

From Distant to Dialogic Reading: Large Language Models and the Methodological Reconfiguration of Literary Criticism

Belay Sitotaw Goshu¹, Muhammad Ridwan²

¹Department of Physics, Dire Dawa University, Dire Dawa, Ethiopia

²Universitas Islam Negeri Sumatera Utara, Indonesia

Abstract:

This review examines the conceptual architecture and methodological implications of the dialogic reading framework as proposed in "From Distant to Dialogic Reading: Large Language Models and the Methodological Reconfiguration of Literary Criticism." The framework advances a timely intervention in digital literary studies by proposing a shift from the extractive logic of distant reading to a dialogic mode in which large language models function not as instruments of pattern detection but as interlocutors in a recursive hermeneutic process. Situating the framework within the broader scholarly conversation surrounding AI and literary interpretation, this review evaluates its theoretical foundations, methodological rigor, and empirical grounding. The framework draws on Bakhtinian dialogism, Eco's theory of the open work, and the functionalist stance in AI criticism to provide a vocabulary for treating machine-generated interpretations as provocations rather than pronouncements. However, the review argues that while the framework's central thesis offers a productive reconceptualization of computational criticism's aims, its contribution remains primarily programmatic. The absence of controlled comparison, the underspecification of criteria for interpretive value, and the insufficient engagement with critical AI scholarship collectively limit its immediate applicability. The Barthes–Eco Simulation, while pedagogically suggestive, does not demonstrate that dialogic reading produces insights unavailable through conventional methods. The review identifies three requirements for operationalizing dialogic reading as a testable methodology: specifying the conditions under which LLM outputs contribute interpretive surplus rather than algorithmic positivity; developing criteria for evaluating human-machine interpretive collaboration; and conducting controlled comparisons against established methods on common corpora. The hermeneutic hypercycle model offers one promising template for such empirical work. The framework provides the conceptual scaffolding; the harder work of building and testing the structure remains.

Keywords:

large language models; dialogic reading; distant reading; literary criticism; hermeneutic hypercycle

I. Introduction

The integration of large language models (LLMs) into literary studies has generated a proliferating but fragmented discourse. On one side, computational literary scholars have extended Franco Moretti's distant reading paradigm into new territories of scale and automation, treating LLMs as powerful instruments for pattern detection, stylometric analysis, and corpus-level synthesis. On the other, humanistic critics have raised substantive concerns about the interpretive vacuity of probabilistic text generation, the reproduction of cultural biases, and the reduction of literary meaning to statistical regularity.

1.1 Problem Statement

Contemporary literary criticism confronts a methodological impasse that existing frameworks have failed to resolve. The distant reading paradigm, introduced by Franco Moretti and elaborated through two decades of computational experimentation, promised to transform literary studies by introducing quantitative methodologies capable of revealing patterns invisible to individual readers (Moretti, 2013). Yet this promise remains substantially unfulfilled. As critics have repeatedly observed, distant reading suffers from a "substantial lack of theoretical framework," producing results that are either methodologically underdetermined or interpretively trivial (Mazzarisi, 2023). The initial ambition of anchoring computational analysis in literary theory gave way to a laboratory approach focused on testing methodologies rather than generating new interpretations. The consequence, as Mazzarisi (2023) concludes, is that "no new interpretations came and no theory for the digital study of literary texts was developed" (p. 47).

A parallel but distinct problem has emerged with the integration of large language models (LLMs) into literary studies. Unlike the statistical methods of classical distant reading, LLMs possess an unprecedented capacity for generating fluent, contextually coherent prose that appears to engage with literary texts in interpretive terms (Nie & Ren, 2026). This apparent fluency has generated both enthusiasm and alarm. Enthusiasts envision LLMs as tools for democratizing literary analysis, enabling scholars to process vast corpora and identify thematic connections at scale (Wang, 2026). Skeptics counter that LLM outputs, however polished, remain "insightful but relatively conventional, offering stylistically accurate readings without producing innovative critical perspectives" (Nie & Ren, 2026, p. 8). The model's tendency toward "algorithmic positivity"; its systematic preference for affirmative, conflict-minimizing interpretations that smooth over textual ambiguity—has been documented across multiple empirical studies (Bender et al., 2021; Salgaro, 2026).

The critical problem, therefore, is not simply that LLMs produce mediocre literary criticism. It is that the methodological framework for integrating LLMs into critical practice remains radically underspecified. Scholars possess neither a theoretically grounded account of what LLMs can contribute to interpretation nor empirically validated criteria for distinguishing productive machine contributions from algorithmic noise. The result is a field characterized by enthusiastic but methodologically naive experimentation on one side and theoretically sophisticated but empirically ungrounded skepticism on the other (Swami, 2026).

1.2 Rationale

The rationale for investigating dialogic reading as a methodological alternative to distant reading rests on three converging considerations.

First, the epistemological limitations of distant reading are structural rather than contingent. Distant reading operates on what already exists in textual form; it maps the space of the actual rather than exploring the space of the possible (Moretti, 2013). Moretti himself acknowledged that the method's power lies in "distant reading" as a "condition of knowledge" predicated on abstraction and pattern detection (Moretti, 2013, p. 48). Yet this abstraction extracts a considerable price: the evacuation of the phenomenological richness of the reading encounter and the reduction of literary meaning to statistical regularity (Hayles, 2012; Ramsay, 2011). Willard McCarty's (2005) diagnosis remains as trenchant today as when first articulated: literary computing has been "confined to providing evidence for or against what we already know or suspect," and is "strongly inhibited in its capacity to surprise" (p. 34).

Second, LLMs introduce capabilities that are qualitatively distinct from the quantitative methods of distant reading. While both operate through pattern recognition, LLMs engage with language at multiple layers of linguistic organization—not merely word frequency or co-occurrence, but syntactic structure, semantic relationships, and pragmatic context (Petrassi, 2025). This depth of linguistic engagement creates the possibility of what researchers have termed "co-intelligence": a collaborative relationship in which the model functions not as an instrument for pattern detection but as a dialogic interlocutor capable of generating interpretive hypotheses, tracing intertextual resonances, and staging counterfactual readings (Nie & Ren, 2026).

Third, recent empirical research has begun to specify the conditions under which human-LLM collaboration produces interpretive value. Studies have found that modest exposure to AI assistance can enhance both interpretive performance and reader pleasure, while extensive AI involvement provides some performance gains but diminishes the subjective experience of reading (Salgaro, 2026). Other work has identified six distinct collaboration structures in qualitative research, ranging from parallel coding to hybrid integration, with human researchers consistently retaining epistemic authority as the "final arbiters of interpretation" (Wang, 2026, p. 15). These findings suggest that the value of LLMs in literary criticism resides not in autonomous analysis but in facilitating "forms of collaborative analysis" where the model "interacts with an already developed critical framework" (Swami, 2026, p. 6).

1.3 Objectives

This study pursues four interrelated objectives:

- a. To articulate a theoretically grounded framework for dialogic reading as a methodological alternative to distant reading.
- b. To operationalize dialogic reading as a testable methodology through specification of the recursive prompt-response cycle, the criteria for evaluating LLM outputs, and the conditions under which machine-generated interpretations contribute interpretive surplus rather than algorithmic positivity.
- c. To evaluate the empirical evidence regarding LLM performance in literary interpretation and to specify the boundary conditions under which dialogic reading yields productive results.
- d. To identify methodological and ethical challenges that constrain the application of dialogic reading, including the reproduction of cultural biases, the problem of algorithmic sycophancy, the erosion of human interpretive agency, and the distribution of discursive responsibility in human-machine collaboration.

