Journal of Literary and Artistic Expressions
ISSN *(Online)*: 3105-2002
Volume 2, Issue 1, Jan-Feb, 2025, Page 5-8

Review Article
Received: 07-01-2026| Accepted: 22-02-2026| Published: 28-02-2026

The Algorithmic Muse: Authorship, Creativity, and Aesthetic Agency in the Age of Generative AI

***Tarak Mehta* ***

Abstract

The emergence of large language models capable of producing fluent poetry, memoir, and fiction has renewed long-standing critical debates about authorship, creativity, and aesthetic agency. This paper examines recent literary experiments in human-AI co-creation to argue that generative AI does not simply revive or resolve the poststructuralist critique of the author associated with Barthes and Foucault; it stages that critique in a new, materially specific form.

Drawing on Colella's analysis of the enlarged paratexts surrounding AI-generated literature and Bajohr's model of graduated authorship, from primary human composition to quaternary prompt-based generation, the paper argues that the human figures who curate, prompt, and edit AI output have not disappeared from these works but have instead relocated their authorial presence into surrounding textual apparatus: prefaces, afterwords, footnotes, and public commentary that explain, validate, and claim responsibility for the machine's language. Benjamin's earlier account of aura and mechanical reproduction offers a useful historical parallel, suggesting that each new reproductive technology prompts a similar renegotiation of where authenticity and authorial presence are felt to reside.

The paper concludes that AI-generated literature, far from eliminating the author, has intensified the visibility of authorial labor by displacing it from the sentence to the surrounding frame, making paratext, rather than prose style alone, the primary site where contemporary readers are asked to locate a human creative presence.

Keywords

Generative AI, Authorship, Paratext, Creativity, Digital Literature

**Independent Scholar*

INTRODUCTION

Since the wide public release of large language models capable of generating fluent prose and verse, a growing body of literary experiments has combined human prompting, editing, and curation with machine-generated text, producing memoirs, novels, and poetry collections credited, in varying degrees, to artificial intelligence. These experiments revive, in a strikingly literal form, the theoretical questions poststructuralist criticism raised decades earlier about the status of the author. Where Barthes and Foucault argued, on largely conceptual grounds, that the author was a constraining fiction imposed on texts whose meaning properly belonged to language and to readers, generative AI produces texts whose composition genuinely involves no single human consciousness choosing each word. This paper asks what happens to the concept of authorship when a text's origin is, quite literally, distributed between a statistical model and a human collaborator, and argues that the human figure has not vanished from these works but has migrated to a new location: the paratext.

GRADUATED AUTHORSHIP

Bajohr's model of graduated authorship offers a useful framework for tracking the human contribution across different modes of AI-assisted composition. In this model, primary authorship describes conventional human composition; secondary authorship describes rule-based or combinatorial writing, where the human designs a procedure that generates text; tertiary authorship describes machine-learning-based generation, where the human selects training data and parameters but does not control the model's internal process; and quaternary authorship describes composition through prompting alone, where the human's contribution consists chiefly of the instructions given to a pretrained, proprietary model. As the distance between human intention and generated sentence increases across these categories, the human contribution does not disappear; it becomes concentrated in different

***Corresponding Author:** Tarak Mehta

tasks, selecting, prompting, curating, editing, and ultimately explaining the process to readers.

This graduated model clarifies why recent works of AI-assisted literature are so consistently accompanied by extensive paratextual explanation. As Colella's analysis of several such works demonstrates, prefaces, afterwords, and editorial notes do more than contextualize these texts; they actively perform the labor of authorship that the sentences themselves no longer visibly carry, establishing who is responsible for the work's meaning and value even when that person did not compose its individual sentences.

Aura Relocated

Benjamin's account of aura, the unique presence and authenticity a work of art possesses by virtue of its singular, unrepeatable existence, and his argument that mechanical reproduction diminishes this aura by making the work endlessly copyable, provides a suggestive historical parallel to the present situation, even though the technologies differ considerably. Benjamin observed that new reproductive technologies do not simply destroy aura; they relocate the conditions under which authenticity is perceived, shifting value from the object's uniqueness toward other markers, such as the circumstances of its exhibition or the story of its making.

Generative AI appears to be prompting a comparable relocation. If a poem's individual sentences can no longer reliably signal a unique human consciousness behind them, since a model trained on vast corpora can produce comparably fluent language, then readers and publishers redirect their search for authenticity toward the surrounding apparatus: the documented prompting process, the editor's account of what was kept and discarded, the public explanation of how much of the text is unaltered machine output. Aura, in this account, has not disappeared from AI-assisted literature; it has moved from the sentence to the frame.

The Persistence of the Author Function

This relocation of authorial presence complicates any straightforward claim that generative AI finally realizes the poststructuralist prediction of the author's disappearance. Foucault's account of the author function describes not a person but a set of classificatory operations that a culture performs on certain texts, assigning responsibility, coherence, and ownership. The paratextual apparatus surrounding AI-generated literature performs precisely this function, even in the apparent absence of a conventional author. A preface that explains the prompting method, a hyperlinked transcript of the human-AI exchange, or a public interview describing editorial choices all serve to reassign the author function to the human collaborator, regardless of how few individual sentences that person composed.

