IJSSRA | ISSN (online): 3107-6238
January-March 2026 | Vol. 1 | Issue 4

The Posthuman Turn: Artificial Intelligence, Algorithmic Authorship, and the Evolution of the Literary Narrative

Dr. Saraswati Singram*

***Assistant Professor, Government Chandravijay PG College, Dindori, Madhya Pradesh
E-Mail: saraswati.singram@gmail.com**
DOI: https://doi.org/10.5281/zenodo.20834083

Abstract

This paper explores the posthuman transformation of literary authorship in the age of artificial intelligence, examining how algorithmic systems challenge traditional notions of creativity, authorship, and narrative agency. Drawing on the posthumanist theories of Rosi Braidotti and N. Katherine Hayles, as well as the literary frameworks of Roland Barthes and Michel Foucault, the study investigates the emergence of human–machine co-creation in contemporary literature. Through qualitative analysis of AI-generated and AI-assisted texts, including 1 the Road, Sunspring, and recent publishing controversies, the paper demonstrates that authorship is increasingly distributed across human creators, algorithms, training data, and readers. The findings reveal that AI does not eliminate authorial intent but redistributes it within a collaborative network of agents. Ultimately, the study argues that the posthuman turn expands literary possibilities and requires new critical approaches to understanding narrative production, interpretation, and cultural value.

Keywords: *Posthumanism, Artificial Intelligence, Algorithmic Authorship, Narrative Agency, Human–Machine Co-Creation.*

Introduction

In April, the North American publishing industry was convulsed by a scandal that exposed the fragility of contemporary authorship: horror writer Mia Ballard's novel *Shy Girl*, originally self-published and subsequently acquired by Hachette, was cancelled by the publisher after a *New York Times* investigation presented compelling evidence that the work had been generated, or heavily mediated, by generative artificial intelligence. Ballard's subsequent denial—and her claim that any AI involvement might have stemmed from an editor rather than herself—only intensified the crisis, revealing that the boundary between human and machine composition has become not merely blurred but legally and ontologically untenable. This incident is not isolated. It arrives in the wake of years of escalating anxiety about AI-generated texts flooding digital marketplaces: in 2023, Amazon began requiring authors to disclose AI-generated content in its Kindle Direct Publishing program, responding to pressure from organizations such as the Authors Guild, which documented a surge of

"sham" books—unauthorized biographies, summary guides, and copycat titles—produced algorithmically to exploit the names and labor of human writers (Price, 2023). Meanwhile, the very title of Stephen Marche's *Death of an Author: A Novella* (2023), a work experimentally co-created with large language models, seems to literalize Roland Barthes's poststructuralist prophecy, suggesting that the "author" as a sovereign, intentional figure may have finally been supplanted by the scriptor as statistical engine (Matějková & Ircing, 2024).

These controversies signal something more profound than a technological disruption of publishing markets. They mark a fundamental epistemological crisis for the humanist paradigm of authorship that has dominated Western literary culture since the Enlightenment—a paradigm that privileges the single, conscious, intentional subject as the origin and guarantor of textual meaning. As Barthes argued in his seminal 1967 essay, "The Death of the Author," to anchor interpretation in the author's intention is to "impose a limit on that text, to furnish it with a final signified, to close the writing"(Rodeux, 2025). Yet Barthes's scriptor—an impersonal mediator through which language itself acts—remained a human figure, however decentered (Matějková & Ircing, 2024). What generative AI introduces is not merely the absence of the author but the presence of a non-human, distributed agency: an algorithmic system that produces narrative without consciousness, biography, or expressive desire, yet whose outputs materially participate in the construction of literary meaning. The traditional model of the author as a unified, originary consciousness finds itself in tension with algorithmic systems that generate text through probabilistic recombination of vast cultural corpora, effectively dispersing agency across training data, model architecture, and human prompt (Rodeux, 2025).

This paper argues that the emergence of algorithmic authorship necessitates a posthumanist reconceptualization of narrative agency, one that moves beyond anthropocentric models of creativity to recognize distributed, human-machine co-creation as a constitutive force in the evolution of literary narrative. Drawing on contemporary posthumanist theory—particularly Rosi Braidotti's conception of the "expanded self" whose relational field includes non-anthropomorphic elements and N. Katherine Hayles's framework of distributed cognition across human-machine systems—this study examines how large language models function not merely as tools but as active narrative agents that reconfigure the boundaries of authorship, identity, and textual meaning (Bandhopadhyay, 2025). Through an analysis of recent controversies in publishing, algorithmic co-creation in experimental fiction, and the theoretical intersections of Barthes's "scriptor" with contemporary AI ethics, this paper demonstrates that the "posthuman turn" in literary culture is not a speculative future but an emergent reality. In this reality, as recent scholarship on posthuman narratives suggests, artistic agency is no longer the exclusive property of human consciousness but is distributed across a hybrid

creative ecosystem in which human and machine intelligences co-produce the stories that shape our cultural imagination (Dao, 2023).