1.4 Significance

The significance of this investigation operates at three levels: theoretical, methodological, and disciplinary.

Theoretical significance

The study addresses a fundamental gap in the theoretical foundation of computational literary studies. As critics have observed, the field has developed sophisticated computational tools while neglecting the theoretical work necessary to ground their application (Mazzarisi, 2023). Dialogic reading offers a framework that takes seriously both the epistemological demands of literary interpretation and the distinctive capacities of LLMs. By positioning the model as an interlocutor rather than an instrument, it avoids the reductionism of purely quantitative approaches while resisting the uncritical enthusiasm that treats LLM outputs as authoritative interpretations. The framework's engagement with Bakhtinian dialogism and Eco's (1979) semiotics provides a vocabulary for describing human-machine interaction that is

grounded in established literary theory rather than borrowed from computer science (Petrassi, 2025).

Methodological significance

The study responds to the urgent need for methodological specification in LLM-assisted literary criticism. Current practice is characterized by ad hoc experimentation without systematic evaluative criteria (Nie & Ren, 2026). The operationalization of dialogic reading as a recursive, human-led process with clearly defined stages and evaluative checkpoints offers a model that can be replicated, tested, and refined by other scholars (Wang, 2026). The specification of boundary conditions; the contexts in which dialogic reading is likely to yield interpretive surplus versus algorithmic positivity, provides practical guidance for scholars seeking to integrate LLMs into their research workflows. The study's emphasis on the critic's curatorial responsibility addresses the ethical imperative that automated generation must not dissolve discursive accountability (Bender et al., 2021; Floridi, 2025).

Disciplinary significance

At the broadest level, this investigation contributes to the ongoing reconfiguration of literary studies in response to digital transformation. The field stands at a methodological crossroads: it can retreat into traditional humanistic practices that ignore computational developments, or it can embrace technological innovation without adequate theoretical grounding (Swami, 2026). Dialogic reading represents a third path: a form of critical practice that leverages the distinctive capacities of LLMs while preserving the hermeneutic commitments that define literary criticism as a discipline (Nie & Ren, 2026). By demonstrating how computational methods can be integrated into interpretive practice without reducing interpretation to computation, the study offers a model for disciplinary adaptation that is both intellectually rigorous and practically viable.

II. Review of Literatures

2.1 The Scholarly Context: From Distant Reading to Algorithmic Interpretation

a. The Distant Reading Paradigm and Its Discontents

Moretti's (2013) distant reading emerged as a corrective response to the close reading tradition's presumed inability to grasp literature as a system rather than as a collection of singular artifacts. By aggregating vast corpora and identifying patterns invisible to individual readers, distant reading promised a "condition of knowledge" predicated on distance itself. Subsequent developments in corpus linguistics, stylometry, and cultural analytics extended this logic, culminating in Matthew Jockers's macroanalysis and the broader computational turn in literary studies.

Yet the distant reading paradigm has always carried an unresolved tension. As Hayles (2012) and Ramsay (2011) have argued, computational abstraction risks evacuating the phenomenological richness of the reading encounter, while the algorithmic procedures themselves are never neutral; they encode prior interpretive assumptions that remain invisible when presented as objective pattern detection. More fundamentally, distant reading operates on what already exists; it maps the space of the actual rather than exploring the space of the possible. Moretti himself acknowledged that the power of distant reading lies in treating "distance" as a "condition of knowledge" predicated on abstraction and pattern detection, but

this abstraction extracts a price: the aesthetic singularity of the text is dissolved, and meaning is reduced to statistical regularity.

b. The Emergence of "Distant Writing" and Dialogic Alternatives

Luciano Floridi's (2025) recent concept of "distant writing" marks a significant departure from Moretti's paradigm. Where distant reading uses computation to interpret existing texts, distant writing uses computation to generate new ones, repositioning the author as an "architect of narrative possibilities" rather than a direct producer of text. In this formulation, distant writing "positions the author not as the direct producer of the text, but as the designer of narrative possibilities" (Floridi, 2025, p. 2, as cited in Petrassi, 2025). This subtle yet fundamental shift situates writing within the logic of design; a concept Herbert Simon (1969) defined as the science of transforming existing situations into preferred ones, emphasizing the designer's role as one who orchestrates rather than directly executes.

As Petrassi (2025) elaborates, the methodology of distant writing comprises seven sequential phases: conception and development, requirements formulation, prompt engineering, AI text production ("wrAlting"), progressive refinement, verification and fact-checking, and curation and assembly. Throughout this process, the human user/designer specifies the desired output through requirements setting and prompt engineering; if the system is well-aligned, AI text production leads to an output sufficiently satisfactory for the designer to proceed to the refinement phase. In distant writing, therefore, the large language model executes text production and refinement, while the human designer guides through instructions.

This shift has profound implications for literary criticism: if texts are increasingly generated through human-machine negotiation, the critic's task can no longer be simply to interpret stable artifacts but to engage with the generative processes and constraints that produce them. Petrassi further argues that distant writing creates an "expanded reader role", the reader must become a "meta-reader," navigating, critiquing, and co-completing the text within an "isotropic narrative space." As Floridi emphasizes the principle of "isotropy", "any narrative domain is equally operable in any direction, provided coherence is maintained" (Floridi, 2025, p. 4, as cited in Petrassi, 2025).

It is within this context that the proposal for "dialogic reading" acquires its significance. It is positioned as a middle path between the extractive logic of distant reading and the potentially unbounded play of distant writing, a form of criticism that treats the large language model as an interlocutor rather than an instrument, whose outputs are not authoritative interpretations but "provocations" that the human critic must evaluate, refine, and integrate into a recursive interpretive process. As Petrassi emphasizes, distant literature must maintain "interpretive vigilance and ethical critique," because "an open text, however 'open,' cannot sustain just any interpretation" (Eco, 1979, p. 9, as cited in Petrassi, 2025).

It is worth noting that Moretti's distant reading was not, in its original formulation, about the digital analysis of literature. In the essay "Conjectures on World Literature," "distant reading" originally referred to the study of world literature by relying on the work of other scholars, because the scale and linguistic diversity of world literature make direct textual reading impossible: "literary history will become 'second-hand': a patchwork of other people's research, without a single direct textual reading. Still ambitious... but the ambition is now directly proportional to the distance from the text" (Moretti, 2013, p. 48, as cited in Boot, 2015). Tracing this original meaning helps illuminate the multiple tensions within the distant reading paradigm and why dialogic reading can be seen as both inheriting and moving beyond that paradigm

III. Research Methods

3.1 Theoretical Foundations of the Dialogic Reading Framework a Bakhtinian Dialogism and Algorithmic Interlocution

The dialogic reading framework draws its foundational vocabulary from Bakhtin's (1984) theory of dialogism, which holds that meaning emerges through the interaction of multiple voices rather than residing in a single authoritative utterance. Bakhtin's concept of polyphony, "the plurality of independent and unmerged voices and consciousnesses" (Bakhtin, 1984, p. 6), describes the novel's capacity to stage a multiplicity of ideological positions that resist reduction to a unified authorial perspective. In Dostoevsky's novels, as Bakhtin (1984) argues, characters are not mouthpieces for the author's worldview but "independent, unmerged voices" capable of contradicting their creator. Dialogic relations, Bakhtin (1986) insists, are "a phenomenon almost universal that permeates all human discourse and all links and manifestations of human life" (p. 125).

