International Journal of Social Sciences and Humanities Invention 12(05): 8598-8613, 2025
DOI: 10.18535/ijsshi/v12i05.04
ISSN: 2349-2031
https://valleyinternational.net/index.php/theijsshi

Distant Writing and The Epistemology of Authorship: On Creativity, Delegation, And Plagiarism in The Age Of AI

Tiegue Vieira Rodrigues

Department of Philosophy, Federal University of Santa Maria, Brazil.

Abstract:

This paper examines the epistemic and ethical dimensions of Luciano Floridi's (2025a) concept of "distant writing," a form of AI-assisted composition in which large language models (LLMs) are used to generate, refine, or structure literary and argumentative texts. Drawing on analytic epistemology and virtue theory, and further informed by Floridi's (2025b) thesis on AI as agency without intelligence, it argues that distant writing constitutes a distributed epistemic activity, wherein the human author retains epistemic agency while delegating generative labor to non-agential systems. Central to this analysis is a distinction between epistemic instrumentality and epistemic agency, which clarifies the role of LLMs as sophisticated tools rather than co-authors. The paper explores how procedural, evaluative, and semantic knowledge inform the author's control over AI-generated content and assesses the epistemic risks introduced by the plausibility and opacity of such outputs. In response to concerns about plagiarism, it proposes a normative reformulation of attribution standards, emphasizing transparency and epistemic conscientiousness over traditional notions of originality. Ultimately, the paper situates distant writing within an emerging epistemic ecology in which creativity, authorship, and intellectual responsibility are collaboratively negotiated between human and machine. It advocates for a taxonomy of AI involvement to guide ethical disclosure and preserve the integrity of authorship in an age of synthetic textual production.

Keywords: Epistemic Agency; Distant Writing; AI authorship; plagiarism; Virtue Epistemology; Luciano Floridi.

1. Introduction

In recent years, the convergence of artificial intelligence (AI) and creative practice has precipitated a significant reconfiguration of the epistemic foundations of authorship, challenging long-standing assumptions about the nature of creativity, intention, and intellectual labor.[1] One of the most compelling and theoretically rich articulations of this transformation is **Luciano Floridi's (2025a) concept of *distant writing***.[2] This term captures a paradigm shift in literary production, wherein the traditional role of the author—as the primary originator and craftsman of narrative—is transformed into that of a designer, orchestrator, or curator of machine-generated textual outputs. In this emerging model, human agents do not compose texts directly in the conventional sense; rather, they interact with generative AI systems—particularly large language models (LLMs)—by crafting precise prompts, refining responses, and iterating through feedback loops to guide the system toward desired narrative outcomes (Floridi, 2025a).[3]

This process does not reduce the author to a passive bystander. On the contrary, the human remains a central epistemic agent, but one whose agency is now exercised through the manipulation and modulation of an autonomous, albeit non-intentional, linguistic system (Floridi, 2025b). The act of prompting, in this context, becomes a form of meta-authorship: it entails the conceptual framing and strategic direction of an algorithmic process whose internal workings are largely opaque, yet whose outputs are sensitive to input in complex, often non-linear ways.[4]The author, then, shifts from writing *in* the text to writing *through* the machine, transforming the nature of the literary creative act from one of direct composition to one of indirect design (Floridi, 2025a).

Floridi's (2025a) framing of distant writing draws a deliberate and illuminating analogy to Franco Moretti's (2000) influential concept of distant reading, which marked a methodological departure from traditional, close literary analysis. Distant reading employed computational tools to analyze large corpora of texts, enabling scholars to identify macro-patterns, stylistic trends, and structural features across vast literary landscapes—patterns that would remain inaccessible through individual textual analysis. Importantly, Moretti's approach was interpretative and retrospective: it sought to understand and map the literary past by aggregating and analyzing extant texts.

[1] This reconfiguration echoes earlier shifts in authorship studies, such as those prompted by post-structuralism, but AI introduces a novel agent (or instrument) into the creative process. For a discussion on the "death of the author" in a different context, see Kreminski (2024).

[2] The concept of "distant writing" builds upon Floridi's extensive work on the philosophy of information and AI ethics, offering a specific lens through which to analyze AI's role in textual production. See Floridi (2023) for a comprehensive overview of AI ethics.

[3] The dynamic interaction between human prompter and LLM can be seen as a form of "generative midtended cognition," where human and AI co-construct meaning through iterative feedback. See Barandiaran & Pérez-Verdugo (2025).

[4] The opacity of LLMs, often referred to as the "black box" problem, poses significant challenges for accountability and understanding. See Lipton (2020) for a discussion of model interpretability.