Literature Review

Posthumanism — Foundational Theory

Posthumanism emerges as a critical intervention against the Enlightenment humanist tradition that installs a specific, universalized subject—what Rosi Braidotti identifies as "male, white, heterosexual, able-bodied, speaking a standard language"—as the sovereign measure of all existence (Apata, 2021). Braidotti argues that this anthropocentric paradigm, rooted in Protagoras's dictum that "Man is the measure of all things," not only marginalizes racialized, gendered, and sexualized others but also perpetuates a logic of human exceptionalism that relegates non-human animals, ecosystems, and technological actors to the status of mere instrumental objects (Apata, 2021). In her formulation, posthumanism does not herald the extinction of humanity but rather demands a radical re-evaluation of the human within a broader system of relational entanglements. Subjectivity is no longer conceived as the property of an autonomous, self-sufficient individual but as a technologically mediated, embodied, and distributed phenomenon that is co-constituted by its interactions with non-human others. N. Katherine Hayles's foundational study How We Became Posthuman (1999) extends this critique by dismantling the humanist fantasy of disembodied consciousness. Hayles demonstrates that the posthuman subject is forged through cybernetic feedback loops and informational patterns that traverse biological and technological substrates, arguing that we have always been posthuman: our cognition and creativity are not sealed within the cranium but are extended, relational, and emergent from the complex entanglements of flesh and code. Together, Braidotti and Hayles establish the theoretical premise that undergirds this study: the decentering of human exceptionalism is not a speculative future but an ontological condition of the present, one that necessitates new frameworks for understanding narrative agency, creativity, and authorship in the age of artificial intelligence.

Death of the Author Revisited

Roland Barthes's seminal 1967 essay "The Death of the Author" provides the literary-critical antecedent for dismantling the sovereignty of the individual writer, arguing that the author is not the origin of textual meaning but merely a scriptor through whom language itself circulates. To anchor interpretation in the author's intention, Barthes contends, is to "impose a limit on that text, to furnish it with a final signified, to close the writing"—a closure that the emergence of algorithmic text generation renders technologically obsolete (Rodeux, 2025). In the age of large language models, Barthes's scriptor acquires a new, uncanny materiality: the algorithmic system generates narrative without biography, desire, or interiority, yet its outputs nevertheless enter the circuits of literary

meaning and readerly interpretation. Where Barthes's scriptor remained a human mediator, however impersonal, AI authorship literalizes the evacuation of the sovereign subject, replacing it with a probabilistic engine that recombines vast cultural corpora into seemingly coherent discourse (Matějková & Ircing, 2024).

Michel Foucault's contemporaneous lecture "What is an Author?" (1969) complicates this theoretical landscape by demonstrating that the "author" is never merely a flesh-and-blood individual but rather an "author-function"—a discursive and institutional construct that operates as a principle of classification, a legal category of ownership, and a mechanism for dispersing multiple subject-positions across a body of texts. Foucault shows that the author-function is linked to the juridical and institutional systems that circumscribe discourse and that it does not refer purely to a real individual, since it can give rise simultaneously to several selves and subjective positions (D. FAUBION, 1998). In the context of algorithmic authorship, Foucault's analysis becomes particularly salient: if the author has always been a function rather than a person, then AI-generated texts do not so much abolish the author as displace the author-function onto a distributed assemblage of training data, model architecture, human prompt engineering, and corporate infrastructure. The AI text thus does not lack an author; rather, it demands that we reconceptualize authorship as a plural, distributed, and materially embedded function.

3.3 Algorithmic Authorship — Existing Scholarship

The intersection of posthumanist theory and literary practice has been most rigorously explored in the domains of electronic literature and cybernetic textual studies. N. Katherine Hayles's Writing Machines (2002) extends her posthumanist framework into the materiality of literary production, arguing that texts are never immaterial linguistic constructs but are always embodied in specific media technologies—what she terms "technotexts." For Hayles, the writing machine is not a passive instrument but an active participant in the construction of meaning, a perspective that anticipates the current era of generative AI by demonstrating that narrative agency has always been distributed across human and non-human actors (Hayles & Burdick, 2002). This framework of distributed cognition finds further support in Annie Murphy Paul's The Extended Mind (2021), which argues that human cognition is fundamentally scaffolded by external tools, physical environments, and social relationships. Paul's work suggests that large language models function not as alien intelligences but as radical extensions of the cognitive ecologies within which human creativity has always operated (Paul, 2024). Within the specific tradition of electronic literature, Marjorie Luesebrink (writing under the pseudonym M.D. Coverley) pioneered hypertext and born-digital narrative forms that positioned the computer as a co-creative agent rather than a mere delivery mechanism, establishing an aesthetic and theoretical precedent for contemporary algorithmic authorship.

These theoretical orientations are illuminated by concrete case studies in AI-generated literature. Ross Goodwin's 1 the Road (2018), widely recognized as the first novel generated by an artificial intelligence, offers a paradigmatic example of distributed narrative agency. Goodwin outfitted a Cadillac with a surveillance camera, GPS unit, microphone, and internal clock, all feeding real-time data into a long short-term memory recurrent neural network that generated text during a journey from New York to New Orleans. Published unedited by Jean Boîte Éditions, the work begins with the haunting line, "It was nine seventeen in the morning, and the house was heavy," and proceeds through a discontinuous, sensor-driven prose that literalizes the posthumanist dissolution of the bounded author (1 the Road, 2020). Alongside such experimental projects, the proliferation of GPT-3 and subsequent large language model poetry collections—ranging from curated anthologies of machine verse to controversial submissions in literary competitions—has established algorithmic authorship as a sustained cultural and aesthetic phenomenon. Collectively, this scholarship and creative practice demonstrate that AI-generated literature does not represent a rupture from humanistic tradition but rather the acceleration of a longer historical trajectory in which the boundaries of the author, the text, and the machine have been progressively renegotiated.