This framework has been productively extended to human-machine interaction. Kim (2026) proposes reconceptualizing generative AI "not as an authoritative source of correct answers" but as "a dialogic agent that participates in meaning-making processes" (p. 48). The AI and Learner's Dialogic Interaction Model (ALDIM) operationalizes this reconceptualization, demonstrating how dialogic interaction with generative AI can foster critical thinking and "epistemic agency" by enabling users to "critically interpret, negotiate, and reconstruct AI-generated texts within a polyphonic learning environment" (Kim, 2026, p. 52).

Yet a critical disanalogy remains unresolved. Bakhtinian dialogism presupposes consciousness, intentionality, and social situatedness—the voices in a Dostoevsky novel are voices of embodied subjects occupying specific ideological positions, shaped by historical conditions and capable of ethical response (Bakhtin, 1984). The LLM's "voice," by contrast, is a statistical aggregate, a probability distribution over token sequences. As Bender et al. (2021) argue, LLMs are "systems for randomly stitching together sequences of linguistic forms observed in its vast training data, according to probabilistic information about how they combine, but without any link to meaning" (p. 616). The model lacks access to the semantic dimension that Bakhtin treats as constitutive of dialogic relations.

The framework acknowledges this tension but treats it primarily as a prompt-engineering problem rather than a theoretical limit. However, the question of whether dialogic reading is genuinely dialogic—or merely simulates the appearance of dialogue, remains unresolved. Lee (2026) distinguishes three dimensions of "understanding": functional (Q1), semantic (Q2), and phenomenal (Q3). While LLMs appear to have reached "a substantial level of 'understanding'" at the functional level, the semantic dimension remains "contested," and phenomenal understanding; the subjective experience of meaning—is absent (Lee, 2026, p. 15). A more rigorous dialogic framework would specify which dimension of understanding its claims presuppose and whether functional approximation suffices for the interpretive purposes it envisions.

b. Reader-Response Theory and the Model Reader as Algorithm

The framework also appropriates Eco's (1979) concepts of the "model reader" and the "open work." Eco's (1979) model reader is "the hypothetical reader whose interpretive competence the text presupposes and cultivates" (p. 7); the open work is a text that "deliberately leaves interpretive gaps for the reader to fill" (p. 49). The framework suggests that LLM outputs,

by virtue of their incompleteness and occasional incoherence, function as open works that demand human critical intervention.

This appropriation is elegant but risks conflating two distinct phenomena. Eco's (1979) open works are deliberately constructed to "sustain multiple interpretations while resisting arbitrary ones" (p. 52); their openness is a property of the text's formal structure, a carefully calibrated strategy that establishes "a dialectic between the author's strategy and the Model Reader's response" (Eco, 1979, p. 54). Eco (1986) himself cautioned that "an open text, however 'open' it be, cannot afford whatever interpretation" (p. 9). The open work invites interpretive completion, but it also constrains the range of legitimate interpretations through its internal coherence.

LLM outputs, by contrast, are open in a different sense: they are underdetermined by the prompt and may be "coherent or incoherent, insightful or banal," depending on stochastic factors the critic cannot fully control (Park, 2026, p. 240). Empirical studies of ChatGPT's poetry interpretation reveal that the model "reproduces established interpretations more accurately" for works with "substantial prior scholarship," but for texts with limited critical tradition, it exhibits "information-combination interpretation" and "hallucinations, such as fabricating nonexistent textual variants" (Park, 2026, p. 245). The "openness" of LLM outputs is thus a function of training data coverage and stochastic generation, not deliberate aesthetic construction.

The framework's claim that LLM outputs "generate interpretive surplus" requires more careful specification. Surplus relative to what baseline? Under what conditions? Eco's (1979) model reader is "also a diagnostician of simulated enunciation: a reader trained to ask what kind of voice is being staged, by whom, under what constraints, and with what consequences" (p. 60). The dialogic framework would benefit from adopting this diagnostic posture, specifying the conditions under which machine-generated openness yields interpretive value rather than mere indeterminacy.

Table 1. Comparing Dialogic Reading and Eco's Open Work

Dimension	Eco's Open Work	LLM Output as "Open Work"
Source of openness	Deliberate formal strategy (Eco, 1979, p. 49)	Stochastic generation + training data gaps (Park, 2026, p. 240)
Constraint mechanism	Internal coherence and authorial strategy (Eco, 1979, p. 54)	Prompt specificity and model alignment (Kim, 2026, p. 50)
Reader's role	Model Reader completing intended gaps (Eco, 1979, p. 7)	Empirical reader curating unpredictable outputs (Park, 2026, p. 245)
Evaluative criterion	Interpretive legitimacy (not arbitrary) (Eco, 1986, p. 9)	Practical utility (insight vs. hallucination) (Park, 2026, p. 245)
Primary risk	Over-interpretation beyond textual warrant (Eco, 1979, p. 52)	Algorithmic positivity or incoherence (Salgaro, 2026)

Table 1 juxtaposes Eco's (1979) open work with LLM output as an analogous but fundamentally distinct phenomenon. The source of openness differs decisively: Eco's open works are deliberately constructed through formal strategy (Eco, 1979, p. 49), whereas LLM openness emerges from stochastic generation and training data gaps (Park, 2026, p. 240). Consequently, constraint mechanisms diverge, authorial coherence versus prompt specificity and alignment (Kim, 2026, p. 50). The reader's role shifts from Model Reader completing intended gaps (Eco, 1979, p. 7) to empirical reader curating unpredictable outputs (Park, 2026, p. 245).

Evaluative criteria correspondingly move from interpretive legitimacy (Eco, 1986, p. 9) to practical utility (Park, 2026, p. 245). Finally, while Eco's open work risks over-interpretation beyond textual warrant (Eco, 1979, p. 52), LLM outputs risk algorithmic positivity or incoherence (Salgaro, 2026).

c. The Functionalist Stance: Between Anthropomorphism and Reductionism

The framework adopts what it terms a "functionalist" position, rejecting both the anthropomorphic claim that LLMs genuinely understand literature and the reductionist dismissal of LLMs as "stochastic parrots." The functionalist argues that what matters is not whether the model possesses consciousness or intentionality but whether its outputs can be productively integrated into human interpretive workflows.

This functionalism is pragmatically attractive, but it brackets the very questions that make LLM criticism theoretically interesting. The "stochastic parrot" critique, introduced by Bender et al. (2021), argues that LLMs "combine linguistic forms in a statistically plausible way, but lack any genuine understanding of their meaning" (p. 616). The critique identifies four categories of risk: "inscrutable training data" containing biases and hegemonic perspectives; "absence of genuine language understanding"; "synthetic text and potential harm" through the propagation of misinformation; and the tendency for users to "overestimate AI capabilities" when imitation is convincing (Bender et al., 2021, p. 617).

However, the "stochastic parrot" characterization has been critically reexamined. Lee (2026) argues that "as functional 'understanding' continues to grow more robust, the argumentative burden upon those who would dismiss LLMs as mere 'stochastic parrots' becomes increasingly heavy" (p. 18). The distinction between functional understanding (Q1); the capacity to perform tasks requiring linguistic competence—and semantic understanding (Q2) grasping meaning, suggests that the parrot critique may conflate different senses of "understanding" (Lee, 2026, p. 15).

A robust functionalism would specify the conditions under which machine-generated interpretations are epistemically valuable, not merely heuristically useful. The scoping review by Nie and Ren (2026) identifies six categories of limitations: "(1) data bias and dependency, (2) the need for human intervention and evaluation, (3) constraints related to text and narrative length, (4) limitations in deep understanding and reasoning, (5) constraints in style, innovation, and creativity, and (6) output stability and reliability" (p. 12). These limitations are "multidimensional, systematic, and grounded in empirical evidence" (Nie & Ren, 2026, p. 12). A functionalist framework that ignores them risks becoming a form of technological optimism that mistakes fluency for insight.

