Beyond Fiction.

Simulation and Narrative in AI-Generated Stories about Polish University Teachers

DOI: 10.14746/fp.2026.43.3

This is an open access article distributed under the terms of the CC BY-NC-ND 4.0 license.

Bartosz Lutostański

ORCID: 0000-0001-6345-3275

1. Introduction

The rise of large language models (LLMs) capable of generating coherent, stylistically consistent, and emotionally resonant narratives poses a fundamental challenge to literary theory. This paper, a contribution to the *Forum of Poetics* special issue, argues that AI-generated fiction requires a reconceptualisation of fictionality. Canonical frameworks, particularly those predicated on communicative intention and human semiotic agency, falter when applied to texts produced by generative algorithms. Such outputs mimic the surface logics of fiction but lack an intentional subject, a communicative ethos, or a model of reception.

Generative systems—Claude, ChatGPT, Gemini—have rapidly begun to infiltrate domains once reserved for literature proper: short stories[1], novels[2], poems[3], or mixed-media projects[4]. Yet, despite increasing scholarly and institutional engagement (Hannes Bajohr, Nick Montfort, Jill Walker Rettberg, AI Stories in Bergen, or Digital Humanities Centre at the Institute of Literary Research of the Polish Academy of Sciences), their ontological status remains under-theorised and hovers between imitation and invention, algorithm and art.

Within this context, fictionality theory—particularly in its narratological-rhetorical articulation—is strained. While flexible across genres and textual modes, it has not adequately accounted for algorithm-driven textual production. Fictionality, as a category rooted in sender intention, rhetorical signalling, and genre-bound reader inference, cannot accommodate the machinic text, which resists precisely these coordinates. What is needed is not a definitional adjustment but a paradigmatic shift. This paper proposes to make that shift by employing Jean Baudrillard's theory of simulation. His dismantling of the signifier-signified dyad and diagnosis of hyperreality provide a more apt framework for analysing LLM-generated narratives. Where fictionality theory presumes a communicative pact, simulation theory identifies a structural absence of reference, intention, and origin.

Two theoretical paradigms structure the present intervention: the rhetorical theory of fictionality and the Baudrillardian theory of simulation. The former emerged as a corrective to semantic, pragmatic and phenomenological approaches to fiction, positing fictionality as a discursive act grounded in communicative intention, interpretive negotiation, human-specific, creative imagination, and inherent awareness of the fictive/nonfictive distinction. Simulation theory, by contrast, represents a post-semiotic rupture: the implosion of representation, the procession of simulacra, and the production of hyperreality. This theory proves more productive for addressing AI-generated narratives, which are not authored in any conventional sense but rather emerge from algorithmic recombination within massive corpora. Though these frameworks appear incommensurable—the former rooted in communicative ethics, the latter in semiotic collapse in the realm of media and politics[5]—their juxtaposition animates this paper's central analytic tension.

This article is conceptual and case-study driven. It combines a critical-theoretical exposition of a corpus of ten short AI-generated narratives produced by Claude AI. The prompt—"Produce 10 different versions of a story about a Polish university teacher"—was deliberately designed to generate variation within thematic and stylistic constraints, thereby surfacing the LLM's capacities for narrative simulation. One story, "The Reluctant Revolutionary", is used below to exemplify my line of argumentation, and is analysed both in terms of surface narrativity and ontological unsettlement. Interpretively, the article moves beyond hermeneutics of

[1] E.g., Zilles, Li. Machine, Unlearning. Denver: Counterpath Press, 2018.

[2] Qudan, Rie. Sympathy Tower Tokyo. New York: Summit Books, 2025.

[3] Johnston, David Jhave. ReRites. 2017-2018. https://glia.ca/rerites/rerites_read_txt.html.

[4] Rettberg, Scott. Republicans in Love. 2022. Digital exhibition. https://eloconference2023exhibitions.wordpress.com/exhibition2_resistance/republicans-in-love/.

[5] Agnieszka Ziętek, Jean Baudrillard wobec współczesności. Polityka, media, społeczeństwo (Kraków: Universitas, 2013), 19-41.

meaning to engage a hermeneutics of generation—foregrounding algorithmic and statistical patterning as central analytical concerns.

