The Ethics of AI-Generated Literature: Originality, Authorship, and Literary Value in the Age of Large Language Models

Lima Antony

Professor, Department of English, St.Xavier's College for Women (Autonomous),Aluva, India

Abstract

The advent of sophisticated large language models capable of generating poetry and fiction has fundamentally challenged traditional conceptions of authorship, originality, and literary value. This paper examines whether AI-generated literature can be considered original creative work and analyzes the implications for authorship in literary theory and practice. Through an examination of recent empirical studies, legal frameworks, and theoretical foundations from Roland Barthes's "Death of the Author" to contemporary digital humanities scholarship, this analysis reveals that AI-generated literature exists in a liminal space between human creativity and algorithmic reproduction. While AI systems can produce texts that are indistinguishable from human-written works to non-expert readers, questions of originality and authorship remain deeply problematic. The paper argues that AI-generated literature challenges but does not necessarily invalidate traditional notions of authorship, instead requiring new frameworks for understanding creative collaboration between humans and machines. The implications extend beyond literary theory to questions of intellectual property, educational integrity, and the fundamental nature of creative expression in the digital age.

Introduction

The question of what constitutes authentic literature has never been more pressing than in our current moment, when artificial intelligence systems can generate poetry indistinguishable from works by celebrated human authors. Recent studies demonstrate that readers cannot reliably differentiate between AI-generated poems and those written by canonical poets like Emily Dickinson, Walt Whitman, and Sylvia Plath (Porter and Machery 2024). This technological capability forces a fundamental reconsideration of concepts that have anchored literary study for centuries: authorship, originality, and creative authenticity. The emergence of AI-generated literature represents more than a technological novelty; it constitutes what David Gunkel describes as a crisis of meaning itself, challenging our assumptions about who or what can be considered an author and whether the concept of individual creative authorship remains viable in an age of algorithmic text generation.

The stakes of this inquiry extend far beyond academic literary theory. As AI writing tools become increasingly sophisticated and accessible, they raise urgent questions about intellectual property, academic integrity, and the value of human creative expression. Educational institutions grapple with policies regarding AI

assistance in writing, while publishers and literary organizations debate whether AI-generated works merit publication and recognition. These practical concerns reflect deeper philosophical questions about the nature of creativity, consciousness, and meaning-making that have implications for how we understand literature's role in human culture.

This paper examines whether AI-generated poetry and fiction can be considered original creative work and analyzes what this technological capability means for traditional concepts of authorship. By engaging with empirical research on reader responses to AI literature, legal frameworks governing authorship, and theoretical perspectives from Roland Barthes to contemporary digital humanities scholars, this analysis demonstrates that AI-generated literature occupies a complex position that neither simply replaces human authorship nor leaves it unchanged. Instead, AI literature necessitates new theoretical frameworks that can account for forms of creativity that emerge from human-machine collaboration while preserving meaningful distinctions between different modes of literary production.

Theoretical Framework: Reconsidering Authorship in the Digital Age

The relationship between authorship and meaning has been contested territory in literary theory since Roland Barthes declared the "death of the author" in his influential 1967 essay. Barthes argued that "writing is that neutral, composite, oblique space... the negative where all identity is lost, starting with the very identity of the body writing" (Barthes qtd. in Gunkel 2025). For Barthes, meaning emerges not from authorial intention but from the reader's encounter with text, making the author's identity and purpose irrelevant to interpretation. This perspective, which emphasized "the primacy of each individual reader's interpretation of the work over any 'definitive' meaning intended by the author," anticipated many of the questions now raised by AI-generated literature (Wikipedia).

However, the emergence of large language models has pushed Barthes's insights to their logical extreme in ways that even he could not have anticipated. As Gunkel observes, AI systems produce "texts that do not have, nor are beholden to, the authoritative voice of an author, and statements whose truth cannot be anchored in and assured by a prior intention to say something." This represents what he calls a more conclusive form of authorial death than Barthes envisioned not merely the separation of text from authorial intention, but the complete absence of a human consciousness behind the writing.