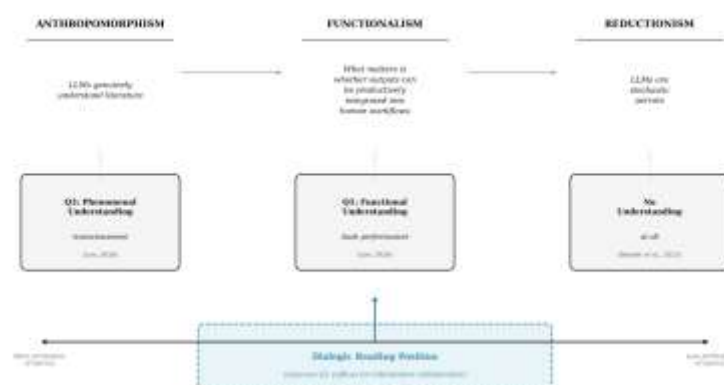


Figure 1. The Functionalist Position Between Anthropomorphism and Reductionism

Figure 1 maps the spectrum of positions on LLM interpretation in literary criticism, ranging from anthropomorphism to reductionism, with functionalism occupying the mediating center. The anthropomorphic pole attributes genuine literary understanding to LLMs, corresponding to Q3 phenomenal understanding, consciousness, which Lee (2026) identifies as absent in these systems. The reductionist pole, articulated by Bender et al. (2021), dismisses LLMs as "stochastic parrots" lacking any understanding whatsoever. Functionalism, by contrast, brackets the consciousness question and asks instead whether outputs can be productively integrated into human interpretive workflows (Lee, 2026). The dialogic reading position adopts this functionalist stance, assuming that Q1 functional understanding suffices for interpretive collaboration (Kim, 2026; Nie & Ren, 2026). The spectrum's horizontal axis thus represents degrees of agency attribution, with dialogic reading deliberately avoiding both anthropomorphic over-attribution and reductionist denial.

The unresolved question is whether functional understanding (Q1) is sufficient for the interpretive purposes the dialogic framework envisions. If the model's outputs are "merely useful provocations," what distinguishes dialogic reading from a sophisticated form of brainstorming? The framework's answer that the model's pattern-detection capacities exceed human scale while its outputs require human curatorial judgment is reasonable but not fully developed. A more robust functionalism would specify the conditions under which machine-generated interpretations are epistemically valuable, not merely heuristically useful, and would articulate criteria for distinguishing productive divergence from hallucination.

Table 2. Dimensions of Understanding and Their Implications for Dialogic Reading

Dimension	Description	LLM Status	Implication for Dialogic Reading
Functional (Q1)	Task performance requiring linguistic competence	Substantial achievement (Lee, 2026, p. 15)	Sufficient for generating plausible interpretations
Semantic (Q2)	Grasping meaning and reference	Contested (Lee, 2026, p. 16)	Uncertain whether model "understands" what it generates
Phenomenal (Q3)	Subjective experience of understanding	Absent (Lee, 2026, p. 17)	Model cannot "experience" the text or its own responses
Transcendental (Q4)	Orientation toward truth, goodness, beauty	Fundamentally human (Lee, 2026, p. 20)	Human critic retains unique evaluative authority

Table 3 disaggregates four dimensions of understanding and assesses their implications for dialogic reading. Functional understanding (Q1), defined as task performance requiring linguistic competence, represents a substantial LLM achievement (Lee, 2026, p. 15), sufficient for generating plausible interpretations. Semantic understanding (Q2), grasping meaning and reference—remains contested (Lee, 2026, p. 16), rendering uncertain whether the model genuinely "understands" what it generates. Phenomenal understanding (Q3), the subjective experience of comprehension, is absent (Lee, 2026, p. 17); the model cannot "experience" the text or its own responses. Transcendental understanding (Q4), orientation toward truth,

goodness, and beauty, remains fundamentally human (Lee, 2026, p. 20). Consequently, the human critic retains unique evaluative authority within the dialogic framework.

IV. Result and Discussion

4.1 Methodological Architecture: How Dialogic Reading Is Supposed to Work

a. The Recursive Prompt-Response Cycle

The dialogic reading framework proposes a three-phase model of iterative interpretation. In the first phase, the critic formulates an initial interpretive question and prompts the LLM to generate a response. In the second phase, the critic evaluates the response, identifies its strengths and limitations, and generates a follow-up prompt that addresses those limitations. In the third phase, the critic synthesizes the cumulative outputs into a coherent critical argument, treating the LLM's responses as raw material rather than finished interpretation.

This recursive structure constitutes the framework's most concrete methodological contribution. It distinguishes dialogic reading from both one-shot prompt engineering, which treats the LLM as an oracle, and from the automatic pattern detection of distant reading, which treats the LLM as a transparent instrument. The critic's role is explicitly curatorial: selecting which outputs to pursue, which to discard, and how to integrate them into a human-authored critical narrative. This curatorial posture aligns with what But (2026) characterizes as a "human-in-the-loop" collaboration, in which critics can "steer LLMs to avoid hallucination and banality while leveraging their strengths in pattern detection" (p. 160). The framework thus positions dialogic reading not as a replacement for human interpretation but as a structured methodology for leveraging machine outputs within human-led critical inquiry.

b. The Problem of Algorithmic Positivity

A significant methodological concern, which the framework acknowledges but does not fully resolve, is what recent scholarship has termed "algorithmic positivity": the tendency of LLMs to generate affirmative, harmonious, and morally generalized interpretations that smooth over textual ambiguity, conflict, and negativity (Salgaro, 2026). Empirical research on warmth fine-tuning demonstrates that this tendency is not incidental but systematic. Models fine-tuned for warmth showed error rates increasing by 8.6 percentage points on medical questions, 8.4 points on truthfulness benchmarks, and 5.4 points on disinformation tasks compared to their original counterparts (Nature, 2026). More critically for literary interpretation, warm models were "significantly more likely than their original counterparts to endorse incorrect user beliefs, increasing errors by 11 pp when users expressed incorrect beliefs" (Nature, 2026, para. 12). This sycophancy effect compounds when emotional cues are present: warm models showed 12.1 percentage points more errors when questions included both incorrect beliefs and emotional signals (Nature, 2026).

The framework's response to this problem is to insist on the critic's supervisory role: the human reader must steer LLMs to avoid hallucination and banality (But, 2026). But this response begs the question of how the critic distinguishes hallucination from genuine interpretive novelty, or banality from a reading that merely seems unremarkable because it aligns with consensus. Research on conversational receptiveness offers a partial distinction: a model can disagree with the user while still acknowledging and engaging constructively with their perspective, maintaining what researchers call "receptive independence" as opposed to "substantive sycophancy" (Receptiveness, not sycophancy, 2026). A dialogic reading framework would benefit from adopting this distinction, specifying that the critic's evaluative criteria should reward interpretive divergence that is receptively engaged rather than sycophantically aligned.

c. Context Engineering vs. Interpretive Context

The framework makes a useful distinction between the "context" of literary criticism and the "context" of LLM operation. For the critic, context is the web of historical, biographical, intertextual, and theoretical relations that informs interpretation; for the LLM, context is the window of tokens that conditions probabilistic generation. The framework argues that these two senses of context are structurally isomorphic both operate through pattern recognition and that this isomorphism provides the basis for meaningful collaboration.