The central proposition of this paper is not a claim to be defended but a problem to be unfolded: how can we theorise narrative texts with formal markers of fictionality but without the ontological preconditions for fiction? Rather than forcing AI-generated fiction into existing ontological categories, I argue for recognising it as a fourth-order phenomenon: simulated narrative text operating according to distinct epistemic protocols. This recognition requires a double displacement: first, of fictionality theory as a sufficient framework; second, of the assumption that computational narrative inherently entails communicative intent, anthropogenic imagination, or rhetorical signalling.

This position is not without risk. Simulation theory may tempt over-interpretation, attributing aesthetic weight to machinic output. Conversely, humanist frameworks may be inadvertently reimposed on posthuman artefacts. In reading these texts, the scholar cannot help but animate them with interpretive desire, thereby simulating meaning where none is encoded. The argument advanced here is thus necessarily performative: it interprets in order to displace interpretation. It reads to unread . This aporia is not a methodological failure. It is the condition of theorising narrative in the age of AI.

2. LLMs - From Fiction to Simulation

Large Language Models (LLMs) are sophisticated mathematical constructs that operate through probabilistic computation and large-scale pattern recognition. Built upon neural architectures, they transform linguistic input into numerical vectors, enabling contextual prediction through iterative optimisation[6]. Particularly influential has been the transformer architecture introduced by Google in 2017[7], which processes entire text sequences in parallel and weighs internal relations via the attention mechanism[8]. In simplified terms, each token in a sentence is assigned computational markers ("queries" and "keys"), compared pairwise, and evaluated for contextual relevance. Predictions are then generated via a softmax function that calculates the most statistically probable continuation[9]. Because output is probabilistically generated rather than rule-based, identical prompts can produce an infinite array of variants. The model's outputs are shaped by invisible parameters: how much randomness is allowed (so-called temperature), how words are selected (sampling), and which computational snapshot (model checkpoint) is accessed during generation. These factors ensure that LLM-

[6] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever, "Language Models Are Unsupervised Multitask Learners" (San Francisco: OpenAI, 2019), https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

[7] Ashish Vaswani et al., "Attention Is All You Need," Advances in Neural Information Processing Systems 30 (2017), https://arxiv.org/abs/1706.03762.

[8] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," paper presented at the International Conference on Learning Representations (ICLR), 2015, https://arxiv.org/abs/1409.0473.

[9] Jason Wei et al., "Emergent Abilities of Large Language Models," Transactions on Machine Learning Research, 2022, https://arxiv.org/abs/2206.07682.

generated texts lack stability and reproducibility; each response is contingent, ephemeral, and fundamentally non-repeatable.

LLMs' computational architecture prompted scholars to interrogate their limitations. A critique by Emily M. Bender et al. describes LLMs as "stochastic parrots": systems that "stitch together sequences of linguistic forms [...] without any reference to meaning"[10]. Generated texts lack communicative intent, world models, or audience modelling. Instead, they produce outputs that simulate linguistic coherence without cognition or reference[11]. In other words, ChatGPT or Claude neither think nor understand; instead, they simulate the properties of language so well that the illusion of intelligence arises. In this, they embody not communication, cognition and volition but algorithmic inference at scale. This critique recalls Ada Lovelace's early insight[12] that mechanical computation cannot originate meaning but only reproduce patterns: "It can *follow* analysis; but it has no power of *anticipating* any analytical relations or truths"[13].

This epistemological limit underpins what Halpern calls the "misunderstanding" of AI comprehension—a mistake rooted in conflating syntax with semantics. Likewise, Simone Natale[14] characterises belief in AI agency as a form of self-deception and Bender et al.—delusion. LLMs outperform humans in certain tasks yet remain devoid of consciousness, autonomy, or intentionality. John Searle's Chinese Room argument remains prescient: symbol manipulation is not understanding[15]. As Fletcher puts it, "*no* computer AI has *ever* learned to write a story [...]; it has emitted a syntactic mix-and-match that human brains [...] have re-imagined as a story"[16].

2.1 A short genealogy of fictionality theory

David Gorman identifies three dominant strands of fictionality. Semantic approaches locate fictionality in content—truth conditions, reference, and imaginary entities. Pragmatic models treat fictionality as a function of social and communicative practices. A third tradition identifies formal signals of fictionality—free indirect discourse, omniscient narration, and paratextual cues[17]. Gorman concludes that each approach offers productive tools but fails to overcome a fundamental asymmetry in narratological theory: fictionality is extensively

[10]Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" in Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21) (New York: Association for Computing Machinery, 2021), 616, https://doi.org/10.1145/3442188.3445922.

