Pujari T, Goel A, Kejriwal D. Ethical and responsible AI in the age of adversarial diffusion models: Challenges, risks, and mitigation strategies. *International Journal Science and Technology*. 2022 Dec 25;1(3):54-68.
- [5] Veluru CS. Harnessing AI for Transformative Journalism: Advancements in Automated Content Creation and Analysis. *European Journal of Advances in Engineering and Technology*. 2022;9(9):85-92.
- [6] Tyler V. SYNTHETIC TRUTH OR FABRICATED LIES? NAVIGATING MISINFORMATION RISKS IN DIFFUSION-POWERED CONTENT GENERATION.
- [7] Mahama T. Assessing cancer incidence rates across age groups using stratified sampling and survival analysis methods. *Int J Comput Appl Technol Res* [Internet]. 2022;10(12):335–49. Available from: <https://doi.org/10.7753/IJCATR1012.1008>
- [8] Chukwunweike J. Design and optimization of energy-efficient electric machines for industrial automation and renewable power conversion applications. *Int J Comput Appl Technol Res*. 2019;8(12):548–560. doi: 10.7753/IJCATR0812.1011.
- [9] O'Hare JJ, Fairchild A, Ali U. Convergence and Disruption in Digital Society--Money, Secure Communication, Digital Objects and Generative AI in Spatial Mixed Reality. arXiv preprint arXiv:2207.09460. 2022 Jul 19.
- [10] Vaccari C, Chadwick A. Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social media+ society*. 2020 Feb;6(1):2056305120903408.

- [11] Panda G, Upadhyay AK, Khandelwal K. Artificial intelligence: A strategic disruption in public relations. *Journal of Creative Communications*. 2019 Nov;14(3):196-213.
- [12] Schudson M. The objectivity norm in American journalism. *Journalism*. 2001 Aug;2(2):149-70.
- [13] Thomson TJ, Angus D, Dootson P, Hurcombe E, Smith A. Visual mis/disinformation in journalism and public communications: Current verification practices, challenges, and future opportunities. *Journalism Practice*. 2022 May 28;16(5):938-62.
- [14] Tiwari TV. Unlocking the Future: Generative AI's Impact on Business.
- [15] Fletcher J. Deepfakes, artificial intelligence, and some kind of dystopia: The new faces of online post-fact performance. *Theatre Journal*. 2018;70(4):455-71.
- [16] Vizoso Á, Vaz-Álvarez M, López-García X. Fighting deepfakes: Media and internet giants' converging and diverging strategies against hi-tech misinformation. *Media and Communication*. 2021 Mar 3;9(1):291-300.
- [17] Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: Mapping the debate. *Big Data & Society*. 2016 Nov;3(2):2053951716679679.
- [18] Belosludtsev A, Dziuba E. Generative Artificial Intelligence in the System of International Relations: Risks, Opportunities, and Regulations. In *International Conference on Topical Issues of International Political Geography 2022* Dec 23 (pp. 187-200). Cham: Springer Nature Switzerland.
- [19] Yu S, Carroll F. Implications of AI in national security: understanding the security issues and ethical challenges. In *Artificial intelligence in cyber security: Impact and implications: Security challenges, technical and ethical issues, forensic investigative challenges 2022* Jan 1 (pp. 157-175). Cham: Springer International Publishing.
- [20] Yigitcanlar T, Corchado JM, Mehmood R, Li RY, Mossberger K, Desouza K. Responsible urban innovation with local government artificial intelligence (AI): A conceptual framework and research agenda. *Journal of Open Innovation: Technology, Market, and Complexity*. 2021 Feb 24;7(1):71.
- [21] Enemosah A, Chukwunweike J. Next-Generation SCADA Architectures for Enhanced Field Automation and Real-Time Remote Control in Oil and Gas Fields. *Int J Comput Appl Technol Res*. 2022;11(12):514-29. doi:10.7753/IJCATR1112.1018.
- [22] Graefe A, Haim M, Haarmann B, Brosius HB. Readers' perception of computer-generated news: Credibility, expertise, and readability. *Journalism*. 2018 May;19(5):595-610.
- [23] Hughes RT, Zhu L, Bednarz T. Generative adversarial networks-enabled human-artificial intelligence collaborative applications for creative and design industries: A systematic review of current approaches and trends. *Frontiers in artificial intelligence*. 2021 Apr 28;4:604234.
- [24] Cordeiro JV. Digital technologies and data science as health enablers: an outline of appealing promises and compelling ethical, legal, and social challenges. *Frontiers in medicine*. 2021 Jul 8;8:647897.
- [25] Zubair T, Raquib A, Qadir J. Combating Fake News, Misinformation, and Machine Learning Generated Fakes: Insight's from the Islamic Ethical Tradition. *Icr Journal*. 2019 Dec 15;10(2):189-212.
- [26] Corbeil JR, Corbeil ME. Teaching and Learning in the Age of Generative AI.
- [27] Pihlajarinne T, Alén-Savikko A, editors. *Artificial intelligence and the media: Reconsidering rights and responsibilities*. Edward Elgar Publishing; 2022 Feb 22.
- [28] Gianni R, Lehtinen S, Nieminen M. Governance of responsible AI: From ethical guidelines to cooperative policies. *Frontiers in Computer Science*. 2022 May 24;4:873437.
- [29] Lacity MC. Fake news, technology and ethics: Can AI and blockchains restore integrity?. *Journal of Information Technology Teaching Cases*. 2022 Oct 4;12(2):121-34.
- [30] Feher K, Katona AI. Fifteen shadows of socio-cultural AI: A systematic review and future perspectives. *Futures*. 2021 Sep 1;132:102817.
- [31] Geis JR, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, Langer SG, Borondy Kitts A, Birch J, Shields WF, van den Hoven van Genderen R. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *Radiology*. 2019 Nov;293(2):436-40.

- [32] Weber-Lewerenz B. Corporate digital responsibility (CDR) in construction engineering—ethical guidelines for the application of digital transformation and artificial intelligence (AI) in user practice. *SN Applied Sciences*. 2021 Oct;3:1-25.
- [33] Just N, Latzer M. Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, culture & society*. 2017 Mar;39(2):238-58.
- [34] Tom E, Keane PA, Blazes M, Pasquale LR, Chiang MF, Lee AY, Lee CS, Force AA. Protecting data privacy in the age of AI-enabled ophthalmology. *Translational vision science & technology*. 2020 Jan 28;9(2):36-.
- [35] Sraml Gonzalez J, Gulbrandsen M. Innovation in established industries undergoing digital transformation: the role of collective identity and public values. *Innovation*. 2022 Jan 2;24(1):201-30.
- [36] Amato G, Behrmann M, Bimbot F, Caramiaux B, Falchi F, Garcia A, Geurts J, Gibert J, Gravier G, Holken H, Koenitz H. AI in the media and creative industries. *arXiv preprint arXiv:1905.04175*. 2019 May 10.