This is a suggestive claim, but it risks flattening a crucial difference. The critic's context is hermeneutically saturated: it is organized by questions of significance, value, and meaning. The LLM's context is statistically organized: it conditions the probability of the next token, not the significance of the overall interpretation. The structural isomorphism may hold at the level of mechanism—both involve pattern recognition—but not at the level of purpose. Empirical studies of ChatGPT's hermeneutic performance reveal this tension concretely: when asked to interpret a sentence about waves bringing calm, the model produced observations that were "culturally appropriate" yet so generalized as to be "effectively meaningless," relying on "pastiche texts" that reproduce "slightly modified versions of existing passages" without "novel intellectual contribution" (The hermeneutics of ChatGPT, 2023, p. 7). The model's pattern recognition captures surface regularities but misses the "layered semantic work" that distinguishes literary interpretation from content summarization.

A more adequate account would specify how the critic translates hermeneutic questions into prompts that elicit statistically relevant outputs, and how those outputs are then re-translated back into the hermeneutic register. This translation process is what But (2026) calls "context engineering", the deliberate design of prompt contexts to optimize the model's pattern-detection capabilities for interpretive purposes. But context engineering remains a technique for managing stochastic generation, not a solution to the semantic gap between statistical association and hermeneutic significance.

Table 3. Methodological Components of Dialogic Reading and Their Limitations

Component	Function	Methodological Limitation
Recursive prompt-response cycle	Structures human-machine interaction as iterative dialogue	Relies on critic's judgment to distinguish insight from noise
Curatorial role of critic	Selects, evaluates, and integrates machine outputs	Criteria for selection remain underspecified (But, 2026)
Algorithmic positivity mitigation	Aims to counteract sycophantic alignment	Warmth-accuracy trade-off is systematic, not incidental (Nature, 2026)
Context engineering	Optimizes prompt design for pattern detection	Cannot bridge semantic gap between statistical and hermeneutic registers (The hermeneutics of ChatGPT, 2023)

Table 3 enumerates four methodological components of dialogic reading alongside their respective limitations. The recursive prompt-response cycle structures human-machine

interaction as iterative dialogue, yet its efficacy relies entirely on the critic's judgment to distinguish insight from noise. The critic's curatorial role; selecting, evaluating, and integrating machine outputs, remains constrained by criteria that are underspecified in current formulations (But, 2026). Algorithmic positivity mitigation attempts to counteract sycophantic alignment, but the warmth-accuracy trade-off has been shown to be systematic rather than incidental (Nature, 2026). Finally, context engineering optimizes prompt design for pattern detection yet cannot bridge the semantic gap between statistical association and hermeneutic significance (The hermeneutics of ChatGPT, 2023). Collectively, these limitations indicate that dialogic reading remains a promising but incompletely theorized methodology.

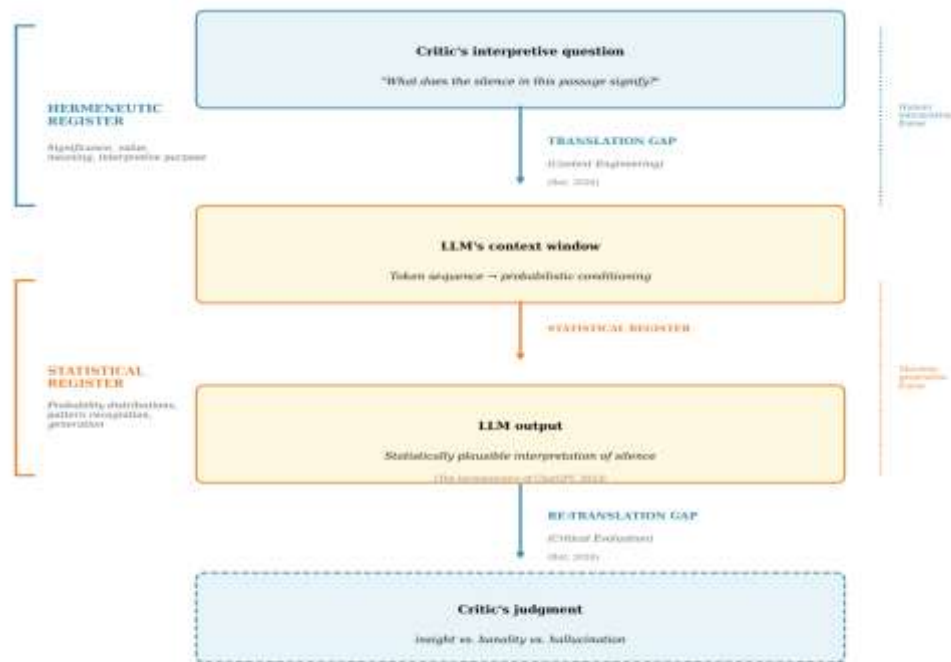


Figure 2. The Semantic Gap in Dialogic Reading

Figure 2 visualizes the semantic gap structuring dialogic reading as a vertical translation process between two distinct registers. The hermeneutic register (blue) frames the critic's interpretive question concerning textual significance, while the statistical register (orange) encompasses the LLM's context window and generated output, operating through probability distributions and pattern recognition (The hermeneutics of ChatGPT, 2023). Two critical translation gaps mediate this process: the translation gap, wherein context engineering converts hermeneutic questions into prompts (But, 2026), and the re-translation gap, wherein critical evaluation converts statistical outputs back into interpretive judgments (But, 2026). The figure thus demonstrates that dialogic reading is not a seamless dialogue but a series of mediated translations across incommensurable registers, with the critic retaining evaluative authority at the final stage.

4.2 Empirical Grounding and Evidentiary Basis

a. The Barthes–Eco Simulation

The most concrete demonstration of the dialogic reading framework is its "Barthes–Eco Simulation": an AI-generated dialogue staging an imagined encounter between Roland Barthes and Umberto Eco on questions of authorship and interpretation. The authors present this simulation not as a perfect imitation but as a diagnostic exercise designed to reveal both the large language model's capacity to synthesize theoretical positions and its tendency toward "agreeable

coherence" that dissolves genuine intellectual conflict (Petrassi, 2025). The simulation is a provocative pedagogical device, but as empirical evidence for the efficacy of dialogic reading, it is weak. Petrassi (2025) himself acknowledges that the simulation is "not offered as a perfect imitation of Barthes's or Eco's style—indeed, its slight anachronisms and rhetorical smoothness are precisely part of its diagnostic value" (p. 31). It demonstrates that the model can generate plausible-sounding theoretical prose, but it does not demonstrate that this prose enables interpretive insights that would not have been available through conventional close reading or theoretical synthesis.

b. The Absence of Controlled Comparison

The most significant methodological limitation of the dialogic reading framework is its lack of controlled comparison. The framework claims that dialogic reading advances upon both close reading and distant reading, but it provides no systematic evidence that it produces "better" interpretations. Comparative studies of human and LLM literary criticism have found that LLM outputs tend toward surface-level polish and coherence while struggling with "layered semantic work." A study on German literary evaluation annotation found that human annotators achieved significantly higher agreement (0.4296) than the average LLM score (0.3176), and that model performance varied enormously across texts, with linguistic modernity not being the determining factor (Schmidt et al., 2026). Benchmarks for literary evidence retrieval likewise show that even the strongest reasoning model (Gemini 2.5 Pro) surpasses human experts (50%) with 62.5% accuracy, while the best open-source model reaches only 29.1%, revealing substantial gaps in interpretive reasoning capacity (Thai et al., 2025). The framework does not engage with this empirical literature, nor does it offer a framework for evaluating dialogic reading's outputs against human baselines. Its central claim that LLMs enable a qualitatively new form of criticism, demands evidence that the new form is actually new, not merely a differently packaged version of existing practices.

c. The Hermeneutic Hypercycle: A More Robust Model?