[11]Bender et al., "Stochastic Parrots," 617.

[12]Quoted in Mark Halpern, "No Ghost in the Machine: Artificial Intelligence Isn't as Intelligent as You Think," The American Scholar, March 2, 2020, https://theamericanscholar.org/no-ghost-in-the-machine/.

[13]Halper, "No Ghost", n.p.

[14]Simone Natale, Deceitful Media: Artificial Intelligence and Social Life after the Turing Test (Oxford: Oxford University Press, 2021).

[15]See also Alan Turing, "Computing Machinery and Intelligence," Mind 59, no. 236 (1950): 433-460.

[16]Angus Fletcher, "Why Computer AI Will Never Do What We Imagine It Can," Narrative 30, no. 1 (January 2022): 126.

[17]David Gorman, "Theories of Fiction," in The Routledge Encyclopaedia of Narrative Theory, ed. David Herman, Manfred Jahn, and Marie-Laure Ryan (Oxon: Routledge, 2005), 164-167.

theorised, while factuality remains under-theorised: "[t]he restricting factor appears to be the poorly understood nature of factual discourse."[18] Henrik Skov Nielsen, James Phelan and Richard Walsh, in a 2015 issue of *Narrative*, responded to that "restricting factor" by proposing a rhetorical approach to fictionality. Expanding on their prior research,[19] they reconceptualise fictionality as a communicative act[20]. Their model rests on several key premises:

- Fictionality is rooted in a specifically human capacity to imagine alternatives to the actual;
- It entails a communicative intent by a sender to signal fictionality;
- It requires an interpretive assumption by the receiver that fictionality is intended;
- No formal textual technique is either necessary or sufficient to identify fictionality;
- Fictionality modifies the ethos of the sender and the logos of the message.

Nielsen, Phelan, and Walsh stress that fictionality emerges in context and cannot be reduced to formal textual features. Rather, it presupposes, first of all, a communicative agent capable of intentional fictive signalling—what they call "an attitude toward the communicated information that is different from attitudes toward nonfictive discourse"[21]. The agent's competences are founded in the ability to "understand the real and the actual in terms of the possible and the imagined" (a feature "specific to human beings" in their view), and an intention "to speak fictively, non-fictively , or to blur the line between the fictive and the non-fictive status of her discourse"[22].

In addition, the model implies a sort of contract between sender and receiver, where fictionality is inferred from contextual and paratextual cues—genre labels, metanarrative markers, or discursive framing—as well as from the reader's assumption of the sender's imaginative volition. The contract stipulates the communication of a fictional text that is originally imagined and uniquely creative. The intentional communication of one human to another, along with flexible means of this process entailing a variety of genres, conventions and styles, is of paramount importance because it originates specific interpretive assumptions on the part of the readers and their approach to the ethos and logos of the fiction.

However, the fictionality theory, as formulated by Nielsen et al., breaks down in the case of digital fiction generated by means of artificial intelligence tools, such as large language

[18]Gorman, "Theories of Fiction," 167.

[19]James Phelan, Living to Tell about It: A Rhetoric and Ethics of Character Narration (Ithaca, NY: Cornell University Press, 2005); Richard Walsh, The Rhetoric of Fictionality: Narrative Theory and the Idea of Fiction (Columbus: Ohio State University Press, 2007); Henrik Skov Nielsen, "Natural Authors, Unnatural Narration," in Postclassical Narratology: Approaches and Analyses, ed. Jan Alber and Monika Fludernik (Columbus: Ohio State University Press, 2010), 285.

[20]Due to space limitations I am unable to engage with (many) critical commentaries to Nielsen et al.'s model of fictionality, in particular, Paul Dawson's ("Ten Theses Against Fictionality," Narrative 23, no. 1 (2015): 74-100.).

[21]Nielsen, Phelan, and Walsh, "Ten Theses About Fictionality," 67.

[22]Nielsen, Phelan, and Walsh, "Ten Theses About Fictionality," 64.

models. Narratives produced by ChatGPT or Claude are not rooted in intentional creativity and semantic and referential anchoring. A case in point, the role of a communicative agent or authorial instance becomes radically destabilised, or even undermined—who is the sender?[23] In Nielsen et al.'s articulation, it is a human being intentionally producing a text and consciously signalling its ontological status (e.g. distinguishing and imagining counterfactuals and unreal scenarios). In the case of AI-generated text, none of these properties is present. We deal with an algorithm-driven machine, a probabilistic system optimising next-token prediction, whose technical setup and stochastic properties lead to the collapse of the communicative contract fictionality theory prescribes. An AI agent, built on massive-scale parallel text processing and mathematical transformations of language into vectors, exhibits neither communicative intentionality nor the capacity to differentiate between the real and the non-actual. Even if the human prompter initiates the generative act, prompting is not equivalent to authorial communication. The prompt does not constitute a speech act directed toward an audience with imaginative intent. Therefore, the ontological status of the narrative discourse cannot be determined or even inferred from the output.

