Available online at www.sserr.ro
Social Sciences and Education Research Review

(8) 1 301 - 311 (2021)

ISSN 2393–1264
ISSN–L 2392–9863

DOI: 10.5281/zenodo.15251980

AI IS LEARNING HOW TO WRITE. ETHICAL PROBLEMS FOR JOURNALISM

Dan Valeriu VOINEA
University of Craiova

Abstract

The integration of Artificial Intelligence (AI) into journalism, particularly through advancements in natural language generation, presents significant opportunities and complex ethical challenges. This paper examines the ethical problems arising from AI-generated content in news production, drawing on established media ethics principles and emerging AI ethics frameworks. Key issues identified include the potential for disseminating misinformation due to AI "hallucinations," the risk of perpetuating or amplifying societal biases embedded in training data, the critical need for transparency and disclosure regarding AI authorship, complexities surrounding accountability for algorithmic outputs, and concerns about labor displacement and the changing roles of journalists. Early examples, such as automation by the Associated Press and errors following AI adoption at Microsoft's MSN, illustrate these tensions. The analysis emphasizes that traditional journalistic values—accuracy, fairness, accountability, transparency—remain paramount. It argues for robust human oversight, treating AI as a tool requiring verification and editorial judgment, rather than an autonomous author. Suggestions for navigating the future include developing dynamic ethical guidelines, enhancing AI literacy through training,

fostering cross-disciplinary collaboration, prioritizing AI applications that augment human capabilities, engaging proactively with regulation, and maintaining an unwavering focus on audience trust. The paper concludes that conscientious, ethical integration is crucial for harnessing AI's benefits while safeguarding journalistic integrity.

Keywords: *Artificial Intelligence, AI, Journalism Ethics, Algorithmic Journalism, Automated Journalism, Media Ethics, Misinformation, Bias, Transparency, Accountability, Labor Displacement*

Introduction

The increasing sophistication of artificial intelligence (AI) presents profound implications for the field of journalism, particularly through algorithms capable of generating written content. News organizations have adopted AI, leveraging natural language generation (NLG) for tasks ranging from automating routine, data-driven reports—such as weather updates, sports recaps, and financial summaries—to early experiments with more narrative AI-written stories. Landmark examples trace back several years, including the Los Angeles Times' "Quakebot," which automatically generated brief news alerts about seismic activity from 2014 onwards, and the Associated Press's (AP) significant investment in automation technology, also starting in 2014. AP's use of Automated Insights allowed it to produce thousands of corporate earnings reports quarterly, a dramatic increase from previous human capacity, thereby freeing journalists for more analytical and investigative work (Colford, 2014). The subsequent evolution of AI, especially the advent of powerful large language models like OpenAI's GPT-3 (released in 2020), has exponentially increased the potential for AI to produce fluent, human-like text across diverse topics, accelerating experimentation and adoption within newsrooms.

However, this technological advancement brings a complex host of ethical dilemmas. Concerns center on the potential for AI systems, if not

meticulously checked, to generate and spread misinformation or misleading content (Marconi, 2020; Wu et al., 2021). Furthermore, algorithms can inadvertently reproduce or even amplify biases present in their vast training datasets, raising fairness concerns. Questions of transparency and potential deception also arise: should audiences be explicitly informed about AI authorship? Who bears responsibility if an AI generates flawed or harmful content—the developer, the news outlet, or the deploying editor? Compounding these issues are anxieties about labor impacts, particularly the potential for AI to displace human journalists in certain roles. The 2020 decision by Microsoft to replace dozens of human news curators with an AI system, which subsequently made errors such as misidentifying individuals in photos, starkly illustrated these intertwined risks of job loss and algorithmic fallibility (Waterson, 2020; Jamie, 2020). While AI offers undeniable efficiency gains, its ethical deployment demands careful navigation, ensuring alignment with journalism's foundational principles: truth, accuracy, fairness, independence, and accountability (SPJ, 2014).