Yet Michel Foucault's response to Barthes suggests important limitations to this framework. Foucault argued that "while the author is not a fixed, consolidated subject that straightforwardly determines meaning, our knowledge of the author still plays an essential role in producing and regulating how texts are used and interpreted" (qtd. in New Statesman 2023). This perspective proves particularly relevant to AI literature, where the absence of human authorship creates interpretive challenges that go beyond simple questions of meaning to encompass issues of value, authenticity, and literary merit.

The legal framework surrounding authorship provides another lens through which to examine these questions. Recent court decisions have consistently held that AI systems cannot be considered authors for copyright purposes, with the U.S. Court of Appeals upholding decisions "denying authorship to AI" (Gunkel 2025). This legal stance reflects broader cultural assumptions about authorship as necessarily involving human agency and intentionality. However, as Mazzi notes in analyzing intellectual property law, the question becomes more complex when considering "the originality of text prompts and AI-generated outputs," suggesting that human involvement in prompting and selecting AI-generated content may constitute a form of creative contribution worthy of legal recognition.

The digital humanities have begun developing new theoretical approaches to these questions that move beyond simple binary distinctions between human and machine authorship. Scholars in this field recognize that AI-generated literature emerges from what can be understood as a form of collaboration between human prompters, algorithmic systems, and the vast corpus of human-written texts on which AI models are trained. This collaborative model suggests that rather than representing the complete death of the author, AI literature might herald the emergence of new forms of distributed or collective authorship that require fresh theoretical frameworks.

Analysis: The Question of Originality in AI-Generated Literature

The question of whether AI-generated literature can be considered original hinges on how we define originality itself. Traditional conceptions of literary originality emphasize novelty, uniqueness, and the expression of individual creative vision. However, empirical research reveals a complex picture that challenges simple assessments of AI creativity. A recent comparative study of AI-generated and human-authored drama found that while AI systems can "merge several literary genres in a comprehensive whole and innovative collage," they

ultimately produce work that "lacks psychological and emotional depth" and appears "devoid of the complexity of human psychology." (Elias et al. 2025).

This limitation points to a fundamental tension in AI creativity. Large language models excel at pattern recognition and recombination, allowing them to generate text that follows established literary conventions with remarkable fidelity. Studies have shown that AI can produce poetry that not only mimics the style of canonical poets but actually receives higher ratings from non-expert readers than authentic human-written poems when evaluated on qualities like "rhythm and beauty" (Porter and Machery 2024). However, this apparent success may reflect the limitations of the evaluation rather than genuine creative achievement.

Critics argue that what appears as AI creativity is fundamentally different from human originality because it lacks intentionality and consciousness. As one analysis notes, AI-generated poems "cannot match the complexity of human-authored verse" and tend toward "formulaic work" that is "more straightforward and generally more accessible" than human poetry (Kuta qtd. in Intellectual Takeout). This accessibility may explain why non-expert readers prefer AI poetry not because it is superior, but because it avoids the "tension, complexity, and mystery of the greatest poetry, which can only come from the equally mysterious human consciousness and its encounter with a meaningful reality."

The question of originality becomes even more complex when considering the training data on which AI systems rely. Large language models are trained on vast corpora of human-written text, leading to concerns about what amounts to sophisticated plagiarism. As one researcher notes, AI systems may "unintentionally generate text that mirrors the style or content of its training data, unbeknownst to users," creating "critical questions about how society perceives and defines plagiarism in the context of AI-generated text" (Zulhusni qtd. in Turnitin). This raises fundamental questions about whether originality can exist when the creative process involves the systematic analysis and recombination of existing human works.

However, some scholars argue for a more nuanced understanding of AI creativity that recognizes different types of originality. The concept of combinatorial creativity the novel recombination of existing elements has long been recognized as a legitimate form of human creativity, and AI systems excel at identifying unexpected connections between disparate concepts. Moreover, the collaborative nature of AI writing, in which human prompters guide and select from AI outputs, suggests a form of co-creativity that may constitute its own type of originality.

The legal perspective on originality provides additional complexity. European copyright law defines originality as requiring that a work be "its author's own intellectual creation," involving "intellectual effort, which is free and creative" (European Court of Justice qtd. in Oxford Academic). Under this standard, AI-generated works cannot be considered original because they lack human intellectual effort. However, works that involve significant human editing or curation of AI output may qualify for protection, suggesting a spectrum of originality that depends on the degree of human involvement.