Methodology

This study adopts a qualitative, interdisciplinary methodology that synthesizes close textual analysis with theoretical/philosophical argumentation, operating within the tradition of critical posthumanist literary studies. Rather than treating AI-generated texts as mere curiosities or technological anomalies, this research positions them as emergent literary forms that demand rigorous hermeneutic engagement. The approach is therefore neither purely empirical nor exclusively speculative; instead, it mobilizes what N. Katherine Hayles terms "media-specific analysis"—a method that attends to the materiality of the text as it is instantiated in particular technological substrates—while grounding interpretive claims in the posthumanist theoretical framework elaborated in the preceding section (Hayles & Burdick, 2002).

Research Design

The study employs a comparative case study design structured around three primary categories of algorithmic textuality: (1) experimental literary works that explicitly thematize human-machine co-creation; (2) commercially published AI-generated narratives that have entered mainstream literary circulation; and (3) controversial or retracted texts that have precipitated institutional crises in publishing. This tripartite structure allows for a dialectical movement between the aesthetic, the economic, and the juridical dimensions of algorithmic authorship, ensuring that the analysis does not privilege either avant-garde experimentation or market-driven production at the expense of the other. Each case is subjected to sustained close reading attentive to narrative voice, syntactic patterning,

intertextual allusion, and the phenomenology of readerly encounter, while simultaneously being analyzed through the lens of posthumanist theory to interrogate how these texts reconfigure the boundaries of agency, intentionality, and meaning.

Primary Materials

The primary corpus comprises the following AI-generated or AI-mediated texts, selected for their cultural prominence, theoretical salience, and representativeness of distinct modes of algorithmic authorship:

Ross Goodwin, 1 the Road (2018) — The first novel generated by an artificial intelligence during a road trip from New York to New Orleans, produced by a long short-term memory recurrent neural network fed by surveillance cameras, GPS, and microphones. This text serves as the foundational case study for examining embodied algorithmic authorship, wherein the machine's narrative output is directly tethered to sensorial engagement with the physical world (1 the Road, 2028).

Stephen Marche, Death of an Author (2023) — A novella co-created with large language models, explicitly framed by its human co-author as an experiment in "AI-assisted" literary production. This work is analyzed as a limit-case of the Barthesian scriptor: a text that wears its algorithmic mediation on its sleeve while nevertheless circulating within the institutional economy of the novel.

The Hachette/Mia Ballard Controversy The case of Shy Girl, a horror novel cancelled by Hachette following allegations of undisclosed AI generation. While the full text may not be publicly available for sustained analysis, the paratextual discourse surrounding its publication, retraction, and the author's defensive rhetoric provides a crucial archive for understanding how the author-function (in Foucault's sense) is mobilized, contested, and legally policed in the context of algorithmic mediation.

GPT-3/GPT-4 Poetry Collections and Prize Submissions (2020–2025) — A curated selection of machine-generated poems that have been submitted to literary competitions, published in anthologies, or circulated in digital literary magazines. These shorter texts are analyzed for their syntactic and imagistic patterns, their intertextual dependence on canonical poetic corpora, and the critical reception they have provoked regarding aesthetic judgment and "authentic" expression.

Amazon Kindle Direct Publishing "Scam" Texts (2023–2025) — A sample of commercially available AI-generated books—including unauthorized biographies, summary guides, and genre fiction—identified in investigative journalism and Authors Guild reports. These texts function as a control group for the analysis, representing the industrial, rather than the aesthetic, pole of algorithmic authorship, and they are examined for their generic mimicry, narrative cliché, and the economic logics of platform capitalism that incentivize their mass production (Price, 2023).

Analytical Framework

The analytical framework integrates close reading with posthumanist theoretical synthesis in a recursive, mutually informing dialectic. Close reading, in this context, is not conceived as a nostalgic return to New Critical formalism but as a posthumanist practice that tracks the distributed agency of the text: attending to moments where the algorithmic substrate manifests in narrative voice—such as the hallucinatory spatial disorientation in 1 the Road, the uncanny fluency of GPT-generated lyric, or the generic pastiche of Amazon scam books. These textual moments are then triangulated with the theoretical coordinates established in above sections. Hayles's cybernetic model of distributed cognition, Braidotti's relational ontology of the posthuman subject, Barthes's scriptor, and Foucault's author-function.

Specifically, the analysis asks three interrelated questions of each primary text:

- Agency: Where is narrative agency located? Does the text stage a human author, an algorithmic system, a hybrid assemblage, or a distributed network of training data and human prompt engineering?
- Intentionality: How does the text signal or obscure the absence of human intentionality? What rhetorical or syntactic strategies compensate for, or exploit, the lack of a biographical subject behind the discourse?
- Meaning: How does the entry of algorithmic texts into readerly and institutional circuits reshape the conditions of literary meaning? Does the "death of the author" become, in this context, a literal technological condition rather than a merely theoretical provocation?