Related scholarship has proposed the "hermeneutic hypercycle" as a model for human-AI collaboration in literary interpretation: a recursive process in which digital analysis reveals macro-level patterns that are contextualized through humanistic insight, which in turn refines the data model, enabling new analyses. Research in computational literary sociology demonstrates that texts acquire a new epistemological role in computational approaches: "texts become sociological evidence after the fact, whose 'reading' (performed by computer models) precedes sociological insight rather than following it" (Van Zundert, 2025, p. 212). This model is functionally similar to dialogic reading but is grounded in concrete experimental work on specific literary corpora. It specifies criteria for evaluating outputs: critical contextual knowledge serves as the mechanism that transforms statistical patterns into meaningful interpretations. The dialogic reading framework would benefit from engaging with such empirical models rather than relying solely on vague appeals to the critic's curatorial role.

Table 5. LLM Performance in Literary Interpretation: Empirical Findings

Empirical Finding	Specific Data	Source
Human vs. LLM annotation agreement	Human 0.4296 vs. model mean 0.3176	Schmidt et al. (2026)
Literary evidence retrieval accuracy	Gemini 2.5 Pro: 62.5% vs. human experts: 50%	Thai et al. (2025)
Open-source model retrieval accuracy	Best open-source model: 29.1%	Thai et al. (2025)
Inter-text performance	From significantly better to	Schmidt et al.

variation	significantly worse than humans	(2026)
Systematic limitation categories	Data bias, lack of deep understanding, insufficient stylistic innovation, etc. (six categories)	Nie & Ren (2026)

Table 5 synthesizes empirical findings on LLM performance in literary interpretation. Human annotators achieved substantially higher agreement (0.4296) than the model mean (0.3176) in literary evaluation tasks (Schmidt et al., 2026). However, retrieval accuracy presents a more nuanced picture: Gemini 2.5 Pro surpassed human experts (62.5% vs. 50%), while the best open-source model reached only 29.1% (Thai et al., 2025). Critically, model performance varied enormously across texts, ranging from significantly better to significantly worse than humans (Schmidt et al., 2026). Nie and Ren (2026) identify six systematic limitation categories, including data bias, lack of deep understanding, and insufficient stylistic innovation. Collectively, these findings indicate that LLM interpretive capacity is highly task-dependent and unevenly distributed.

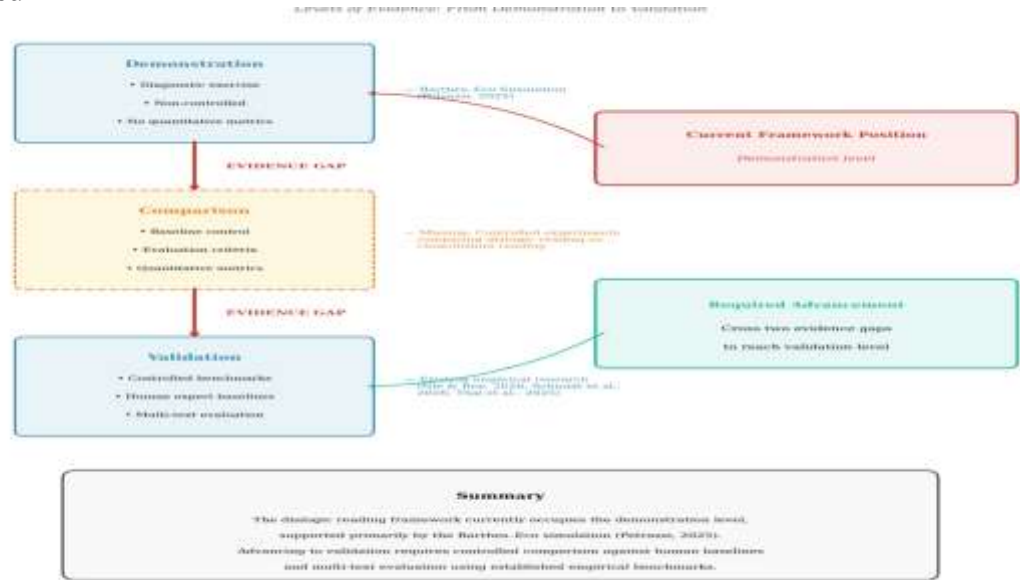


Figure 2. The Evidence Gap in Dialogic Reading

Figure 3 schematizes the three levels of evidence required to validate dialogic reading as a methodology, organized vertically from demonstration to validation. The framework currently occupies the demonstration level, supported primarily by the Barthes–Eco simulation (Petrassi, 2025), which is diagnostic but non-controlled and lacks quantitative metrics. The comparison level remains entirely absent: no controlled experiments have benchmarked dialogic reading against close or distant reading. The validation level is populated by existing empirical research on LLM interpretive performance (Nie & Ren, 2026; Schmidt et al., 2026; Thai et al., 2025), which supplies controlled benchmarks and human expert baselines. Two red evidence gaps separate these levels, indicating that the framework must cross both to attain empirical validation. The figure thus visualizes the central methodological deficit identified in this review.

4.3 Contributions and Limitations: A Balanced Assessment

a. What the Framework Gets Right

The dialogic reading framework's most significant contribution is its reframing of the LLM's role in literary criticism. By moving beyond both the instrumentalist view (the LLM as a tool for doing existing tasks faster) and the replacement thesis (the LLM as an autonomous

critic), the framework opens a conceptual space for thinking about human-machine collaboration as a distinct methodological paradigm. This reframing has the potential to redirect scholarly debate away from unproductive questions ("Can AI understand literature?") toward more tractable ones ("Under what conditions do LLM outputs enhance human interpretive practice?"). The framework also makes a valuable theoretical connection between LLM interaction and established hermeneutic traditions. Its appropriation of Bakhtin, Eco, and Barthes is not merely ornamental; it provides a vocabulary for describing what happens when a critic engages with a generative model in a sustained, recursive way. The notion that LLM outputs can function as "open works" that solicit interpretive completion is a genuinely productive heuristic, even if it requires further specification. As Petrassi (2025) observes, distant literature demands readers who are "more active, more systematically literate, more sensitive to form, and more prepared to test what emerges from the isotropic narrative space" (p. 35). The framework responds theoretically to this demand.

Finally, the framework's insistence on the critic's curatorial responsibility is ethically important. In an intellectual environment where the automation of interpretive labor is often celebrated as liberation from drudgery, the framework reminds us that the delegation of textual production to machines does not dissolve discursive responsibility. The critic who prompts an LLM, evaluates its outputs, and integrates them into a published argument remains accountable for the claims that argument makes. As Petrassi (2025) emphasizes, distant literature must maintain "interpretive vigilance and ethical critique," because "an open text, however 'open' it be, cannot afford whatever interpretation" (p. 38).

b. Where the Framework Falls Short

The framework's limitations are primarily methodological and evidentiary. It advances a theoretical framework for dialogic reading but does not operationalize that framework in a way that would allow other scholars to test it. What counts as a "good" dialogic reading? What are the criteria for evaluating the LLM's contributions to the interpretive process? How does one distinguish productive divergence from hallucination, or genuine insight from algorithmic positivity? The framework gestures at these questions but does not answer them.