Critical Evaluation: Human-Machine Collaboration and Literary Value

The emergence of AI-generated literature challenges traditional binary distinctions between human and machine creativity, suggesting instead a spectrum of collaborative relationships that resist easy categorization. Contemporary practice reveals multiple models of human-AI literary collaboration, from simple prompt-based generation to sophisticated iterative processes involving human editing, curation, and refinement. Each model raises different questions about authorship, creativity, and literary value that require careful analysis.

At one end of the spectrum lies what researchers term "human-out-of-the-loop" AI generation, where systems produce text with minimal human intervention beyond initial prompting. Studies of this approach reveal significant limitations in the resulting literature. AI systems operating independently tend to produce work that, while technically competent, lacks the interpretive complexity that characterizes valued human literature. As one analysis notes, such systems cannot "provide accurate quotes and sources" and demonstrate a "tendency to introduce factual error," highlighting their disconnection from the epistemic frameworks that ground human literary creation (Amirjalili et al. 2024).

More promising is the "human-in-the-loop" model, where human agents actively participate in selecting, editing, and refining AI-generated content. Research indicates that this collaborative approach can produce work that is "indistinguishable from human-written poetry" when evaluated by non-expert readers (Köbis and Mossink 2021). However, this apparent success masks important questions about the nature of the human contribution and whether such collaboration constitutes genuine co-creation or sophisticated tool use.

The question of literary value proves particularly complex in evaluating AI-generated literature. Empirical studies reveal a paradox: while readers often rate AI-generated poems higher than human-written works on metrics like accessibility and formal structure, critics argue that this preference reflects the limitations of AI

literature rather than its strengths. AI poems receive higher ratings because they are "more straightforward and generally more accessible," avoiding the "complexity and mystery" that characterizes the most valued human poetry (Porter and Machery 2024). This suggests that reader preference may be a poor metric for evaluating literary achievement, particularly when readers lack expertise in poetry analysis.

The institutional response to AI literature provides another lens for understanding its literary value. Major literary publications and academic journals have generally prohibited AI-generated content, with organizations like the Science family of journals explicitly stating that "text generated from AI, machine learning, or similar algorithmic tools cannot be used in papers published in Science journals" (Environmental Factor). Similarly, the Committee on Publication Ethics maintains that "AI tools cannot be listed as an author of a paper" because they "cannot take responsibility for the submitted work" (COPE). These policies reflect institutional commitments to human agency and accountability that extend beyond simple questions of textual quality.

However, some scholars argue for a more nuanced approach to AI literature that recognizes its potential contributions while maintaining meaningful distinctions from human creativity. The concept of "symbiotic cooperation of human and machine" suggests possibilities for genuine collaboration that preserves human agency while leveraging AI capabilities (Rosa et al. qtd. in Elias et al. 2025). Under this model, AI systems serve as sophisticated creative partners that can "assist writers by providing suggestions for character development" and enable "exploration of unconventional narratives and diverse voices," while humans retain ultimate responsibility for creative decisions and meaning-making.

The educational implications of AI-generated literature reveal additional complexities in evaluating its value and appropriate use. While AI writing tools offer potential benefits for learning and creativity, concerns about academic integrity and the development of human creative capacities remain paramount. As researchers note, "over-reliance on AI tools, without critical human oversight, can lead to issues such as unintentional plagiarism, loss of individual authorship, and a decline in the quality of written work" (Turnitin). This suggests that the value of AI literature may depend significantly on how it is integrated into broader practices of human learning and creativity.

Implications: Toward New Frameworks for Digital Literature

The emergence of AI-generated literature necessitates new theoretical and practical frameworks that can accommodate forms of creativity that transcend traditional boundaries between human and machine agency. Rather than simply extending existing literary theory, these developments require fundamental reconceptualization of concepts like authorship, originality, and literary value that have structured literary study for centuries.

One promising direction lies in developing what might be termed "collaborative authorship" models that recognize the distributed nature of AI literary creation. Such frameworks would acknowledge that AI-generated literature emerges from complex interactions among human prompters, algorithmic systems, training datasets composed of human writing, and the broader cultural contexts that shape both human and machine "reading" practices. This perspective suggests moving beyond individual authorship toward understanding literature as an inherently social and technological practice that has always involved multiple agents and influences.