By contrast, distant writing is generative and prospective (Floridi, 2025a). It does not aim to interpret or explain existing literary works but to create new ones by leveraging the probabilistic and statistical capacities of LLMs.[5] While distant reading reflects upon cultural artifacts that have already been authored, distant writing participates in the active construction of new cultural artifacts, often with aesthetic and semantic properties that surpass the anticipatory intentions of the human prompter. In this sense, distant writing is not merely a methodological innovation; it is a transformation in the ontology of authorship. It represents a future-facing practice concerned with producing rather than interpreting, designing rather than deciphering, and co-creating rather than composing.[6] This shift raises complex questions not only about creative agency and textual ownership, but also about the epistemic conditions under which authorship and originality are now constituted.

This conceptual shift—wherein human creativity is mediated through artificial systems—raises profound and urgent epistemological questions. At the core of this transformation lies the disruption of long-held assumptions about the relationship between knowing, creating, and authoring. If the creative act is now facilitated or even partially executed by generative systems—such as large language models (LLMs)—whose internal mechanisms are not fully transparent even to their developers,[7] what does it mean for a human agent to know a work, to create it, or to claim authorship? These are not merely semantic queries but foundational challenges to our epistemic frameworks. When an individual inputs a prompt into an AI system and receives a text in return, are they engaging in an epistemic act? If so, what kind of act is it—interpretive, creative, manipulative, inferential, or something entirely new?

To address these questions, we must turn to key constructs within analytic epistemology. The first is epistemic agency—the capacity of a subject to be an active knower, responsible for their beliefs and assertions (Zagzebski, 1996). Traditionally, epistemic agency is understood as the capacity to exercise rational control over one's cognitive processes: to deliberate, form justified beliefs, and generate knowledge through reliable methods. However, in AI-mediated creativity, much of the linguistic and conceptual processing is performed by a system devoid of intentional states or rational deliberation, a form of "agency without intelligence" (Floridi, 2025b). The human prompter no longer generates every word or idea directly, but rather steers a probabilistic engine of language. This complicates the attribution of epistemic agency: Is the prompter merely initiating a causal process, or are they genuinely responsible for the epistemic status of the final product? Does crafting a prompt and selecting outputs constitute a form of epistemic authorship, or merely a kind of editorial oversight?

The second concern relates to the distribution of cognitive labor—a concept central to social epistemology, particularly in the work of Alvin Goldman (1999; 2011). In traditional knowledge practices, cognitive tasks are often distributed among individuals with specialized roles—writers, editors, researchers, etc.—but these agents are typically intentional and epistemically accountable. In AI-assisted writing, the labor of generating sentences, developing plotlines, or refining prose is partially delegated to a non-human system. This raises the question of how cognitive responsibility is allocated when human and machine collaborate. Does the AI perform epistemic labor, or is it merely a tool that facilitates the human's epistemic project, as Floridi (2025b) suggests? If the latter, how should we describe the human's relationship to the machine's output? Are they its author, its supervisor, its co-creator?[8]

Finally, this epistemic entanglement intersects with the normative question of intellectual credit and testimonial justice, as explored by Miranda Fricker (2007). If the authorial process is no longer fully attributable to a single agent, how should we understand claims of ownership or originality? Traditional models of intellectual credit presuppose that authors deserve recognition in proportion to their epistemic contribution. Yet in distant writing, the boundaries of contribution are increasingly ambiguous. The AI generates prose that may be syntactically sophisticated and semantically rich, but it does so without understanding, intention, or creative purpose (Floridi, 2025b). Still, its outputs may be instrumental in shaping the final work. Thus, we face a challenge in determining what kind of credit is owed, and to whom. While the AI cannot be considered a moral or epistemic agent, its role complicates the legitimacy of claiming exclusive authorship, particularly in contexts where the human author's labor is minimal.

Moreover, this evolution in literary methodology intersects with urgent normative concerns about plagiarism. Traditional conceptions of plagiarism rely on clear distinctions between original and derivative work, and between the agent who creates and the agent who appropriates. Yet in the context of distant writing, such boundaries become blurred (Floridi, 2025a). If an AI model trained on massive textual datasets generates a passage that is then selected, edited, and published by a human author, who holds the epistemic and ethical claim to that content? To what extent can such collaboration be deemed original, and under what conditions might it be considered a form of misattribution (Roig, 2006)? Moreover, this evolution in literary methodology—where AI systems

[5] Italo Calvino, in works like *The Castle of Crossed Destinies* (1973) or his essay "Cybernetics and Ghosts" (1986), explored the combinatorial nature of literature and the idea of the "literature machine" long before modern LLMs, presaging some of the creative and philosophical questions raised by distant writing. See also Lima (2023a, 2023b) and Scolari (2023) for recent discussions on Calvino and computing/AI.

[6] Bernard Stiegler's work on "neganthropology" and the impact of an "automatic society" provides a critical framework for understanding the broader societal shifts implied by practices like distant writing, particularly concerning the future of human labor and knowledge (Stiegler, 2019).