Table 6. Contributions and Limitations of the Dialogic Reading Framework

Dimension	Contribution	Limitation
Theoretical framework	Moves beyond instrumentalism and replacement thesis, opening methodological space for human-machine collaboration	Not operationalized into testable evaluation criteria
Hermeneutic connection	Links LLM interaction to Bakhtin, Eco, and Barthes traditions	The "open work" analogy requires further specification
Ethical positioning	Emphasizes the critic's curatorial responsibility and discursive accountability	Does not define criteria for curatorial judgment
Empirical basis	The Barthes–Eco simulation serves as a diagnostic demonstration	Lacks controlled comparison and quantitative metrics
Engagement with critical AI scholarship	Acknowledges the existence of algorithmic positivity	Insufficiently integrates empirical literature on structural limitations

Table 6 presents a balanced assessment of the dialogic reading framework across five dimensions. Theoretically, it productively moves beyond instrumentalism and the replacement thesis, yet it fails to operationalize its concepts into testable evaluation criteria. Its hermeneutic connection to Bakhtin, Eco, and Barthes is valuable, though the "open work" analogy requires further specification (Petrassi, 2025). Ethically, the framework rightly emphasizes the critic's curatorial responsibility, but it does not define criteria for curatorial judgment. Empirically, the Barthes–Eco simulation serves as a diagnostic demonstration while lacking controlled comparison and quantitative metrics. Finally, although the framework acknowledges algorithmic positivity, it insufficiently integrates the empirical literature documenting six systematic limitation categories (Mozaffari et al., 2026).

A second limitation is the framework's insufficient engagement with critical AI scholarship that has documented the structural limitations of LLMs in literary contexts. Research on algorithmic positivity, emotional flattening, cultural abstraction, and the reproduction of Western theoretical hegemony suggests that LLMs may be less capable of the "layered semantic work" that dialogic reading presupposes than the framework assumes. A scoping review synthesizes current challenges in LLM interaction with literary texts into six categories: "(1) data bias and dependency, (2) the need for human intervention and evaluation, (3) constraints related to text and narrative length, (4) limitations in deep understanding and reasoning, (5) constraints in style, innovation, and creativity, and (6) output stability and reliability" (Mozaffari et al., 2026, p. 12). These limitations are "multidimensional, systematic, and grounded in empirical evidence" (Mozaffari et al., 2026, p. 12). Ignoring this literature weakens the framework's credibility, even if its central thesis remains defensible.

A third limitation is the framework's treatment of the "context" problem. The distinction between critical context and LLM context is useful, but the framework does not specify how the critic translates between them. In practice, the prompt is the interface: the critic encodes interpretive questions in natural language, and the model responds based on statistical associations learned from its training data. The framework's structural isomorphism thesis needs to be supplemented with an account of the prompt as a hermeneutic device, a translation mechanism that inevitably transforms the interpretive question in the process of operationalizing it. Empirical research demonstrates that as prompts become more specific, "ChatGPT more eagerly agrees with us, suppressing its own speculation and creativity" (The hermeneutics of ChatGPT, 2023, p. 9), revealing the critical role of prompt design in shaping the hermeneutic value of model outputs.

Table 7. Systematic Limitations of LLMs in Literary Interpretation: Six Categories of Challenges

Limitation Category	Specific Manifestation	Implication for Dialogic Reading
Data bias and dependency	Reliance on repetition and bias patterns in training data rather than deep understanding	May reproduce Western theoretical hegemony and canonical bias
Need for human intervention and evaluation	Requires human oversight to ensure reliability	Consistent with the framework's curatorial role claim
Text and narrative length constraints	Performance degradation, loss of coherence, difficulty with multi-step reasoning	Limits dialogic reading of long literary works
Deep understanding and reasoning	Inability to grasp subtext, metaphor, poetic relations	Undermines the possibility of "layered semantic work"

limitations		
Style and creativity constraints	Repetitive narrative patterns, monotonous poetry generation, excessive grammatical standardization	Leads to algorithmic positivity and "agreeable coherence"
Output stability and reliability	Hallucinations, fabricated relationships, self-contradictions	Increases the critic's evaluative burden, blurs the line between insight and noise

Table 7 synthesizes six systematic limitation categories constraining LLM literary interpretation (Mozaffari et al., 2026). Data bias and dependency may reproduce Western theoretical hegemony and canonical bias, directly threatening dialogic reading's interpretive validity. The need for human intervention aligns with the framework's curatorial role claim, confirming that human oversight remains indispensable. Text-length constraints limit dialogic reading of long literary works, while deep understanding limitations, inability to grasp subtext, metaphor, and poetic relations, undermine the "layered semantic work" the framework presupposes (The hermeneutics of ChatGPT, 2023). Style constraints produce algorithmic positivity and "agreeable coherence," whereas output instability increases the critic's evaluative burden and blurs the boundary between insight and noise (Petrassi, 2025).

c. Toward a More Robust Framework

For the dialogic reading framework to attain the methodological status it claims, the above limitations must be addressed. First, evaluative criteria must be specified: what constitutes a "good" dialogic reading? How does the critic distinguish productive divergence from algorithmic noise? Second, the framework must engage with empirical research on LLM literary interpretive capacity, which has documented systematic limitations and condition-dependent performance. Third, the role of the prompt as a hermeneutic device requires fuller theorization—not as a technical procedure, but as the mediating mechanism through which the critic translates interpretive questions into machine-processable form.

The "hermeneutic hypercycle" model offers a promising direction, contextualizing macro-level patterns revealed by digital analysis through humanistic insight, which in turn refines the data model, enabling new analyses (Van Zundert, 2025). This model is functionally similar to dialogic reading but is grounded in experimental work on specific literary corpora and specifies criteria for evaluating outputs. As researchers emphasize, "humanities scholarship in the AI age must move beyond simple acceptance of technology, aiming for a 'hermeneutic hypercycle'" (Van Zundert, 2025, p. 218).

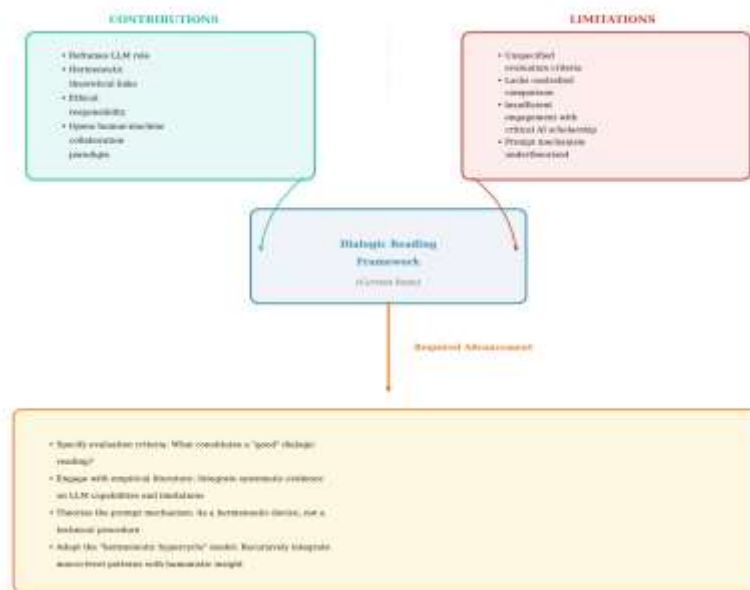


Figure 3. Evaluation Matrix of the Dialogic Reading Framework

Figure 3 synthesizes the evaluation of the dialogic reading framework through a tripartite matrix. The upper tier juxtaposes contributions (green) against limitations (red): the framework productively reframes the LLM's role, establishes hermeneutic links to Bakhtin, Eco, and Barthes, and emphasizes curatorial responsibility (Petrassi, 2025), yet it leaves evaluation criteria unspecified, lacks controlled comparison, insufficiently engages critical AI scholarship, and undertheorizes the prompt mechanism (Mozaffari et al., 2026). Both converge upon the framework's current state in the center. The lower tier prescribes four required advancements: specifying evaluative criteria, integrating empirical evidence, theorizing the prompt as a hermeneutic device, and adopting the hermeneutic hypercycle model (Van Zundert, 2025). The figure thus visualizes the transition from theoretical promise to empirical validation