Ethical Frameworks and Emerging Issues

Navigating the ethical landscape of AI in journalism requires drawing upon both established media ethics and emerging principles from AI ethics. Longstanding journalistic codes, such as the Society of Professional Journalists (SPJ) Code of Ethics (2014), provide essential guidance. Core tenets like "Seek Truth and Report It," "Minimize Harm," "Act Independently," and "Be Accountable and Transparent" are directly applicable to content regardless of its origin, human or algorithmic. Social Responsibility Theory, which posits that media have an obligation to serve the public good and operate ethically, further underscores the need for critical evaluation and responsible implementation of AI tools.

Concurrently, the field of AI ethics offers relevant frameworks emphasizing values like fairness (avoiding bias), accountability (assigning

responsibility), interpretability (understanding AI decisions), and reliability. The concept of algorithmic accountability, as articulated by scholars like Diakopoulos (2015), is particularly pertinent. It stresses that human creators, editors, and news organizations—not the algorithms themselves—must remain responsible for the outputs and impacts of AI systems. This perspective aligns with the consensus in both academic literature and early newsroom guidelines: AI, lacking consciousness or moral intent, functions as a tool, and the ethical burden remains squarely on the humans who design, train, deploy, and oversee it (Dörr, 2016; Zhang & Ting, 2021; van Dalen, 2021).

Synthesizing these perspectives reveals several key ethical challenges prominent in literature and early practice:

Misinformation and Accuracy: A primary concern is the propensity of AI models, particularly large language models, to generate "hallucinations"—statements that sound authoritative but are factually incorrect or nonsensical. Because these models predict text based on patterns rather than understanding truth, publishing their output without rigorous human verification poses a significant risk of disseminating misinformation. This directly contravenes journalism's fundamental commitment to accuracy (Marconi, 2020). Early experiments and uses of AI in news generation, even when subject to editorial review, demonstrated that subtle inaccuracies could slip through, necessitating corrections and underscoring the critical need for robust, perhaps even enhanced, fact-checking protocols specifically designed for AI-generated content. The core challenge is ensuring that any content published under a news organization's banner meets traditional standards of veracity, irrespective of its automated origins.

Bias and Fairness: AI systems learn from the data they are trained on. If this data reflects existing societal biases—related to race, gender, geography, ideology, or other factors—the AI is likely to replicate and potentially amplify those biases in its output (Smith, 2018; Bender et al., 2021). This can manifest

subtly, through biased language or skewed framing, or more overtly, such as AI systems demonstrating a lack of cultural sensitivity or failing to accurately represent or distinguish between individuals from minority groups, as highlighted by the MSN photo incident (Jamie, 2020). Algorithmic bias undermines journalistic goals of impartiality, fairness, and equitable representation. Addressing this requires proactive measures throughout the AI lifecycle, including careful curation and auditing of training data, ongoing testing of AI outputs for bias, fostering diversity within the teams developing and deploying AI, and maintaining vigilant human oversight focused on fairness considerations.

Transparency and Disclosure: A cornerstone of journalistic ethics is transparency—being open with the audience about newsgathering and production processes. Applied to AI, this principle mandates disclosing when content is generated or significantly assisted by algorithms (Hmmelmann, 2019; Chen et al., 2021). Failing to do so can be seen as deceptive, misleading readers about the nature and potential limitations of the information presented. While early studies on audience perceptions of automated content yielded mixed results, with some finding little difference in perceived credibility but noting a lack of stylistic appeal (Clerwall, 2014), the prevailing ethical consensus strongly favors clear labeling. Disclosure respects audience autonomy, allowing readers to apply appropriate context or skepticism, and fosters accountability by making AI use an open practice subject to public scrutiny. Attempting to hide AI involvement risks severe reputational damage if discovered.

Accountability and Control: Determining responsibility for AI-generated content is crucial. Ethical frameworks and emerging best practices firmly place accountability on human actors—the editors, journalists, and the news organization itself—rather than the algorithm (Diakopoulos & Koliska, 2017). AI systems should be treated as tools, powerful assistants whose work requires verification and approval. Establishing clear editorial workflows, where designated humans vet and take ownership of AI outputs before publication, is

essential. Using AI as a scapegoat for errors ("the algorithm did it") is ethically untenable. Accountability also involves due diligence in selecting and implementing AI tools, understanding their limitations, and establishing clear internal guidelines that reinforce ultimate human editorial control over all published content.