[7] The planetary costs and power dynamics embedded in the development and deployment of large AI models are extensively detailed by Crawford (2021) in *Atlas of AI*. These material conditions underpin the seemingly immaterial outputs of distant writing.

[8] The notion of AI as a "co-creator" or "collaborator" is prevalent in popular discourse, but careful philosophical distinctions, such as Floridi's (2025b) "agency without intelligence," are crucial to avoid anthropomorphizing these systems.

function as co-creative agents—intersects with pressing normative concerns surrounding plagiarism, a concept foundational to academic and literary integrity. Traditional frameworks for understanding plagiarism rest upon relatively clear ontological and moral distinctions: between what is original and what is derivative, between the agent who creates and the agent who copies or appropriates. These distinctions underpin both the attribution of intellectual credit and the enforcement of ethical norms in academic, journalistic, and literary domains. However, in the context of distant writing, where human authors rely on generative AI systems to produce substantial portions of textual content, these once-stable categories become increasingly ambiguous, even unstable.

One primary difficulty lies in determining the origin of AI-generated content. Large language models, such as GPT-4 or Claude, are trained on vast corpora of human-authored text, often scraped from the internet without the explicit consent of original authors (see, e.g., Bommasani et al., 2021). Although these models do not reproduce specific passages verbatim under normal conditions, their generative processes are shaped by the statistical regularities of this training data.[9] Consequently, when an AI generates a passage that is later selected, edited, and published by a human user, it is not always clear whether the resulting text is genuinely novel or merely a paraphrased recombination of existing material. This complicates traditional plagiarism detection frameworks, which rely on identifying lexical or syntactic overlap with pre-existing texts. In many cases, AI outputs may exhibit a high degree of semantic dependency on their training data without crossing the threshold of verbatim reproduction, thereby falling into a gray area epistemologically and ethically.

This raises the further question: who holds the epistemic and ethical claim to the content thus produced? If the AI is not a moral or epistemic agent—it lacks understanding, intention, and consciousness (Floridi, 2025b)—it cannot be held accountable for its outputs, nor can it claim authorship in any meaningful sense. At the same time, the human user who prompted the AI may not have exercised sufficient creative control to justify full authorial credit, especially if their role was limited to selecting and minimally editing the machine-generated output. This ambiguity challenges the applicability of conventional standards for determining authorship and originality.

Furthermore, this issue intersects with the concept of misattribution, as discussed by Roig (2006), which involves the improper assignment of credit for intellectual or creative work. In the age of distant writing, misattribution can take subtle and complex forms—not only in cases of direct plagiarism, but also in instances where the human author implicitly claims full creative authorship of work that was largely machine-generated. If the human contributor presents AI-assisted writing as the product of their own unaided intellect, this can constitute a form of epistemic misrepresentation, even if no single source text has been copied. The core ethical question becomes: under what conditions does AI-assisted authorship require disclosure, and what constitutes fair recognition of the respective roles played by human and machine in the production of a text?

Compounding the problem is the fact that existing academic and publishing norms have yet to establish consistent or robust guidelines for addressing these questions. For instance, while some institutions have begun to issue policies on AI-assisted work (see, e.g., Nature, 2023), these often focus on procedural compliance (e.g., requiring disclosure) rather than engaging with the deeper normative question of what constitutes intellectual responsibility in a human–machine collaboration.[10] Until such frameworks are clearly articulated, the boundaries between original creation, acceptable reuse, and unethical appropriation will remain contested and fluid, particularly in creative domains where stylistic innovation and content recombination are both expected and encouraged.[11]

In this context, the practice of distant writing functions as both a methodological innovation and a normative provocation (Floridi, 2025a). It forces scholars, editors, and creators to revisit the foundations of attribution, originality, and ethical authorship in light of emerging epistemic agents—however non-traditional or non-conscious—that increasingly participate in the creative process.[12] Thus, the intersection of AI-generated writing and the ethics of plagiarism is not simply a matter of policy compliance, but one of rethinking the conceptual architecture that underlies our notions of literary and academic integrity.

This paper explores these questions through the lens of analytic epistemology, focusing on the concepts of epistemic agency, the division of epistemic labor, and the normative structure of plagiarism. It argues that distant writing, as conceptualized by Floridi (2025a), is not merely a technological innovation but an epistemic and ethical provocation that compels us to rethink the foundations of authorship, creativity, and intellectual responsibility. In doing so, it aims to provide a principled framework for assessing the epistemic status of AI-assisted literary production and for navigating the contested terrain of ownership and originality in the age of generative artificial intelligence.

2. The Epistemic Architecture of Distant Writing

[9] The training data for LLMs often reflects societal biases, leading to outputs that can perpetuate stereotypes or skewed representations. This is part of what Mihalcea et al. (2024) and Atari et al. (2023) refer to when discussing why AI is "WEIRD" (Western, Educated, Industrialized, Rich, Democratic).