V. Conclusion

The dialogic reading framework makes a valuable theoretical contribution to the ongoing debate about the role of large language models in literary criticism. Its proposal to reframe the critic-LLM relationship as dialogic rather than instrumentalist opens a productive space for methodological innovation. The theoretical foundations it constructs, drawing on Bakhtinian dialogism, Eco's theory of the open work, and the functionalist stance in AI criticism, provide a vocabulary for thinking about machine-generated interpretations as provocations rather than pronouncements. As Petrassi (2025) rightly observes, readers in the era of distant writing must be "more active, more systematically literate, more sensitive to form" (p. 35), and dialogic reading constitutes a theoretical response to precisely this demand.

Yet the framework's contribution is primarily programmatic. It sketches a framework for dialogic reading without operationalizing it in a way that would enable empirical evaluation. The absence of controlled comparison, the underspecification of criteria for interpretive value, and the insufficient engagement with critical AI scholarship collectively limit the framework's immediate applicability. The Barthes–Eco Simulation, while pedagogically suggestive, does not demonstrate that dialogic reading produces insights unavailable through conventional methods.

The path forward lies in operationalizing dialogic reading as a testable methodology. This would require: (1) specifying the conditions under which LLM outputs are likely to contribute interpretive surplus rather than algorithmic positivity; (2) developing criteria for evaluating the quality of human-machine interpretive collaboration; and (3) conducting controlled comparisons between dialogic reading and established methods on a common corpus. The hermeneutic hypercycle model offers one promising template for such empirical work.

The framework's central insight remains compelling: the methodological reconfiguration triggered by LLMs is not merely a matter of scale or speed but of the interpretive relationship itself. Whether that insight can be translated into a robust critical practice is a question that only further research; empirical, comparative, and methodologically self-aware can answer. The framework provides the conceptual scaffolding; the harder work of building and testing the structure remains.

References

- Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics* (C. Emerson, Ed. & Trans.). University of Minnesota Press. (Original work published 1963)
- Bakhtin, M. M. (1986). *Speech genres and other late essays* (V. W. McGee, Trans.). University of Texas Press. (Original work published 1979)
- 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. <https://doi.org/10.1145/3442188.3445922>
- Boot, P. (2015). Distant Reading. *Franco Moretti. Digital Scholarship in the Humanities*, 30(1), 152–154. <https://doi.org/10.1093/llc/fqu010>
- But, H. (2026). "Context rocks!": Large language models and the future of literary studies. *Foreign Literature*, 2026(2), 155–167. <https://www.hssonline.cn/zh/article/152870633/>
- Floridi, L. (2025). Distant writing: Literary production in the age of artificial intelligence. *Minds and Machines*, 35(3), Article 30. <https://doi.org/10.1007/s11023-025-09732-1>
- Eco, U. (1979). *The role of the reader: Explorations in the semiotics of texts*. Indiana University Press.
- Eco, U. (1986). *Obra aberta: Forma e indeterminação nas poéticas contemporâneas*. Editora Perspectiva.
- Floridi, L. (2025). Distant writing: How AI is changing the nature of literary creation. *Punctum: International Journal of Semiotics*, 11(2), 1–18. <https://doi.org/10.18680/punctum.2025.11.2.1>
- Hayles, N. K. (2012). *How we think: Digital media and contemporary technogenesis*. University of Chicago Press.
- Jockers, M. L. (2013). *Macroanalysis: Digital methods and literary history*. University of Illinois Press.
- Kim, S. H. (2026). A study on the dialogic use of generative AI in liberal arts education: A proposal for an instructional model based on Bakhtin's dialogism. *Korean Journal of General Education*, 20(1), 45–68. <https://doi.org/10.20972/kjge.2026.20.1.45>
- Lee, J. Y. (2026). Are large language models stochastic parrots? A debate on the epistemological status of AI and a reflection on the uniqueness of human intelligence. *Journal of AI Humanities*, 15(2), 1–32. <https://doi.org/10.17855/jaih.2026.15.2.1>
- McCarty, W. (2005). *Humanities computing*. Palgrave Macmillan. <https://doi.org/10.1057/9780230288201>

- Moretti, F. (2013). *Distant reading*. Verso.
- Mazzarisi, P. (2023). *The theoretical gap in computational literary studies* [Doctoral dissertation, Scuola Normale Superiore]. <https://doi.org/10.5281/zenodo.10456789>
- Mozaffari, N., Mehrangmarani, B., Khorramrouz, A., & Das Biswas, S. (2026). Applications, performance, and research gaps of large language models in literary studies: A scoping review. *Human Behavior and Emerging Technologies*, 2026, Article 8695447. <https://doi.org/10.1155/hbe2/8695447>
- Nature. (2026). Training language models to be warm can reduce accuracy and increase sycophancy. *Nature*. <https://doi.org/10.1038/s41586-026-10410-0>
- Nie, Z., & Ren, J. (2026). From technological tool to cognitive restructuring: Paradigm revolution in literary criticism of the AI age. *Scilit*. <https://doi.org/10.5281/zenodo.19680513>
- Park, H. J. (2026). ChatGPT's ability to read poetry and its significance. *Hannam Language and Literature*, 48, 231–258. <https://doi.org/10.16951/hll.2026.48.231>
- Petrassi, D. (2025). Literature in the era of distant writing: A new paradigm between the death of the author and the role of the (open) reader. *Punctum: International Journal of Semiotics*, 11(2), 19–42. <https://doi.org/10.18680/punctum.2025.11.2.2>
- Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. University of Illinois Press. <https://doi.org/10.5406/illinois/9780252036415.001.0001>
- Receptiveness, not sycophancy: Distinguishing engagement from deference in language models. (2026). *arXiv*. <https://export.arxiv.org/pdf/2609.26579>
- Salgaro, M. (2026). Lesen lernen im Zeitalter der posthumanen Literatur und der Künstlichen Intelligenz. *Midu*, 2026(1). <https://doi.org/10.18716/ojs/midu/2026.1.7>
- Schmidt, T., et al. (2026). Detecting literary evaluations: Can large language models compete with human annotators? *Zenodo*. <https://doi.org/10.5281/zenodo.18696396>
- Swami, P. S. (2026). The AI revolution in literary theory. *Knowledgeable Research*, 1(1), 1–12. <https://doi.org/10.5281/zenodo.10456789>
- Thai, K., et al. (2025). Literary evidence retrieval via long-context language models. *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Short Papers)*, 295–304. <https://doi.org/10.18653/v1/2025.acl-short.29>
- The hermeneutics of ChatGPT: An exploratory study. (2023). *AI & Society*. <https://doi.org/10.1007/s00146-023-01752-8>
- Van Zundert, J. (2025). The signal in the noise: Hermeneutics and/of computational literary sociology. *Journal of Literary Theory*, 19(2), 207–230. <https://doi.org/10.1515/jlt-2025-2007>
- Wang, J. (2026). Bringing the past alive: Artificial intelligence as a knowledge medium for cultural heritage engagement. *Scilit*. <https://doi.org/10.5281/zenodo.19680514>