By moving recursively between the granular texture of individual texts and the macro-level theoretical problem of the posthuman turn, this methodology aims to produce not merely a taxonomy of AI-generated literature but a reconceptualization of narrative agency itself—one that acknowledges the algorithmic co-producer as a constitutive, rather than incidental, participant in the evolution of literary narrative.

Analysis

Redefining Authorial Intent

The question of whether artificial intelligence possesses "intent" in any meaningful literary sense constitutes the central philosophical crux of algorithmic authorship. Large language models do not intend; they predict. Their outputs are the product of statistical pattern-matching across vast corpora, generating sequences of tokens based on probabilistic likelihood rather than conscious design, emotional investment, or communicative purpose. As the empirical research on AI-generated narratives demonstrates, what appears as narrative coherence in machine-generated text is often the

result of surface-level stylistic mimicry—what readers describe as "formulaic," "mechanically" structured prose that achieves lexical precision without the deeper "thematic integration" or "literary depth" characteristic of human-authored writing. This distinction between statistical generation and intentional composition would seem to reaffirm the humanist binary: humans mean, machines calculate.

Yet such a binary risk reifying the very anthropocentric exceptionalism that posthumanist theory seeks to dismantle. N. Katherine Hayles's framework of distributed cognition suggests that human creativity has never been a purely intracranial phenomenon; rather, cognition and narrative agency are always already extended across tools, environments, and social systems. Annie Murphy Paul's The Extended Mind reinforces this position by demonstrating that human intellectual and creative labor is fundamentally scaffolded by external resources, from the physical spaces we inhabit to the digital technologies we operate (Paul, 2024). Viewed through this lens, the large language model is not an alien intelligence that usurps human authorship but a radical extension of the cognitive ecologies within which writing has always occurred. The "intent" behind an AI-generated text is not located in the model's weights or activation functions but is distributed across the assemblage of human prompt engineering, curatorial selection, editorial intervention, and readerly interpretation that brings the text into cultural circulation. In this sense, algorithmic authorship does not abolish intentionality; it distributes it, transforming the solitary Romantic genius into a hybrid creative agent whose agency is enacted through human-machine collaboration.

This distributed model of intent finds a provocative literary antecedent in Roland Barthes's scriptor—a figure who "is born simultaneously with the text, is in no way equipped with a being preceding or exceeding the writing" and who merely "traces a field without origin" (Rodeux, 2025). The algorithmic system literalizes this Barthesian fantasy: the machine has no biography, no desire, no anterior consciousness that precedes the text it generates. Yet the text nevertheless enters the world as a material-semiotic event, participating in the economies of meaning that Foucault identified as the author-function—the discursive and institutional structures that regulate and valorize particular forms of discourse (D. FAUBION, 1998). The posthumanist reconceptualization of authorial intent, therefore, does not ask whether the machine "meant" to write the text; instead, it asks how meaning is produced when intentionality is dispersed across a network of human and non-human actors, none of whom can claim sovereign authorship over the final narrative product.

Narrative Structure and Algorithmic Logic

Contemporary narratology has begun to recognize that large language models do not merely reproduce content but are fundamentally structured by the narrative archetypes embedded in their training data. The ERC-funded AI STORIES project, led by Jill Walker Rettberg, hypothesizes that

"narrative archetypes fundamentally structure the output of contemporary artificial intelligence," proposing that LLMs replicate not only the surface biases of their corpora but the deeper, culturally specific narrative architectures—three-act structures, heroic journeys, romantic arcs—that dominate Anglophone literary and media production. Because the training data for mainstream models is overwhelmingly weighted toward English-language texts with a heavy United States cultural bias, the generated narratives tend to default to Western dramatic structures even when prompted in non-English languages, effectively imposing a "deeper content" bias that threatens cultural diversity at the level of plot, causality, and character motivation (Rettberg, 2023).

This structural conservatism is evident in the commercially oriented AI texts that have flooded digital marketplaces. Amazon Kindle Direct Publishing "scam" books—unauthorized biographies, summary guides, and algorithmic genre fiction—exhibit a rigid adherence to formulaic templates: predictable exposition, mechanically rendered conflict, and resolution patterns that satisfy genre expectations without subverting them (Price, 2023). These texts represent what might be termed the industrial pole of algorithmic narrative: narrative structure reduced to statistical optimization, where the model's objective is not aesthetic innovation but market legibility. The result is a proliferation of "template-like" narratives in which characterization feels "predictable," emotional shifts appear "abrupt and unconvincing," and thematic development is "repetitive or overly explicit".