[10] For a broader discussion on the need for governance frameworks for AI, see Watson, Mökander, & Floridi (2024).

[11] The tension between innovation and ethical boundaries in AI-assisted creation is further complicated by the potential for "model collapse," where AI models trained on synthetic data may degrade in quality over time. See Dohmatob et al. (2024), Gerstgrasser et al. (2024), and Kazdan et al. (2024).

[12] This revisitation of authorship is also happening in critical discussions within literary communities. See, for example, Nguyen (2023) "Literature Machines" in *Boston Review*.

Floridi's (2025a) conception of distant writing introduces a profound disruption to the classical model of authorship, which traditionally centers on the image of a solitary, expressive individual—the author-genius—as the exclusive source of creative and epistemic production. In this older paradigm, authorship entails not only the act of generating content but also the expression of a unique intentionality, perspective, and cognitive labor. Distant writing, by contrast, displaces the locus of creativity from a unitary human mind to a distributed system involving both human and machine components.[13] In this hybrid configuration, the creative process is no longer synonymous with the internal deliberations of a conscious agent; rather, it unfolds across a collaborative interaction in which the human user engineers prompts, curates outputs, and iteratively refines results generated by an artificial system. Understanding the epistemic structure of such a process requires unpacking where, how, and by whom knowledge claims are made, transformed, and justified (see Sparrow, 2007; Coeckelbergh, 2020).

To do this, we must distinguish between epistemic agency and epistemic instrumentality—a critical distinction in analytic epistemology, particularly within virtue epistemology and the epistemology of cognitive systems (Sosa, 2007; Greco, 2010; Carter & Palermos, 2016). Epistemic agency refers to the capacity of an entity to form, evaluate, and take responsibility for beliefs. It implies intentionality, rational responsiveness to reasons, and the ability to assess evidence. An epistemic agent is someone who not only holds beliefs but does so because they are justified, reasoned, or warranted. In contrast, epistemic instruments are tools or mechanisms that facilitate the generation or transmission of beliefs without themselves being epistemically responsible or reflective. They contribute to epistemic processes, but they do not possess or exercise epistemic virtues (Burge, 2010; Pritchard, 2010). Floridi (2025b) argues that AI systems function as a new form of artificial agency, but one that lacks genuine intelligence or understanding in the human sense, thus positioning them as sophisticated instruments rather than peers in epistemic endeavors.

In the context of distant writing, the large language model (LLM)—for instance, GPT-4 or similar systems—is epistemically instrumental (Floridi, 2025b). It operates without consciousness, intentionality, or rational justification. Its outputs are not the result of beliefs formed through inferential reasoning but rather the result of mathematical optimizations over large datasets (Buckner, 2019).[14] It does not *know* in any propositional sense; instead, it models the likelihood of word sequences based on statistical correlations in its training data (Marcus & Davis, 2019).[15]

To illustrate, consider a user prompting an LLM to "Write a philosophical dialogue in the style of Plato, discussing the ethics of artificial intelligence." The resulting text may plausibly mimic the structure of a Socratic dialogue, employ quasi-Platonic language, and articulate positions that seem coherent or even insightful.[16] However, the model itself has no understanding of Plato, the Socratic method, or the ethical challenges posed by AI. It does not comprehend the underlying philosophical content or the historical significance of its style (Floridi, 2025b). Rather, it outputs sequences that approximate what it has seen in similar contexts during training. The resemblance is algorithmic, not intentional; it is analogous to a mirror that reflects a face without understanding what a face is (Sparrow, 2007; Gunkel, 2018).[17]

This distinction becomes clearer through analogy. Consider a typewriter used by a novelist: the machine plays a vital role in producing the text, but no one would suggest that the typewriter is a co-author. Similarly, a spell-checker that suggests corrections may influence the final form of a sentence, yet it has no intentional grasp of grammar or meaning. The LLM is more sophisticated than these tools, but in epistemic terms, it remains an instrument (Buckner, 2019; Floridi, 2025b). Its complexity does not entail epistemic agency; it merely extends the scope and precision of the author's tools. However, the analogy reaches its limits in one key respect: LLMs can generate entire passages, arguments, or stories autonomously in response to minimal prompts, thereby shaping the final product to a much greater degree than traditional tools (Marcus & Davis, 2019). This raises the stakes for understanding the human evaluator's role in distant writing.

The epistemic significance of the AI's output is not self-contained; it becomes meaningful only when the human interprets, selects, and endorses certain results as coherent, relevant, or truthful (Floridi, 2025a). In this sense, the human remains the ultimate epistemic agent in the interaction, responsible for exercising judgment over the instrumental products of the system. This interaction may be likened to the relationship between a curator and a generative art installation. The installation, driven by algorithmic processes,

[13] This distribution of cognitive labor in human-AI systems can be analyzed through various philosophical lenses, including Dennett's (1991) concept of "real patterns" to describe the predictive utility of attributing certain cognitive states or abilities to complex systems, even if they lack genuine understanding.