Employment and Professional Roles: The potential for AI to automate tasks previously performed by humans raises significant concerns about job displacement within journalism, particularly for roles involving routine data processing or writing (Feng et al., 2021). Microsoft's 2020 decision to replace human editors with AI served as a stark example of this possibility, sparking unease about the future of journalistic labor (Waterson, 2020). From an ethical standpoint, the ideal integration of AI involves augmentation—using technology to handle repetitive tasks, thereby freeing human journalists for more complex, creative, and investigative work that requires critical thinking, ethical judgment, and human empathy, as demonstrated in the AP's initial automation strategy (Colford, 2014). A purely cost-driven approach to replacing humans risks devaluing journalistic labor and potentially diminishing news quality. The transition necessitates ethical consideration of the workforce, including potential retraining and redefining roles towards human-AI collaboration.

Discussion and Conclusion

The integration of AI into journalism signifies a critical juncture, offering pathways towards both enhanced efficiency and potential ethical compromise. Unchecked deployment risks eroding public trust through the propagation of errors, the amplification of bias, or a lack of transparency. Conversely, ethically implemented AI holds the potential to broaden coverage and free human journalists for higher-value work. Navigating this requires steadfast adherence to the core ethical tenets of journalism: accuracy, fairness, accountability, and transparency must guide every step of AI adoption.

Preserving public trust in an era of AI-assisted news necessitates proactive and principled action. Clear, conspicuous disclosure of AI involvement in content creation is paramount. While potentially met with initial audience skepticism, honesty is ultimately less damaging than deception discovered later. Robust human oversight and editorial control are non-negotiable; AI must remain a tool, subordinate to human judgment, with journalists and editors retaining full accountability for all published material. To operationalize this, news organizations should develop and enforce explicit AI usage policies detailing guidelines for disclosure, oversight, bias mitigation, verification protocols, and acceptable use cases. Collaboration across the industry, through professional associations and initiatives, can further help by establishing shared standards, disseminating best practices, and providing collective leverage when engaging with technology providers.

While comprehensive regulatory frameworks specifically governing AI in journalism are still evolving (Helberger et al., 2020), ethical self-regulation by the industry is crucial in the interim. The strategic focus should be on leveraging AI to augment human capabilities—automating routine tasks to empower journalists to pursue in-depth reporting, analysis, and storytelling—rather than viewing AI primarily as a means to reduce labor costs at the expense of quality or ethical diligence. Investing in training and fostering AI literacy among journalists is essential, equipping them to work effectively, critically, and ethically alongside these powerful new tools.

In conclusion, AI presents transformative potential for journalism, but its ethical integration demands conscious effort and vigilance. By rigorously applying journalistic values, maintaining unwavering human accountability, embracing transparency, and investing in robust ethical frameworks and workforce training, the news industry can strive to harness AI's benefits without sacrificing the integrity, quality, and societal trust upon which responsible journalism depends. The fundamental challenge is to ensure that technological

innovation serves, rather than subverts, the core mission of providing reliable, fair, and accurate information essential to democratic life. Continued research is vital to monitor evolving audience perceptions, assess the efficacy of different oversight mechanisms, and understand the complex, dynamic relationship between journalists and AI tools in newsrooms globally.

Future Developments and Suggestions

Looking ahead, the trajectory of AI in journalism will depend on both technological advancements and the strategic choices made by the industry. Several potential developments and actionable suggestions emerge:

Advancing AI Capabilities: Future AI models may become more accurate, context-aware, and capable of citing sources reliably, potentially mitigating some current accuracy concerns. Development of specialized AI tools for journalistic tasks like verification, data analysis, and bias detection could further enhance ethical deployment.

Dynamic Ethical Guidelines: News organizations should treat AI ethics policies not as static documents but as living guidelines, regularly reviewed and updated to reflect technological changes, emerging best practices, and lessons learned from real-world applications. Establishing internal ethics committees or roles focused on AI can facilitate this.

Enhanced AI Literacy and Training: Continuous education for journalists is vital. Training should go beyond basic operation to cover the underlying mechanics of AI, common pitfalls (like hallucination and bias), verification techniques for AI output, and the ethical implications of different AI applications in the news workflow. Journalism schools also need to integrate AI literacy into their curricula.