The legal implications of this shift require careful consideration of how intellectual property frameworks might evolve to accommodate new forms of creative collaboration. Current copyright law's insistence on human authorship reflects important values about agency and responsibility, but it may prove inadequate for addressing the realities of human-AI collaboration. Scholars suggest developing "flexible licensing models" and new forms of "digital rights management" that can protect human creative contributions while enabling legitimate uses of AI assistance (Mazzi 2024). Such approaches might distinguish between works that involve significant human creative input and those that represent primarily algorithmic generation.

Educational institutions face particular challenges in developing appropriate policies for AI-generated literature. Rather than blanket prohibitions, educators might focus on promoting "transparency" and "full disclosure of use" while emphasizing the development of critical evaluation skills (Environmental Factor). This approach would treat AI as a powerful tool that requires careful integration into educational practice rather than something to be avoided entirely. Students would learn not only how to use AI writing tools effectively but also how to maintain their own creative agency and critical judgment in the process.

The implications for literary criticism and scholarship are equally significant. Critics and scholars must develop new methodologies for analyzing texts whose origins may be uncertain or collaborative. This might involve greater attention to textual analysis that focuses on formal and aesthetic qualities rather than authorial intention, ironically returning to some of the insights of mid-twentieth-century formalist criticism while accommodating new technological realities. Additionally, scholars might develop expertise in what could be

called "digital philology" the analysis of how AI systems process and recombine textual materials in ways that reveal new patterns and possibilities in literary tradition.

The broader cultural implications of AI-generated literature extend to questions about the value and purpose of human creative expression. Rather than viewing AI as a threat to human creativity, some scholars argue for understanding it as a tool that can "enhance human imagination" and enable new forms of creative collaboration (Analytics Vidhya). This perspective suggests that the value of human creativity lies not in its exclusivity but in its capacity for meaning-making, emotional depth, and cultural connection qualities that may become more rather than less important as AI capabilities expand.

However, these developments also raise important questions about the future of literary culture and the role of human creativity within it. If AI systems can produce literature that satisfies many readers' preferences for accessibility and formal competence, what happens to the market for human literary creation? How do we maintain space for the kinds of complex, challenging literature that may not appeal to immediate algorithmic optimization but that serves important cultural functions? These questions suggest the need for deliberate cultural policies that preserve space for human creative expression while embracing the possibilities of technological augmentation.

Conclusion

The emergence of sophisticated AI-generated literature represents a watershed moment in literary history that demands fundamental reconsideration of concepts that have long structured our understanding of creative expression. While empirical research demonstrates that AI systems can produce text that is often indistinguishable from human writing to non-expert readers, the question of whether such work can be considered truly original or authentic literature remains contentious and complex.

This analysis reveals that AI-generated literature exists in a liminal space between human creativity and algorithmic reproduction that resists simple categorization. While AI systems lack the consciousness, intentionality, and lived experience that have traditionally been associated with literary authorship, they demonstrate capabilities for pattern recognition, linguistic manipulation, and creative recombination that produce works of genuine interest and, in some cases, aesthetic merit. The most promising developments emerge from collaborative models that preserve human agency while leveraging AI capabilities, suggesting possibilities for new forms of creative partnership rather than simple replacement of human authors.

The theoretical implications of these developments extend Roland Barthes's insights about the death of the author in directions that even he could not have anticipated. AI literature does indeed represent a form of writing without a speaking subject, but rather than simply liberating readers to create meaning, it raises new questions about the sources and nature of meaning itself. The absence of human consciousness behind AI-generated text creates interpretive challenges that require new critical methodologies and theoretical frameworks.

The practical implications are equally significant. Educational institutions, publishers, and cultural organizations must develop policies that can distinguish between appropriate uses of AI assistance and forms of technological dependence that undermine human creative development. Legal frameworks must evolve to accommodate new forms of collaborative creativity while preserving important principles of human agency and responsibility. Critics and scholars must develop new methodologies for analyzing texts whose origins may be uncertain or distributed across human and machine agents.