[14] The computational intensity of training these models has significant environmental implications. See Crawford (2024) and Van Wynsberghe (2021) for discussions on the sustainability of AI. Hasselbach & Van Wynsberghe (2025) also address AI, power, and sustainability within legal and ethical frameworks.

[15] The debate over whether LLMs truly "understand" language or merely manipulate symbols echoes earlier philosophical discussions, such as those surrounding Wittgenstein's philosophy of language. See Wittgenstein (2009) *Philosophical Investigations* for his concept of "language-games" and meaning as use, which can be contrasted with the statistical pattern-matching of LLMs. Mollema (2024) explicitly links Wittgenstein's language user with Calvino's literature machine in the context of social AI.

[16] The capacity of LLMs to generate stylistically coherent philosophical text has been empirically demonstrated, for example, by Schwitzgebel, Schwitzgebel, & Strasser (2024) in their project to create an LLM of a philosopher.

[17] The analogy of the mirror, or the "simulacra" as Baudrillard (1994) might term it, highlights the representational yet potentially hollow nature of AI-generated text when devoid of genuine understanding or grounding.

continuously produces new visual forms. The curator, observing the outputs, selects those that best align with a particular vision or concept. While the system contributes materially to the final aesthetic product, the curator is the one who imposes meaning, coherence, and value (Elgammal et al., 2017).[18] Likewise, in distant writing, the human author's epistemic agency is exercised not through direct composition, but through the direction, interpretation, and responsibility for machine-mediated content.

Distant writing exemplifies a novel epistemic configuration: a form of distributed cognition in which the human remains the site of agency, but the process of content generation is partially outsourced to an instrument that is non-conscious yet generative (Floridi, 2025a; 2025b). Appreciating this distinction is crucial for developing an epistemology adequate to AI-mediated creative practices, where the roles of agent and instrument are increasingly entangled, and the normative implications for authorship, credit, and responsibility are correspondingly complex (Sparrow, 2007; Gunkel, 2018).

This raises a key epistemological question: Is prompting and curating the outputs of a generative AI system an epistemically generative act on the part of the human author? In other words, does the human's engagement with the system constitute a genuine form of knowledge production, or is it better understood as a merely facilitative or mechanical task? At first glance, prompting might seem analogous to issuing a set of instructions to a skilled assistant—where the human provides direction, and the assistant executes. In this view, the creative and epistemic labor lies primarily in the articulation of a goal or vision, while the AI (like the assistant) handles the mechanical execution. However, this analogy risks understating the epistemic subtlety and creativity of the interaction involved in distant writing (Greco, 2010; Floridi, 2025a).[19]

A more illuminating analogy can be found in the domain of collaborative music. Consider a composer working with a highly improvisational pianist. The composer might begin with a thematic motif—a few bars of melody or a harmonic progression—and then ask the pianist to explore variations on that theme. The pianist improvises, offering spontaneous iterations, modulations, and embellishments. Crucially, it is the composer who listens attentively, assesses what aligns with the broader musical vision, selects segments that resonate, modifies others to better fit the desired mood, and rejects those that diverge too far. The creative and epistemic act here lies not solely in the initial motifs or in the pianist's improvisations, but in the process of selection, transformation, and integration carried out by the composer. The final composition reflects a synthesis of contributions, guided by the composer's evaluative and creative judgment (Sosa, 2007; Carter & Palermos, 2016).

In this analogy, the pianist stands in for the LLM: a system capable of generating complex, stylistically rich outputs in response to flexible cues. But like the pianist, the LLM lacks the broader thematic vision, the awareness of long-form coherence, or the teleological sense of where the composition is "going" (Floridi, 2025b). The human, by contrast, is the bearer of intentionality, meaning, and evaluative criteria. She must determine whether a generated paragraph fits the tone of the piece, whether a character's voice remains consistent, whether a plot twist is believable, or whether a particular metaphor enriches or distracts from the central theme. These judgments are not reducible to rote instruction; they involve epistemic sensitivity, a capacity to track norms of truth, coherence, style, and genre (Burge, 2010; Hookway, 2003).

Seen in this light, the practice of prompting and curating LLM outputs in distant writing can be reinterpreted as a form of epistemic orchestration (Floridi, 2025a). The human author does not relinquish control but relocates it from the sentence-by-sentence level to the level of higher-order creative and epistemic strategy. This process is supported by what Clark and Chalmers (1998) famously term epistemic scaffolding, a phenomenon within the extended mind thesis. Epistemic scaffolding occurs when part of an individual's cognitive labor is offloaded onto an external resource—be it a notebook, calculator, or software system—not just for efficiency, but to augment or transform the range of cognitive possibilities available (Clark & Chalmers, 1998).[20]

In the context of distant writing, prompt engineering functions precisely in this way. A carefully crafted prompt can elicit narrative styles, thematic variations, or rhetorical structures that the human writer might not have consciously imagined but can recognize, evaluate, and reconfigure. For example, a prompt such as "Describe a city at the end of the world, in the voice of Virginia Woolf" mobilizes the model to draw on its internal representations of both apocalyptic imagery and Woolf's distinctive narrative style. The model may return with richly textured, stream-of-consciousness prose that invites human reflection and editorial intervention. The author may extract certain sentences, recontextualize others, and reject some entirely—much as a sculptor might chisel away at raw marble to reveal a coherent form embedded within (see Gunkel, 2018).