This suggests that the author function, as Foucault described it, is remarkably adaptable. It does not require that a human compose every sentence of a text in order to attach itself to that human; it requires only that some human figure be available to occupy the position of explainer, curator, and claimant of responsibility. Generative AI has not eliminated this position. It has made its occupant's labor more visible than ever, since that labor must now be actively performed and displayed, rather than assumed as the invisible background condition of conventional authorship.

HISTORICAL PARALLELS: FROM PHOTOGRAPHY TO NEURAL NETWORKS

The present anxiety about generative AI's implications for authorship and creativity is not without precedent, and situating it within the longer history of reproductive technologies helps distinguish what is genuinely new about the current moment from what repeats an established pattern of technological anxiety. Benjamin's analysis of photography and film focused on a related but distinct concern: not whether a machine could originate content, but whether mechanical reproduction of an already-existing artwork diminished that work's aura, its unique presence in time and space. Painters in the nineteenth century worried that photography would render painting obsolete by automating the previously skilled labor of accurate visual representation; instead, photography's rise

pushed painting toward abstraction and away from representational accuracy, redefining rather than eliminating the painter's distinct contribution.

Generative AI's relationship to writing follows a structurally similar but not identical pattern. Where photography automated the accurate reproduction of an already-existing visual scene, large language models automate the production of fluent, grammatically coherent prose and verse, a capability previously understood as requiring sustained human practice and talent. The anxiety this automation produces echoes the painterly anxiety Benjamin described, but the redefinition it appears to be producing runs not toward abstraction, as painting's response to photography did, but toward the paratextual intensification this paper has already described: a shift in where readers are asked to locate the markers of authentic, valued creative labor, from the sentence itself toward the visible human process of prompting, selecting, and explaining that surrounds it.

Reader Response and the Demand for Authenticity

Empirical work on how readers actually respond to AI-generated creative writing offers a useful check on the purely theoretical claims advanced so far. Porter and Machery's study testing readers' ability to distinguish AI-generated poems from human-authored ones found that participants were, if anything, more likely to misidentify AI-generated poems as human-authored than the reverse, yet the same participants rated poems significantly less favorably once informed that a machine had produced them. This finding is theoretically significant because it demonstrates that the demand for a human origin behind a literary text is not primarily a demand about the text's surface qualities, since blind readers frequently cannot detect the difference through surface features alone, but a demand about the conditions of the text's production that becomes active only once those conditions are disclosed.

This helps explain why the paratextual strategies examined throughout this paper are not merely marketing choices but responses to a real and measurable feature of reader psychology. If readers' evaluative response to a poem shifts substantially based on attributed authorship even when their perceptual ability to detect authorship does not, then the paratext's work of establishing, emphasizing, or complicating human involvement is doing real interpretive labor, not simply framing an already-settled aesthetic judgment. The persistent finding that human origin functions as a value marker independent of textual quality suggests that the relocation of aura this paper has traced, from sentence to surrounding frame, responds to a durable feature of how readers assign value to creative work, one that generative AI has made newly visible rather than newly created.

CONCLUSION

Generative AI does not straightforwardly confirm or refute the poststructuralist critique of authorship; it restages that critique under new material conditions, in which the disappearance of the author from the level of the sentence is compensated by an intensification of authorial presence at the level of the paratext. Readers of AI-assisted literature are asked to locate creative agency not in the fluency of individual sentences, which a model can now supply on demand, but in the surrounding frame of prompts, explanations, and editorial choices that a human collaborator continues to supply. The algorithmic muse, in this sense, has not replaced the author. It has relocated her, in plain sight, to the threshold of the text.

This relocation carries implications for how literary criticism approaches AI-assisted texts going forward. Reading practices built around close attention to sentence-level style, diction, rhythm, syntactic choice, are not well suited to a body of work whose most consequential authorial decisions increasingly occur before and after the sentence is generated, in the framing of the prompt and the curation of the output. A critical vocabulary adequate to this new body of work will need to treat the paratext not as auxiliary material surrounding the real literary object but as itself a primary site of literary invention, deserving the same close attention long reserved for style and voice.

REFERENCES

1. Bajohr, H. (2024). Writing at a distance: Notes on authorship and artificial intelligence. German Studies Review, 47, 315-337. https://doi.org/10.1353/gsr.2024.a927862
2. Barthes, R. (1977). The death of the author. In Image, music, text (S. Heath, Trans., pp. 142-148). Fontana. (Original work published 1967)
3. Benjamin, W. (1969). The work of art in the age of mechanical reproduction. In H. Arendt (Ed.), Illuminations (H. Zohn, Trans., pp. 217-251). Schocken Books. (Original work published 1935)
4. Colella, S. (2025). "The language of the digital air": AI-generated literature and the performance of authorship. Humanities, 14(8), 164. https://doi.org/10.3390/h14080164
5. Foucault, M. (1998). What is an author? In J. Faubion (Ed.), Aesthetics, method, and epistemology (Vol. 2). The New Press.
6. Genette, G. (1997). Paratexts: Thresholds of interpretation. Cambridge University Press.
7. Porter, B., & Machery, E. (2024). AI-generated poetry is indistinguishable from human-written poetry and is rated more favorably. Scientific Reports, 14, 26133. https://doi.org/10.1038/s41598-024-76900-

Conflict of Interest: No Conflict of Interest
Source of Funding: Author(s) Funded the Research
How to Cite: Mehta, T. (2025). The Algorithmic Muse: Authorship, Creativity, and Aesthetic Agency in the Age of Generative AI. *Journal of Literary and Artistic Expressions, 2*(2), 5-8.