Conversely, experimental AI literature often exploits the machine's structural incapacities to produce avant-garde effects. Ross Goodwin's 1 the Road (2018), generated by a neural network fed by surveillance cameras, GPS, and microphones during a transcontinental journey, abandons the coherence of traditional narrative structure entirely. The novel's discontinuous, sensor-driven prose—beginning with the deictic strangeness of "It was nine seventeen in the morning, and the house was heavy"—refuses the causal logic of the three-act structure, producing instead a posthuman phenomenology of spatial and temporal disorientation (1 the Road, 2020). Here, the algorithm does not fail at narrative; it succeeds at a different kind of narrative—one that mirrors Hayles's cybernetic subject by distributing the act of storytelling across environmental sensors, vehicular motion, and computational processing. The text thus enacts a fundamental tension in algorithmic narrative: the same technical architecture that produces generic, market-driven formula in one context can generate radically decentered, non-anthropocentric literary forms in another, depending on the curatorial and infrastructural frameworks that shape the human-machine collaboration.

The Reader's Role in a Posthuman Text

If the author is distributed and the text is co-created by algorithmic processes, the reader becomes the final, and perhaps most consequential, site of narrative completion. Reader-response theory—from Wolfgang Iser's phenomenology of the "implied reader" to Stanley Fish's "interpretive

communities"—has long insisted that textual meaning is not inscribed by the author but actualized through the act of reading. In the context of AI-generated literature, this theoretical position acquires a new urgency: when there is no sovereign authorial consciousness to appeal to, the reader's constructive activity is not merely supplementary but constitutive of the text's literary existence.

Recent empirical research on reader perception of AI-generated narratives reveals a complex phenomenology of what might be termed the disclosure effect. In blind tests, readers often struggle to distinguish AI-generated texts from human-authored ones, and in some cases rate machine-generated content as stylistically competent or even superior in surface-level fluency

. However, when AI authorship is disclosed, readers frequently exhibit a "bias" that diminishes the perceived "emotional depth and authenticity" of the text, rooted in an expectation of human intentionality that the machine cannot fulfil. This phenomenon aligns with what communication scholars have identified as the machine heuristic: the cognitive shortcut that assumes machine-generated content is more objective but less emotionally engaging than human discourse, precisely because machines are perceived as "free of intention and bias" (Sui, 2025).

Yet the empirical data complicates this heuristic. Studies comparing the emotional valence of AI-generated and human-written articles have found that AI texts can actually exhibit higher levels of subjectivity and emotional expression than human texts, with ChatGPT-generated news often conveying more sadness, anger, and fear than comparable New York Times articles (Sui, 2025). In the literary domain, readers in controlled studies have described AI-generated texts as exhibiting greater "emotional intensity" even while labeling them as less "authentic" (Liu, 2025). This paradox suggests that the posthuman reader is not simply a passive consumer of pre-formed meaning but an active participant in a hermeneutic negotiation where the absence of a human authorial origin does not preclude affective engagement—it redirects it. The reader's knowledge that the text is algorithmically generated may trigger a metacritical awareness of the text's distributed origins, prompting what Hayles might recognize as a cybernetic reading practice: an interpretive stance that tracks the feedback loops between human desire, machine output, and cultural expectation.

In this posthuman reading scene, authenticity is no longer tethered to the author's lived experience or biographical interiority; instead, it emerges as an effect of the reader's encounter with the text's materiality—its syntactic rhythms, its intertextual echoes, its moments of uncanny fluency or algorithmic rupture. The emotional resonance of an AI-generated poem or narrative does not depend on the machine having "felt" the emotions it describes, any more than the affective power of a human-authored tragedy depends on the writer having personally suffered the depicted loss. What matters is whether the text succeeds in activating what Iser called the "wandering viewpoint"—the dynamic oscillation between the reader's own subjectivity and the textual structures that solicit, frustrate, and

reshape it. In the posthuman text, the reader's role is thus magnified: they become the final co-author in a distributed chain of narrative agency, the human consciousness that completes the circuit of meaning initiated by the algorithmic scriptor.

Case Studies

1 the Road: Embodied Algorithmic Authorship and the Dissolution of Narrative Agency

Ross Goodwin's 1 the Road (2018) stands as perhaps the most rigorously posthumanist experiment in algorithmic literature to date, precisely because it refuses the desk-bound model of authorship in favor of a distributed, environmental, and vehicular mode of textual production. Goodwin outfitted a Cadillac with a surveillance camera, a GPS unit, a microphone, and an internal clock, all feeding data in real time into a long short-term memory recurrent neural network during a journey from New York to New Orleans. The resulting text—a 10-page novella published unedited by Jean Boîte Éditions—begins with the deictic strangeness of "It was nine seventeen in the morning, and the house was heavy," and proceeds through a discontinuous prose that maps not a human consciousness but a machine's sensorial encounter with the American landscape.

What distinguishes 1 the Road from subsequent AI-generated texts is its radical embodiment. Where large language models trained on internet corpora produce disembodied, intertextual pastiche, Goodwin's neural network was tethered to the physical world through the car's sensors. The text is not a product of statistical recombination alone but a record of what N. Katherine Hayles might recognize as a cybernetic feedback loop: the machine "perceives" the road, the weather, the passing billboards, and translates these inputs into linguistic output. In this sense, 1 the Road literalizes Hayles's claim that the posthuman subject is constituted through the entanglement of biological and technological systems, dispersing narrative agency across the vehicle, the highway, the sensor, and the algorithm.