Perhaps most importantly, the emergence of AI-generated literature clarifies rather than diminishes the distinctive value of human creative expression. The qualities that make human literature most valuable its capacity for emotional depth, cultural insight, moral complexity, and meaning-making become more rather than less important in an age of algorithmic text generation. AI systems excel at pattern recognition and recombination, but they cannot replicate the lived experience, moral imagination, and cultural understanding that enable literature to serve its most important functions in human life.

The path forward requires neither wholesale embrace nor blanket rejection of AI-generated literature, but rather the development of nuanced frameworks that can preserve what is most valuable about human creativity while remaining open to the possibilities of technological augmentation. This will require ongoing dialogue among literary scholars, technologists, educators, and policymakers to ensure that the development of AI literature serves human flourishing rather than diminishing it. As we navigate this transition, the fundamental questions raised by AI-generated literature about the nature of creativity, the sources of meaning, and the value of human expression remind us that literature's ultimate importance lies not in its production methods but in its capacity to illuminate the human condition and connect us more deeply to one another and to the world we share.

The ethics of AI-generated literature ultimately depend not on resolving these questions definitively but on maintaining them as living concerns that guide our choices about how to integrate these powerful new tools

into our literary culture. By doing so, we can ensure that technological innovation serves to expand rather than constrain the possibilities for human creative expression and cultural understanding.

Reference

Elias, S., B. S. Alshammari, K. N. Alfraidi, and others. 2025. "Rethinking Literary Creativity in the Digital Age: A Comparative Study of Human versus AI Playwriting." *Humanities and Social Sciences Communications* 12: 689. https://doi.org/10.1057/s41599-025-04999-2.

Amirjalili, Fatemeh, Mahdieh Neysani, and Atefeh Nikbakht. 2024. "Exploring the Boundaries of Authorship: A Comparative Analysis of AI-Generated Text and Human Academic Writing in English Literature." *Frontiers in Education* 9. https://doi.org/10.3389/feduc.2024.1347421.

Committee on Publication Ethics. 2025. "Authorship and AI Tools." *COPE Position Statement*, February 24, 2025. https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools.

Environmental Factor. 2023. "Science Journals Set New Authorship Guidelines for AI-Generated Text." *Environmental Factor*, March 2023. https://factor.niehs.nih.gov/2023/3/feature/2-artificial-intelligence-ethics.

Gunkel, David J. 2025. "AI Signals the Death Of The Author." *NOEMA*, June 4, 2025. https://www.noemamag.com/ai-signals-the-death-of-the-author/.

Holyoak, Keith. 2023. "Can AI Write Authentic Poetry?" *The MIT Press Reader*, March 2, 2023. https://thereader.mitpress.mit.edu/can-ai-write-authentic-poetry/.

Köbis, Nils C., and Luca D. Mossink. 2021. "Artificial Intelligence versus Maya Angelou: Experimental Evidence That People Cannot Differentiate AI-Generated from Human-Written Poetry." *Computers in Human Behavior* 114: 106553. https://doi.org/10.1016/j.chb.2020.106553.

Mazzi, Margherita. 2024. "Authorship in Artificial Intelligence-Generated Works: Exploring Originality in Text Prompts and Artificial Intelligence Outputs Through Philosophical Foundations of Copyright and Collage Protection." *The Journal of World Intellectual Property* 27 (3). https://doi.org/10.1111/jwip.12310.

Wikipedia. 2025. "The Death of the Author." *Wikipedia*, May 25, 2025. https://en.wikipedia.org/wiki/The_Death_of_the_Author.

New Statesman. 2023. "ChatGPT and the Death of the Author." *New Statesman*, April 17, 2023. https://www.newstatesman.com/the-weekend-essay/2023/02/chatgpt-death-author-big-tech-artificial-intelligence.

Porter, Brian, and Edouard Machery. 2024. "AI-Generated Poetry Is Indistinguishable from Human-Written Poetry and Is Rated More Favorably." *Scientific Reports* 14. https://doi.org/10.1038/s41598-024-76900-1.

Turnitin. n.d. "AI Writing in Journals: Preserving Research Integrity." *Turnitin Blog*.

Fritz, Johannes. 2025. "Understanding Authorship in Artificial Intelligence-Assisted Works." *Journal of Intellectual Property Law & Practice* 20 (5): 354–364. https://doi.org/10.1093/jiplp/jpae119