Cross-Disciplinary Collaboration: Stronger partnerships between newsrooms, AI developers, ethicists, and academic researchers are needed. This collaboration can help ensure AI tools are designed with journalistic needs and

ethical considerations in mind from the outset, rather than being retrofitted later. Feedback loops from journalists using AI tools are crucial for iterative improvement.

Focus on Augmentation, Not Just Automation: The most ethically sound and potentially most valuable approach involves using AI to augment human journalists' capabilities. This means prioritizing AI applications that handle laborious tasks (data sifting, transcription, initial drafting from structured data) to free up human reporters for critical thinking, investigation, interviewing, nuanced analysis, and storytelling – tasks requiring human judgment and empathy.

Proactive Regulatory Engagement: The journalism industry should actively engage with policymakers to help shape sensible regulations around AI. This includes advocating for rules that mandate transparency (e.g., clear labeling) across the information ecosystem while ensuring that regulations do not stifle beneficial journalistic innovation.

Exploring New Journalistic Forms: AI may enable novel forms of journalism, such as highly personalized news experiences, interactive data visualizations generated on the fly, or large-scale investigations analyzing datasets previously too vast for human teams alone. Ethical frameworks must evolve to encompass these new possibilities.

Prioritizing Audience Trust: Ultimately, all efforts must center on maintaining and building audience trust. This requires consistent transparency, demonstrable accuracy in AI-assisted content, swift accountability for errors, and a clear commitment to using AI in service of quality journalism rather than as a replacement for it. Ongoing research into audience perceptions of AI in news and effective communication strategies is warranted.

By embracing these suggestions and anticipating future developments with a commitment to ethical principles, the journalism community can navigate

the complexities of AI integration and work towards a future where technology supports and enhances the vital role of journalism in society.

References

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623.

Chen, V. Y., Huang, J., & Su, C. (2021). Is the article generated by AI? The effect of AI authorship disclosure on news credibility. Computers in Human Behavior Reports, 7, 100214.

Clerwall, C. (2014). Enter the Robot Journalist: Users' perceptions of automated content. Journalism Practice, 8(5), 519-531.

Colford, P. (2014, July 15). A leap forward in quarterly earnings stories. AP Definitive Source.

Diakopoulos, N. (2015). Algorithmic Accountability: Journalistic investigation of computational power structures. Digital Journalism, 3(3), 398-415.

Diakopoulos, N., & Koliska, M. (2017). Algorithmic transparency in the news media. Digital Journalism, 5(7), 809-828.

Dörr, K. N. (2016). Mapping the field of algorithmic journalism. Digital Journalism, 4(6), 700-722.

Feng, Y., Cambria, E., & Li, Q. (2021). Artificial Intelligence in Media: How AI is Changing the Media Landscape. IEEE Computational Intelligence Magazine, 16(3), 10-17.

Helberger, N., Leerssen, P., & Dolata, U. (2020). The Algorithmic Ecosystem: The Interplay of Content Moderation, Recommendation Systems, and Advertising Platforms. Internet Policy Review, 9(4).

Hmmelmann, N. (2019). Deepfakes: A Looming Challenge for Privacy, Democracy, and National Security. Belfer Center for Science and International Affairs.

Jamie, P. (2020, June 1). Jade Thirlwall slams MSN for using wrong picture of Little Mix star. Mirror.

Marconi, F. (2020). Newsmakers: Artificial Intelligence and the Future of Journalism. Columbia University Press.

Smith, R. (2018). Algorithmic Bias in Crime Reporting.

Society of Professional Journalists (SPJ). (2014). SPJ Code of Ethics. https://www.spj.org/ethicscode.asp

van Dalen, A. (2021). Considering the 'human touch': An analysis of how journalists frame the implications of automated journalism. Digital Journalism, 10(1), 117-135.

Waterson, J. (2020, May 30). Microsoft sacks journalists to replace them with robots. The Guardian. https://www.theguardian.com/technology/2020/may/30/microsoft-sacks-journalists-to-replace-them-with-robots

Wu, T., Ribeiro, M. T., Heer, J., & Weld, D. S. (2021). AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–19.

Zhang, Y., & Ting, K. F. (2021). Human-machine communication research: Conceptual foundations, key concepts, methods, and future directions. Annals of the International Communication Association, 46(3), 161-180.