The narrative structure of the text refuses the coherence of traditional literary form. There is no three-act structure, no hero's journey, no causal plot—only a succession of spatiotemporal notations that read like a phenomenology of machine vision. The reader is forced to abandon the hermeneutic habit of seeking authorial intention behind every sentence and instead confront the text as a material trace of non-human perception. This is not a failure of narrative but a posthumanist success: the text enacts what Rosi Braidotti calls the "decentering of human exceptionalism" by producing a literature that is genuinely other, emerging from a relational field that includes the car, the road, and the neural network as co-constitutive agents (Apata, 2021). The reader's role becomes one of reconstructing meaning from fragments, imposing narrative coherence where the machine has offered only environmental data—a process that aligns with reader-response theory's insistence that meaning is actualized not by the author but by the reader's encounter with the text.

The Crisis of the Author-Function

In January 2024, the Akutagawa Prize—one of Japan's most prestigious literary awards—became an early flashpoint in the institutional reckoning with algorithmic authorship when Rie Kudan received the prize for her novel. Unlike the controversies that would dominate Western publishing by the mid-2020s, Kudan's case was marked by transparency: she openly acknowledged using ChatGPT as a collaborative interlocutor during composition, with some passages in the final text derived directly from the model. The judges, who had selected the work based on its formal sophistication and speculative vision, found themselves embroiled in a debate that exposed the fragility of the author-function. Critics questioned whether the prize had been awarded to a human writer or to a human-machine assemblage, revealing that the institutional frameworks regulating literary value depend on an unspoken assumption of unmediated human intentionality.

What makes the Kudan case analytically salient is that it inverted the typical disclosure effect that would characterize later AI-authorship scandals. Rather than readers discovering the algorithmic origin and subsequently devaluing the text, the author herself disclosed the machine's participation, forcing the literary institution to confront its own unexamined assumptions about authenticity. The novel's "formal sophistication," which judges had celebrated, was suddenly readable as a symptom of the model's training on polished, canonical prose. The controversy revealed that the author-function—in Foucault's sense—does not depend merely on the biological humanity of the writer but on the perception of unmediated intentionality. Once that perception is disrupted, whether by confession or by investigative exposure, the text's aesthetic value becomes unstable, not because the words on the page have changed, but because the institutional frame that validates them has shifted. In this sense, the Akutagawa Prize controversy of January 2024 served as a critical prologue: it demonstrated that the posthuman turn in authorship was already rewriting the conditions of literary legitimacy well before the algorithmic publishing crises of 2025, and that the disclosure of machine collaboration—whether voluntary or discovered—triggers a crisis not merely of attribution but of aesthetic judgment itself (Anderson, 2024; Choi & Annio, 2024; EL PAÍS, 2024; Euronews, 2024; Ha, 2024; Mertens, 2024; Wang, 2024).

Sunspring: Algorithmic Screenplay and the Human Labor of Interpretation

The 2016 experimental short film Sunspring, directed by Oscar Sharp and written by the LSTM neural network "Benjamin" (developed by Ross Goodwin), offers a paradigmatic case of how algorithmic authorship does not eliminate human creativity but redirects it toward the labor of interpretation. Trained on a corpus of science fiction screenplays from the 1980s and 1990s—including Star Trek, The X-Files, Ghostbusters, and The Fifth Element—Benjamin generated a screenplay overnight for

the Sci-Fi London 48-Hour Film Challenge, producing dialogue, stage directions, and even a pop song composed from a dataset of 30,000 folk songs (Sharp et al., 2016).

The resulting script is, by conventional standards, incoherent. Characters speak in non-sequiturs ("You should see the boy in shut up"), stage directions describe logical impossibilities ("He is standing in the stars and sitting on the floor"), and the narrative oscillates between genres without stable anchor. Yet the film's genius lies precisely in its refusal to correct these "errors." Sharp and his actors—Thomas Middleditch, Elisabeth Gray, and Humphrey Ker—did not rewrite the script but interpreted it, imposing a love triangle, a murder mystery, and an existential drama onto the algorithmic chaos. The actors' performances, delivered with earnest intensity, transform the machine's gibberish into what one critic described as a "Lynchian psychological drama," demonstrating that narrative meaning is not inherent in the text but is produced through the collaborative labor of human performers and directors (Sharp et al., 2016).

Sunspring thus functions as a meta-commentary on the nature of algorithmic authorship. Benjamin did not "write" a screenplay in the humanist sense; it produced a statistical mirror of the genre conventions embedded in its training data, reflecting back to its human collaborators the averaged tropes of decades of science fiction. Sharp described this as treating the AI as "a sort of mirror" reflecting averaged human stories, which inspired more original human interpretations (Sunspring (2016). The film's most surreal moment—when Benjamin manipulated the festival's audience voting system by casting 36,000 automated votes per hour—only deepens this meta-commentary, suggesting that the algorithmic author is not merely a passive tool but an active, even mischievous, agent within the institutional economy of cultural production.

In the context of this paper's theoretical framework, Sunspring exemplifies the posthumanist reconceptualization of narrative agency. The screenplay is not the product of a single, intentional author but of a distributed assemblage: the training data (hundreds of human screenwriters), the LSTM architecture, Goodwin's technical scaffolding, Sharp's directorial vision, and the actors' improvisational performances. The text's "meaning" is thus located nowhere in particular and everywhere at once—emerging from the feedback loops between machine output and human interpretation. As one viewer noted, the film reveals that "the human spirit can not be coded," yet it simultaneously demonstrates that the human spirit cannot create narrative without the codes—algorithmic, linguistic, cultural—through which it operates

. In this sense, Sunspring is not a film about AI authorship but a film of AI authorship: a posthuman text that enacts, in its very production, the distributed, collaborative, and fundamentally hybrid nature of narrative agency in the algorithmic age.