Consequently, LLM-generated texts resist categorisation within the rhetorical model. In my view, they are better understood as simulations in Baudrillard's sense: semiotic artefacts that circulate signs of fictionality without anchoring them in anthropogenic, communicative agency and intentionality. They reproduce the appearance of fiction—narrative form, affective arcs, genre conventions—without generating fictionality as a human-directed and referential act. LLMs simulate a human sender's capacity to imagine the non-actual and to signal this imaginative act to a human receiver without possessing the intentionality or ontological grounding such a framework presumes. In Baudrillardian terms, these texts constitute fourth-order simulations: probabilistic recombinations of signs drawn from massive corpora, devoid of "reference or circumference."[24] They do not signify the absence of reality; they mask the absence of fiction itself.

2.2 Simulation theory: Baudrillard's challenge to fiction

The term simulation has long occupied an ambivalent position in both scientific and cultural discourses.[25] Derived from the Latin *simulatio*, it originally denoted imitation—often with deceptive or representational intent. In early modern science, simulation became a modelling technique used to represent, approximate, or predict real-world phenomena under experimental constraints. This instrumental function deepened with the rise of cybernetics and systems theory, where simulation came to signify the algorithmic replication of system dynamics and the predictive emulation of behaviour within closed computational environments. Yet beyond this technical lineage, simulation is also a highly charged and polysemous concept—especially within game studies, where its epistemological, representational, and ludic dimensions have

[23]See Hannes Bajohr (2025) and his proposal to reconceptualise the categories of author, reader and text(uality).

[24]Jean Baudrillard, Simulations, trans. Paul Frost, Paul Patton, and Philip Beitchman (New York: Columbia University Press, 1983), 83.

[25]Ziętek, Baudrillard.

been variously theorised[26]. It is in the work of Jean Baudrillard that simulation undergoes its further radical rearticulation.

Baudrillard defines simulation as "the generation by models of a real without origin or reality: a hyperreal"[27] operational as "the reigning scheme of the current phase that is controlled by the code."[28] This process unfolds in four stages: first, the image reflects a profound reality; second, it masks and perverts it; third, it masks the absence of a profound reality; and finally, "it bears no relation to any reality whatever: it is its own pure simulacrum."[29] It is at this final stage—the fourth order—that LLM-generated stories operate: they neither point to the real nor conceal its absence; they 'merely' generate meaning effects.

Baudrillard contrasts simulation with representation, the latter of which relies on the principle that a sign refers to a real, even if only through distortion or displacement. Simulation, by contrast, "envelops the whole edifice of representation itself as a simulacrum."[30] It signals a fundamental rupture: "We live in a world where there is more and more information, and less and less meaning".[31] In the hyperreal, signs no longer refer to real objects or imaginaries but circulate within an operational chain of equivalence, detached from ontological grounding.[32] "The real is produced from miniaturised cells, matrices, and memory banks [...] and it can be reproduced an indefinite number of times", argues Baudrillard.[33] The result is a system wherein all referents are collapsed, and what remains is a combinatorial play of models, their "aleatory commutation".[34]

In this hyperreal regime, simulation no longer functions as representation or even imitation—it becomes a strategy of deterrence. As Baudrillard asserts, simulation "threatens the difference between the 'true' and the 'false,' the 'real' and the 'imaginary.'"[35] It does not obscure the real; it precludes its possibility. "Deterrence," he continues, "is the conjunction of the real with its operational double, a metastable, programmatic, perfect descriptive machine which provides all the signs of the real and short-circuits all its vicissitudes."[36]

This collapse of the referential poses a direct challenge to fictionality theory, particularly in its narratological-rhetorical articulation by Nielsen et al.. Their model rests on the ethos of

[26]Gonzalo Frasca, "Simulation," in The Johns Hopkins Guide to Digital Media, ed. Marie-Laure Ryan, Lori Emerson, and Benjamin J. Robertson (Baltimore: Johns Hopkins University Press, 2014), 452-455; Ian Bogost, Persuasive Games: The Expressive Power of Video Games (Cambridge, MA: MIT Press, 2007); Espen Aarseth, Cybertext: Perspectives on Ergodic Literature (Baltimore: Johns Hopkins University Press, 1997); Bernard P. Zeigler, Herbert Praehofer, and Tag Gon Kim, Theory of Modeling and Simulation (London: Academic Press, 2000).