Importantly, this process is not epistemically passive. The human author must constantly make evaluative judgments: Is this metaphor apt? Does this sentence advance the narrative? Is the character's motivation intelligible? These are deeply epistemic questions, grounded in the author's background knowledge, narrative intentions, and aesthetic sensibility. Insofar as the final text embodies these judgments—these acts of selection, coherence-checking, and thematic shaping—it expresses not just information,

[18] The role of metaphor in understanding complex systems like AI is itself an important epistemic consideration. See Black (1955, 1977) for foundational work on metaphor, and Kompa (2021) for the epistemic role of metaphor in science.

[19] The interaction in distant writing can be conceptualized as a form of "thinging with thinging things," where human cognition is dynamically coupled with the generative capacities of AI tools. See Barandiaran & Pérez-Verdugo (2025).

[20] The extended mind thesis, while typically applied to simpler cognitive tools, finds new relevance in the context of LLMs, which significantly augment human cognitive and creative capacities. See Menary (2010) for further exploration of cognitive integration.

but authored knowledge (Greco, 2010). The human is not merely responding to machine suggestions but actively filtering them through a conceptual and normative lens, shaping raw output into meaningful discourse.

Indeed, the concept of meta-control helps capture the human author's role here. Meta-control refers to the ability to regulate and direct a process without micromanaging every subcomponent. Much like a film director does not act out every role, design every costume, or build every set—but coordinates all these contributions to realize a cohesive vision—the distant writer exercises epistemic agency by defining constraints, setting aesthetic goals, evaluating outputs, and revising the generated content in light of those standards (Floridi, 2025a; Sosa, 2007).

The human role in distant writing is not trivialized by the presence of the AI model. Rather, it is reframed. The human author remains the epistemic center of gravity, responsible for guiding, evaluating, and finalizing the work. The machine functions as a non-agential generative instrument—much like a sophisticated improvisational tool—whose outputs acquire epistemic and creative significance only through the author's interpretive and transformative labor (Floridi, 2025b). This reframing invites us to recognize distant writing not as a loss of authorship, but as a novel configuration of authorship: one in which epistemic agency is exercised at the level of orchestration, synthesis, and critical curation (Dreyfus, 1992; Friedman & Hendry, 2019).[21]

The epistemic function of curation in distant writing is especially significant and deserves sustained analysis. Without it, the process risks devolving into mere content sampling, where the role of the human is minimized to passive consumption. However, when executed with critical engagement, curation transforms into an act of epistemic judgment, where the human author actively determines the value, relevance, and coherence of machine-generated content in relation to broader narrative or argumentative aims. This evaluative role is neither superficial nor perfunctory—it requires the integration of substantive knowledge, narrative sensitivity, and aesthetic judgment (Baehr, 2011; Kvanvig, 2003).

For instance, consider a scenario in which an author prompts an LLM with a request such as: "Generate a conversation between a grieving father and an estranged daughter who reunite after ten years." The model may produce five paragraphs of dialogue, each syntactically correct and emotionally charged. However, the human author must assess: Do these paragraphs maintain internal emotional consistency? Is the pacing appropriate? Are the metaphors used in the dialogue thematically aligned with the story's emotional arc?[22] Does the daughter's characterization remain coherent with her previous appearances in the narrative?

Such judgments are cognitively complex and epistemically rich. They require the author to possess a deep understanding of narrative structure, character development, emotional plausibility, and stylistic nuance. The AI may generate text with surface-level fluency—grammatical correctness, plausible diction, and apparent continuity—but it is the human interpreter who must determine whether that fluency translates into narrative truth, coherence, or expressive meaning. This function closely mirrors that of a museum curator, who does not create the artworks themselves but selects, arranges, and frames them to provoke insight, elicit affective responses, or express a thematic vision (Elgin, 2017; Goodman, 1976). The curator's labor is epistemic in nature: she must be able to discern what is meaningful, what deserves prominence, and how individual pieces relate to a broader conceptual framework.