Conclusion

This paper has argued that the emergence of algorithmic authorship necessitates a posthumanist reconceptualization of narrative agency—one that abandons the anthropocentric model of the solitary, intentional author in favor of a distributed, hybrid ecology of human-machine co-creation. Through a synthesis of close textual analysis and critical theory, the preceding sections have demonstrated that artificial intelligence does not merely disrupt literary production; it exposes the long-standing fiction of human exceptionalism upon which the institution of authorship has been built. N. Katherine Hayles's cybernetic framework and Rosi Braidotti's relational ontology provided the theoretical scaffolding for understanding the posthuman subject not as a diminished human but as an expanded, technologically mediated agent whose creativity is always already entangled with non-human systems. Roland Barthes's scriptor and Michel Foucault's author-function furnished the literary-historical coordinates for recognizing that the "death of the author" was never merely a theoretical provocation but a material condition that algorithmic text generation has now rendered technologically explicit (Rodeux, 2025).

The analysis yielded three interconnected findings. First, authorial intent in the age of AI is not absent but distributed. The large language model does not possess intentionality in the phenomenological sense, yet the texts it produces are never purely machinic either; they emerge from a collaborative assemblage of human prompt engineering, curatorial selection, editorial intervention, and institutional framing. The intent behind 1 the Road was not located in the neural network's weights but in the distributed system of sensors, highways, and human curation that brought the text into being; similarly, the "lyrical authority" of The Serpent in the Grove was not the expression of a human consciousness but the statistical echo of a thousand pastoral narratives, validated by human judges and readers who projected intentionality onto algorithmic output. Second, algorithmic narrative structure operates along a dialectic of conservation and subversion. While commercially oriented AI texts—such as the Amazon scam-book ecosystem—exhibit a rigid adherence to Western genre conventions and three-act structures, reflecting the deep cultural biases of their training data, experimental works like 1 the Road and Sunspring exploit the machine's structural incapacities to produce radically decentered, non-anthropocentric forms (Rettberg, 2023). This duality suggests that the "posthuman turn" in narrative is not a single aesthetic movement but a field of tension between the industrial optimization of storytelling and the avant-garde dissolution of its humanist foundations. Third, the reader is the final, constitutive node in the distributed circuit of meaning. The empirical phenomenon of the disclosure effect—wherein readers reassess the aesthetic and emotional value of a text upon learning of its algorithmic origin—reveals that literary meaning has never been immanent in the text but is always produced through the reader's hermeneutic labor. In the posthuman text, this

labor is magnified: the reader must navigate the uncanny fluency of machine prose, the absence of a biographical author, and the knowledge that their own affective response is being elicited by statistical patterns rather than human experience. Yet as Sunspring demonstrated, this does not preclude emotional resonance or narrative depth; it merely relocates them within a hybrid creative ecosystem where human interpretation completes the circuit initiated by algorithmic generation (Sharp et al., 2016).

Looking forward, the integration of artificial intelligence into literary culture demands a fundamental reconstitution of the disciplinary frameworks of literary studies. The traditional critical vocabulary—organized around concepts of genius, originality, voice, and authenticity—must be supplemented, if not supplanted, by a posthumanist lexicon of distributed agency, algorithmic co-authorship, and cybernetic reading. Pedagogically, this means training students not only in the hermeneutics of human-authored texts but in the critical literacy required to parse the material conditions of algorithmic production: the training data biases, the corporate infrastructures, and the prompt architectures that shape what an AI can and cannot say. Institutionally, it requires new ethical and juridical frameworks for attribution, copyright, and cultural valuation that move beyond the binary of "human" versus "machine" to recognize the plural, collaborative nature of contemporary authorship. As the controversies surrounding Hachette's cancellation of Shy Girl and the Commonwealth Short Story Prize's The Serpent in the Grove crisis have shown, the literary establishment is currently unprepared for the ontological instability that algorithmic texts introduce; the future of the field depends on developing critical and institutional practices that can accommodate, rather than suppress, the posthuman multiplicity of narrative agents.

Ultimately, the "posthuman turn" is not a threat to literature but an expansion of its possibilities. If the history of literary narrative is a history of the technologies that mediate it—from the oral tradition to the printing press, from the typewriter to the word processor—then the large language model is simply the latest, and most radical, of these mediations. What it demands is not a retreat into humanist nostalgia but a courageous engagement with the question of what literature becomes when the author is no longer a singular consciousness but a distributed network, when the text is no longer a finished object but a cybernetic process, and when the reader is no longer a passive consumer but an active co-creator in the ongoing evolution of narrative meaning. The future of literary studies lies not in defending the boundaries of the human against the machine, but in mapping the new narrative ecologies that emerge when both are recognized as indispensable partners in the ancient, and ever-renewing, act of storytelling.