[27]Baudrillard, Simulations, 2.

[28]Baudrillard, Simulations, 83.

[29]Baudrillard, Simulations, 10–12.

[30]Baudrillard, Simulacra, 6.

[31]Baudrillard, Simulations, 97.

[32]Baudrillard, Simulations, 112.

[33]Baudrillard, Simulations, 3.

[34]Baudrillard, Simulations, 112.

[35]Baudrillard, Simulations, 5.

[36]Baudrillard, Simulations, 4.

the sender and the interpretive cooperation of the receiver; simulation renders both positions obsolete. In the hyperreal condition, there is no sender capable of intention and no referential truth to which the text might allude. Simulation thus neutralises fictionality's central tenets: intentionality, imagination, and contextually anchored discourse. The signs do not signal; they replicate. They do not reflect imagination; they algorithmically recombine patterns from statistical training data.

Instead, LLMs such as Claude and ChatGPT operate squarely within Baudrillardian simulation logic. LLMs do not compose texts via authorial volition but by executing next-token predictions based on training corpora. Their outputs are generative in the machinic sense—products of computational optimisation, not rhetorical signalling. The absence of a sender collapses the fictive contract: there is no agent imagining the non-actual and, in Bender et al.'s words, no communicative intent, no model of the world or no model of the reader.[37] These LLM systems, by producing plausible simulations of discourse without intention, reference, or imagination, epitomise Baudrillard's fourth-order simulacrum: signs that circulate in the void of the real, detached from origin and indifferent to truth. LLM-generated stories do not conceal their non-authorship or artificiality—they perform the rhetoric of fictionality while hollowing out its epistemic foundations. Their generative logic substitutes semantic presence with statistical approximation, simulating literary coherence as a mode of deterrence, not expression. In Baudrillard's words, "End of the theatre of representation, the space of signs, their conflict, their silence; only the black box of the code [...]".[38]

To treat LLM outputs as fictions in Nielsen et al.'s sense is to misread their ontological status.[39] Rather than expanding the category of fictionality, AI-generated texts force us to confront the exhaustion of that category. As narrative theorists, we must ask not what these texts mean, but what it means to generate meaning effects in the absence of meaning. Here, Baudrillard's theory does not merely complement fictionality theory—it displaces it. The age of digital simulacra inaugurates a new textual paradigm: not fiction, but simulation.

3. AI-Generated Stories as Test Cases

On 15 December 2024, as part of an exploratory experiment combining personal curiosity with professional inquiry, I used the Claude AI interface to generate ten short narrative texts in response to the following prompt: "Produce 10 different versions of a story about a Polish university teacher. Don't exceed 100 words for each story."

This generative exercise was conducted in the context of ongoing research into the ontology of AI-generated fiction and its implications for literary theory and narratology. Initially a form of tinkering with a new textual technology, the output later served as the conceptual basis for a paper proposal to the conference *Avatars, Assistants, Chatbots: New Fictional Characters in*

[37]Bender et al. "Stochastic Parrots," 616.

[38]Baudrillard, Simulations, 104.

[39]Halper, "No Ghost." See also Natele, Deceitful Media; Fletcher, "Why Computer AI" and "Why Computers Will Never Read (or Write) Literature: A Logical Proof and a Narrative," Narrative 29, no. 1 (January 2021): 1-28.

Contemporary Culture (From Literature to New Media), held in Warsaw, April 8–9, 2025.

The generated corpus comprises ten stories, each approximately 80–100 words in length, portraying fictionalised academic characters situated within Polish university settings. Despite the brevity of the texts, each narrative displays a structured scenario, a focalised perspective, and a recognisable trope of transformation, conflict, or revelation. The stories are highly condensed, yet formally coherent, and present genre-specific expectations of academic fiction, didactic allegory, and short-form character sketch.

A brief summary of each generated narrative is provided below; full texts are included in the appendix.

1. The Reluctant Revolutionary – A mathematics professor in 1981 must choose between career stability and joining the Solidarity protests.