The curatorial model challenges classical views of authorship that associate it predominantly with expressive generation—crafting language directly from the author's mind onto the page. In distant writing, the author's primary creative intervention may not lie in lexical invention but in the design, selection, and recombination of outputs derived from a generative model (Floridi, 2025a). This shift in activity raises a profound epistemological tension: if the human role in distant writing is primarily curatorial, directive, and evaluative—rather than expressive in the traditional sense—should we reconceptualize the human as a designer or editor rather than a writer in the conventional sense? (Baehr, 2011; Nagel, 1986).

This question gains further depth when viewed through the lens of epistemic calibration, a term developed by thinkers like Pritchard (2016) to describe processes in which agents interact with epistemic tools—scientific instruments, computer models, or statistical algorithms—in order to ensure their outputs are accurate, reliable, or meaningful within a given epistemic context. In this model, the human does not simply receive information from the tool but engages in a recursive process of tuning and interpreting the tool's output. The agent forms hypotheses, runs the instrument, evaluates its results, adjusts the input or method, and repeats the process. The agent is, in effect, managing a feedback loop in which understanding and reliability are incrementally secured through critical engagement with a semi-autonomous system (Pritchard, 2016; Hempel, 1966).

Such a framework maps remarkably well onto the process of distant writing. The author begins with an initial prompt—a hypothesis about what kind of output might be produced or what stylistic or thematic direction is desired. The AI model returns an output, which the author then evaluates: Does it conform to the desired tone? Is it thematically aligned? Are there inconsistencies or clichés? Based on this evaluation, the prompt may be revised, constraints tightened, stylistic cues adjusted. The model is prompted again, and the process iterates. This recursive engagement mirrors practices found in scientific modeling, where knowledge is refined through repeated interaction with an instrument or system (Kuhn, 1962; Giere, 2006).

Here, the creative act is distributed across time and across system boundaries. It resembles scientific modeling more than romantic authorship. Just as a climate scientist iteratively refines a simulation to ensure that it captures key features of atmospheric behavior,

[21] This reframing also prompts reflection on the "human reader" and how their expectations and interpretive practices might change in an age of AI-generated or co-created texts. See Gibson (2023).

[22] The aptness of metaphor, as discussed by Black (1955, 1977), becomes a critical evaluative criterion for the human author when curating AI-generated text.

the distant writer refines generative conditions to align the model's output with their narrative vision. The result is not a straightforward expression of interior subjectivity but a structured interaction with a cognitive artifact, mediated by epistemic criteria such as relevance, plausibility, novelty, and coherence (Bloor, 1991; Longino, 1990).

It is worth noting, however, that the iterative prompt-and-refine loop at the heart of distant writing is not inherently epistemically virtuous. While it offers the potential for dynamic creativity and increased productivity, it also opens the door to subtle and pervasive epistemic hazards. A poorly designed prompt can easily elicit outputs from an LLM that are factually inaccurate, logically inconsistent, or conceptually muddled (Jakesch, Lederer, & Meyer, 2023). The author, especially one unfamiliar with the domain in question or overly reliant on the AI's surface plausibility, may fail to detect these defects. As a result, such outputs may be inadvertently incorporated into the final product, thereby compromising the epistemic integrity of the work (Friedman & Nissenbaum, 1996; Lipton, 2020).[23]

This problem is compounded by a unique feature of LLM-generated text: its stylistic fluency and rhetorical polish. Modern generative models are extraordinarily adept at mimicking the conventions of human-written language—correct grammar, natural cadence, and contextually appropriate word choice. But this very fluency can create an epistemic illusion: it makes unreliable content appear trustworthy, poorly reasoned claims seem coherent, and fabricated references appear authentic. In essence, these models can produce what might be called epistemic simulacra—outputs that simulate the form of knowledge without its substance (Baudrillard, 1994; Gunkel, 2020).[24]

An instructive analogy can be drawn to forgery in art. A forged painting may replicate the style and brushwork of a famous artist so convincingly that it deceives even trained viewers. Yet despite its surface resemblance, the forgery lacks the authentic historical and creative provenance that gives the original its value. Similarly, a paragraph generated by an LLM may resemble an insightful literary analysis or a well-reasoned philosophical argument, but upon closer inspection, it may lack grounding in accurate sources, fail to follow logical argumentation, or misconstrue key concepts (Jakesch, Lederer, & Meyer, 2023; Milan, 2024).