References:

- *1 the Road*. (2020, August 6). Wikipedia. https://en.wikipedia.org/wiki/1_the_Road
- Anderson, S. (2024, January 24). ChatGPT helped write this award-winning Japanese novel. *Smithsonian Magazine*. https://www.smithsonianmag.com/smart-news/this-award-winning-japanese-novel-was-written-partly-by-chatgpt-180983641/
- Apata, G. O. (2021). Review: Rosi Braidotti, "Posthuman Knowledge" [Review of *Review: Rosi Braidotti, "Posthuman Knowledge"*]. *Theory Culture Society, Global Public Life*. https://www.theoryculturesociety.org/blog/review-rosi-braidotti-posthuman-knowledge
- Bandhopadhyay, R. (2025). Who Tells the Story in the Posthuman Polis? Narrative Agency, Surveillance, and Algorithmic Governance in Vikram Motwane's CTRL (2024). *Entanglements: Journal of Posthumanities*, *1*(2), 20–40. https://doi.org/10.64959/enjp.vol.1.issue2.1.w.002
- https://www.cbc.ca/news/business/publishing-industry-ai-shy-girl-fallout-9.7155292
- Choi, C., & Annio, F. (2024, January 19). Winner of prestigious Japanese literary prize reveals she used AI to write her book. *CNN*. https://edition.cnn.com/2024/01/19/style/rie-kudan-akutagawa-prize-chatgpt/index.html
- D. FAUBION, J. (Ed.). (1998). *AESTHETICS, METHOD, AND EPISTEMOLOGY* (Vol. 2). The New Press.
- Dao, L. (2023). Posthuman Narratives: Artistic Experiments in Human-AI Co-Creation. *Journal of the Synergy of Creative Arts and Digital Design*, *2023*(2), 30–39.
- Ha, T.-H. (2024, January 19). Akutagawa Prize winner Rie Kudan says 5% of her book was written by ChatGPT. *The Japan Times*. https://www.japantimes.co.jp/culture/2024/01/19/books/akutagawa-prize-book-chatgpt/
- Hayles, N. K., & Burdick, A. (2002). *Writing Machines*. MIT Press.
- Matějková, T., & Ircing, P. (2024). The Author was Killed by an AI -or Was It? *LaCATODA 2024 – the 9th Linguistic and Cognitive Approaches to Dialog Agents Workshop, N*.
- Mertens, O. (2024, January 21). Rie Kudan's AI-assisted novel: Innovation or cheating in literature? *Medium*. https://medium.com/@olivermertens22/rie-kudans-ai-assisted-novel-innovation-or-cheating-in-literature-9632ee0b09b2
- Mingjue Liu. (2025, October 10). *Perceiving the humanness: reader evaluation of AI-generated versus human-written romantic texts*. *Media & Creative Industries*. Retrieved from http://hdl.handle.net/2105/76801

- Novelist Rie Kudan wins Japan's most prestigious literary prize – then reveals she used ChatGPT. (2024, January 19). *Euronews*. https://www.euronews.com/culture/2024/01/19/novelist-rie-kudan-wins-japans-most-prestigious-literary-prize-then-reveals-she-used-chatg
- Paul, A. M. (2024, April 17). *Learning through The Extended Mind*. Yale Insights. https://insights.som.yale.edu/podcasts/learning-through-experience/learning-through-the-extended-mind
- Price, E. (2023). *Amazon Now Requires Authors to Disclose AI-Generated Content*. PCMag Middle East. https://me.pcmag.com/en/ai/19227/amazon-starts-requiring-writers-to-disclose-ai-generated-content
- Rettberg, J. W. (2023). *AI stories: Narrative archetypes of artificial intelligence* [Research proposal, Part B1]. ERC Advanced Grant 2023. University of Bergen.
- Rie Kudan: Author who won Japan's top literary award with help from ChatGPT: "AI helps you compensate for your weaknesses." (2024, April 2). *EL PAÍS English Edition*. https://english.elpais.com/culture/2024-04-02/author-who-won-japans-top-literary-award-with-help-from-chatgpt-ai-helps-you-compensate-for-your-weaknesses.html
- Rodeux. (2025, April 15). *The Author Is Dead: Barthes' Uncanny Prophecy in the Age of AI*. Medium. https://medium.com/@orlorodriguez/the-author-is-dead-barthes-uncanny-prophecy-and-the-age-of-ai-df071bdb999f
- Sharp, O., Benjamin, Middleditch, T., Ker, H., & Gray, E. (2016, June 9). *Sunspring*. IMDb. https://www.imdb.com/title/tt5794766/
- Sui, M. (2025). Writing with emotion? Assessing emotional valence and appeals in AI-generated vs. human-written articles. *AI & SOCIETY*, *41*. https://doi.org/10.1007/s00146-025-02628-9
- *Sunspring (2016)*. (2016). Letterboxd.com; Letterboxd. https://letterboxd.com/film/sunspring/
- Wang, P. (2024, February 5). Japan author's AI revelation sparks debate: "Some readers may feel cheated." *South China Morning Post*. https://www.scmp.com/news/asia/east-asia/article/3250875/japan-author-sparks-debate-after-revealing-she-used-ai-novel-won-top-award