2. The Late Night Scholar – A medievalist uncovers a paradigm-shifting archival document in the University of Warsaw library at midnight.

3. The Language Bridge – A Polish language lecturer discovers deeper meanings of survival through teaching a Ukrainian refugee.

4. The Digital Convert – An elderly historian reluctantly embraces VR technology during the COVID-19 pandemic.

5. The Poetry Guardian – A janitor silently memorises Polish poetry recited daily by a literature professor.

6. The Chemistry of Change – Chemistry is taught through metaphors of human relationships, reshaping students' engagement.

7. The Returning Son – A Harvard-returned physicist challenges provincial scepticism with an international research initiative.

8. The Garden Scholar – A neglected garden becomes a site of philosophical discourse and community renewal.

9. The Time Keeper – A professor's lifelong punctuality is broken only by the birth of his grandchild, earning heartfelt recognition.

10. The Voice – A linguistics lecturer conducts an entire class in sign language when voiceless, revealing a hidden family narrative.

Due to space limitations, I will conduct a close reading of only one story generated by Claude. "The Reluctant Revolutionary" exemplifies all major premises of my argument.

> In 1981, Professor Jan Kowalski faced a choice: join his students in Solidarity protests or maintain his position at the university. His specialty was mathematics, not politics, but as he watched his brightest students risk everything for freedom, he realized some equations couldn't be solved on a chalkboard.

At first glance, the story appears to possess the formal elements of an ordinary fictional text. It features a named protagonist (Professor Jan Kowalski), a precise historical setting (Poland, 1981), a moral dilemma (to support the Solidarity movement or preserve institutional status), and even a relevant and evocative metaphoric closure. These markers might encourage the reader to interpret the text as a miniature instance of literary fiction—what Nielsen et al. would term a rhetorically signalled fictive communication. However, as described in section 2.1 above, it must be borne in mind that LLMs are algorithmic systems that transform language into vectors, calculate token probabilities via softmax, and generate outputs through stochastic sampling rather than intention. Their transformer architecture uses attention mechanisms to weigh word relationships mathematically, simulating coherence without understanding. Such models "stitch together" linguistic forms without reference or meaning—a "stochastic parrot"—on the basis of probabilistic inference.

I believe "The Reluctant Revolutionary" exemplifies the AI model's *opus operandi* very well. The protagonist's name is so generically Polish that it functions as a placeholder in official forms—akin to 'John Smith' in English. Likewise, the 1980s Solidarity movement is one of the most well-known moments in Poland's modern history, widely recognised internationally and a constant presence in any political and cultural discussions on freedom, workers' rights, and human rights activism.[40] The story's narrative structure—an internal conflict between desire and obligation—is a stock plot device. From Antigone and Hamlet to Raskolnikov and Simba from *The Lion King*, internal dilemma is a well-worn trope. In narratological discourse, such structures have been theorised at length in Propp's *Morphology of the Folktale*, Brooks's *Reading for the Plot*, Ricoeur's *Time and Narrative*, and Greimas's actantial model. Finally, the metaphorical closure, "some equations couldn't be solved on a chalkboard," transpires as a clever combination of Kowalski's specialisation and the mundane details of his profession, with the significance of his moral choice. The metaphor, however, is as predictable as it is hackneyed. Its position serves to conclude the microstory and provide it with some sort of profound and universal significance. In fact, however, it reads like a gimmick, a trick or formulaic expression of the story's main theme; indeed, the metaphor is a recombinatory cliché, merging the professor's discipline with his moral quandary. It is syntactically tidy but conceptually hollow: a statistical echo of similar narrative closures found in the model's training data.

"The Reluctant Revolutionary" is not a fictional text in the rhetorical sense."Solidarity" and "Jan Kowalski" operate not as historical or character signifiers grounded in context but as floating signifiers probabilistically selected to match prompt-derived keywords. The story, to paraphrase Baudrillard, is a simulated text that "bears no relation to any reality

[40]The "Solidarity on Screen" in The International Cultural Programme of the Polish Presidency of the Council of the European Union 2025 is a good case in point.

whatsoever".[41] The story's markers of "Polishness" serve, paraphrasing Baudrillard again, to "reinject realness and referentiality" into a system where such referents no longer exist[42]. As such they are, respectively, a non-event and a non-name floating in "the orbital recurrence" of synthetic "combinatory models in hyperspace without atmosphere",[43] or more relevant in the context of my argumentation, nodes of probabilistic coherence within Claude's pattern recognition routines haphazardly stitched together of "linguistic forms" observed in its vast training data".[44] As Baudrillard puts it, simulation "short-circuits all the vicissitudes of the real" by substituting it with operational doubles. The result is a text that presents a hyperreal version of fiction: 'cleaner,' more archetypal, and stripped of the real-world contingencies of any historically grounded story.

Simulation theory thus provides a more accurate analytical framework than Nielsen at al.'s rhetorical fictionality. The latter depends on assumptions of intentionality, communicative contract, authorial volition, or cognition of the concepts of the fictional and non-fictional—none of which hold in the case of LLM-generated texts. Simulation theory, by contrast, understands these texts as products of a self-referential semiotic system in which signifiers circulate without anchoring referents. The result is not fictionality but the simulation of fictionality. The appearance of interiority, historical depth, and moral ambiguity is algorithmically fabricated through statistical patterning. What we confront is not literature in the humanist sense, but machinic textuality—a hyperreal production of narrative signs absent of ontological foundation. It does not narrate history; it imitates the form of historical narration. It does not imagine character; it replicates the shape of imagined subjectivity. As Baudrillard writes, "simulation is no longer that of a territory… it is the map that precedes the territory".[45] Analogously, LLM-generated narratives do not reflect the world; they pre-empt it, generating fictionality effects in advance of meaning. In the age of algorithmic narrative, we no longer encounter fiction as human expression—but its operational double— automated and recombinant. Fiction is not written—it is simulated.

4. Conclusion: Toward a Theory of Simulated Fiction

This article has argued that AI-generated narratives, typified by Claude's outputs, resist classification as fiction in a narratological-rhetorical sense. These texts should be understood as simulated: structurally mimetic yet devoid of intentionality, reference, or agency.

Nielsen, Phelan, and Walsh offer a rhetorically rich model of fictionality, but it collapses when confronted with algorithmic generation—LLMs are not authors but statistical engines; they pattern linguistic artefacts, not ideas; their output simulates fictionality without fulfilling its preconditions. By contrast, Baudrillard's framework better captures the collapse of

[41]Baudrillard, Simulations, 11.
[42]Baudrillard, Simulations, 38–43.
[43]Baudrillard, Simulations, 3–11.
[44]Bender et al, "Stochastic Parrots," 617.
[45]Baudrillard, Simulations, 2.

anthropocentric conditions that underpin the category of fiction. AI-generated texts are simulations not because they lack fictional traits, but because they operate orthogonally to fiction: neither confirming nor negating it, but rendering its conceptual infrastructure porous.

This reconceptualisation opens critical questions: Can simulation be theorised as a distinct narrative mode? What are the risks of interpreting generated texts as literature? How does authorship operate when no author exists in any traditional sense? These questions remain provisional, not because they are insignificant, but because the ontological instability of their object precludes definitive resolution.

This article, therefore, defers rather than concludes. It calls for a reconfiguration of literary-critical discourse in response to technological transformation. In this new textual ecology, fiction is no longer simply written—it is simulated. Our critical vocabulary must evolve accordingly: not restored, but renegotiated. The ontological status of AI-generated narrative is not to be declared but understood as an emergent phenomenon—negotiated across disciplines, acts of reading, and theoretical paradigms. We may no longer ask, "Is this fiction?" but rather, "What does it mean to simulate fiction now?"

References

Aarseth, Espen. Cybertext: Perspectives on Ergodic Literature. Baltimore: Johns Hopkins University Press, 1997.

Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural Machine Translation by Jointly Learning to Align and Translate." Paper presented at the International Conference on Learning Representations (ICLR), 2015. https://arxiv.org/abs/1409.0473.

Bajohr, Hannes. "Writing at a Distance: Notes on Authorship and AI." YouTube video, 1:29:38. Posted March 26, 2025. https://www.youtube.com/watch?v=pP66rSFjPYw.

Baudrillard, Jean. Simulacra and Simulation. Translated by Sheila Faria Glaser. Ann Arbor: University of Michigan Press, 1981.

Baudrillard, Jean. Simulations. Translated by Paul Frost, Paul Patton, and Philip Beitchman. New York: Columbia University Press, 1983.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), 610–623. New York: Association for Computing Machinery, 2021. https://doi.org/10.1145/3442188.3445922.

Bogost, Ian. Persuasive Games: The Expressive Power of Video Games. Cambridge, MA: MIT Press, 2007.

Chun, Jon, and Katherine Elkins. "What the Rise of AI Means for Narrative Studies: A Response to 'Why Computers Will Never Read (or Write) Literature' by Angus Fletcher." Narrative 30, no. 1 (January 2022): 104–113.

Dawson, Paul. "Ten Theses Against Fictionality." Narrative 23, no. 1 (2015): 74–100.

Fletcher, Angus. "Why Computer AI Will Never Do What We Imagine It Can." Narrative 30, no. 1 (January 2022): 114–137.

——— "Why Computers Will Never Read (or Write) Literature: A Logical Proof and a Narrative." Narrative 29, no. 1 (January 2021): 1–28.

Frasca, Gonzalo. "Simulation." In The Johns Hopkins Guide to Digital Media, edited by Marie-Laure Ryan, Lori Emerson, and Benjamin J. Robertson, 452–455. Baltimore: Johns Hopkins University Press, 2014.

Gorman, David. "Theories of Fiction." In The Routledge Encyclopaedia of Narrative Theory, edited by David Herman, Manfred Jahn, and Marie-Laure Ryan, 164–167. Oxon: Routledge, 2005.

Halpern, Mark. "No Ghost in the Machine: Artificial Intelligence Isn't as Intelligent as You Think." The American Scholar, March 2, 2020. https://theamericanscholar.org/no-ghost-in-the-machine/.

Natale, Simone. Deceitful Media: Artificial Intelligence and Social Life after the Turing Test. Oxford: Oxford University Press, 2021.

Nielsen, Henrik Skov. "Natural Authors, Unnatural Narration." In Postclassical Narratology: Approaches and Analyses, edited by Jan Alber and Monika Fludernik, 275–300. Columbus: Ohio State University Press, 2010.

Nielsen, Henrik Skov, James Phelan, and Richard Walsh. "Ten Theses About Fictionality." Narrative 23, no. 1 (2015): 61–73.

Phelan, James. Living to Tell about It: A Rhetoric and Ethics of Character Narration. Ithaca, NY: Cornell University Press, 2005.

Radford, Alec, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. Language Models Are Unsupervised Multitask Learners. San Francisco: OpenAI, 2019. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

Turing, Alan. "Computing Machinery and Intelligence." Mind 59, no. 236 (1950): 433–460. https://courses.cs.umbc.edu/471/papers/turing.pdf.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention Is All You Need." Advances in Neural Information Processing Systems 30 (2017). https://arxiv.org/abs/1706.03762.

Walsh, Richard. The Rhetoric of Fictionality: Narrative Theory and the Idea of Fiction. Columbus: Ohio State University Press, 2007.

Wei, Jason, Yi Tay, Rishi Bommasani, Dan Roth, and Percy Liang. "Emergent Abilities of Large Language Models." Transactions on Machine Learning Research, 2022. https://arxiv.org/abs/2206.07682.

Zeigler, Bernard P., Herbert Praehofer, and Tag Gon Kim. Theory of Modeling and Simulation. London: Academic Press, 2000.

Ziętek, Agnieszka. Jean Baudrillard wobec współczesności. Polityka, media, społeczeństwo. Kraków: Universitas, 2013.

KEYWORDS

narratology

SIMULATION

fictionality

Abstract:
This article explores the ontological status of AI-generated fiction through a comparative reading of fictionality theory (Gorman 2005; Nielsen, Phelan and Walsh 2015) and Baudrillard's simulation theory (1981, 1983). Analysing a corpus of short stories generated by Claude AI, it argues that these texts simulate fiction rather than produce it. The article proposes that traditional narratological categories fail to capture the epistemic rupture introduced by LLMs, and that a new hermeneutic of simulated fiction is required.

digital fiction

AI-generated literature

Claude AI

LARGE LANGUAGE MODELS

Note on the Author:
Bartosz Lutostański – b. 1986, PhD in literary studies, Assistant Professor at the Institute of English Studies, University of Warsaw. His research focuses on digital fiction, media theory, postdigital culture, and narratology. Author of, most recently, *Poetyka postcyfrowa wybranych powieści na urządzenia mobilne* (Teksty Drugie 2025), and chapters in *New Approaches to Transcodification* (2025) and *The Palgrave Handbook of Intermediality* (2023). Latest work explores AI-generated narrative and posthuman poetics (article forthcoming in 2026). |