The risk is not hypothetical. Consider a case in which a writer working on a philosophical essay uses an LLM to generate a brief explanation of Gettier problems in epistemology. The model may produce a coherent-sounding paragraph that invents a non-existent counterexample or misattributes it to the wrong philosopher. If the human author lacks the background knowledge to verify the content, they may include it in their essay, thereby propagating false epistemic claims under the veneer of credible exposition (Grasso, 2023; Strzelecki, 2023).[25]

In this context, the epistemic loop of distant writing becomes more than a site of creative potential—it becomes a zone of cognitive vulnerability. Authors must navigate this terrain with a set of intellectual virtues, as described by epistemologists such as Zagzebski (1996) and further explored by Battaglia (2022). These include: *Vigilance*, the virtue of maintaining a watchful stance toward outputs that may be superficially plausible but epistemically weak. This involves actively questioning the reliability of AI-generated content and cross-verifying facts when necessary; *Critical self-reflection*, the capacity to interrogate one's own role in the generative process, including one's reliance on the model, susceptibility to confirmation bias, or inclination to accept convenient outputs without sufficient scrutiny; and *Epistemic humility*, the virtue of developing an awareness of the limitations of one's own knowledge and the epistemic limitations of the tools being used. This involves recognizing that no matter how sophisticated the AI appears, it does not "understand" its outputs and should not be treated as an epistemic peer (Danaher, 2022; Nyholm, 2023).[26]

These virtues are essential not only to maintain the coherence and truthfulness of AI-assisted writing, but also to safeguard against a creeping erosion of epistemic standards. When authors delegate large portions of creative or intellectual labor to LLMs without exercising these virtues, they risk producing texts that are not merely aesthetically hollow but epistemically defective (Gillespie, 2020; Susser, Roessler, & Nissenbaum, 2019).[27]

Thus, while distant writing can enhance creative exploration, it also demands a heightened ethical and epistemic awareness. It calls upon the author not just to guide the machine, but to critically evaluate and take responsibility for what emerges from that collaboration. The human author, in this paradigm, is not absolved of epistemic responsibility by the involvement of an AI. Rather, their responsibility is reframed: not to produce every word, but to ensure that the words selected reflect a standard of truthfulness, coherence, and intellectual integrity (Floridi, 2023; Johnson & Verdicchio, 2017).

Another instructive analogy comes from the domain of architecture, as explored by Bates, O'Connor, and Pennington (2024) in the context of AI art. Architects rarely construct buildings themselves; they design, model, test, and revise. The physical realization of

[23] Beyond factual inaccuracies, the reliance on LLMs without critical oversight can lead to a homogenization of thought or style, potentially stifling genuine human creativity if not carefully managed.

[24] Frankfurt's (2005) analysis of "bullshit" as discourse produced without concern for truth offers another lens through which to view LLM outputs that are plausible but ungrounded.

[25] The creation of "digital duplicates" or sophisticated AI models of specific individuals also raises complex ethical questions, related to but distinct from authorship, concerning identity, representation, and consent. See Danaher & Nyholm (2024).

[26] The call for epistemic humility is particularly salient given the tendency to over-attribute understanding or sentience to sophisticated AI systems.

[27] Stiegler (2019) warns of the "proletarianization" of mind that can occur when cognitive tasks are overly delegated to automated systems without corresponding human skill development or critical engagement.

the building—through contractors, materials, environmental constraints—is often unpredictable. Yet we do not hesitate to credit architects with authorship, so long as the structural and aesthetic unity of the design remains intelligibly attributable to them. Similarly, distant writing may allow for AI-mediated realization of narrative form, but the coherence, originality, and epistemic credibility of the work must ultimately trace back to the human agent's intentional design (Bates, O'Connor, & Pennington, 2024; Lawson, 2006).

In sum, the epistemic architecture of distant writing is layered and distributed. The human author remains the epistemic center of gravity, not by virtue of direct linguistic production, but by maintaining oversight, direction, and critical evaluation of outputs. The AI functions as a generative instrument—highly capable but epistemically inert (Floridi, 2025b). This new modality invites us to revise our understanding of authorship: not as an act of isolated creation, but as a higher-order epistemic practice encompassing orchestration, supervision, and conceptual integration. Floridi's (2025a) contribution, then, is not simply to introduce a novel literary technique, but to prompt a re-examination of the cognitive economy of writing itself in an age of synthetic agency (Hayles, 1999).[28]

3. Knowledge, Creativity, and the Division of Epistemic Labor

The emergence of distant writing invites a reconfiguration of the relationship between knowledge and creativity, particularly with regard to how epistemic labor is divided and attributed across human and artificial agents. Traditionally, creative authorship has been seen as an epistemically unified activity: the author is the originator of both the ideas and their expression. However, when writing becomes a composite process involving prompt engineering, model sampling, and iterative refinement, the intellectual labor becomes distributed. This raises two critical questions for epistemology: (1) What kinds of knowledge are involved in distant writing? and (2) How should epistemic credit be allocated within this hybrid system? (Vee, 2022; Thakur, Loe, & Lee, 2024).

Distant writing, as conceptualized by Floridi (2025a), engages multiple layers of epistemic activity that go beyond traditional notions of textual authorship. At least three distinct types of knowledge are operative in this AI-mediated process: procedural knowledge, evaluative knowledge, and semantic knowledge. Each plays a critical role in shaping the integrity and value of the resulting text, and their interplay underscores the complex epistemic demands placed on the human author in this hybrid mode of creativity (Battaglia, 2022; Johnson & Verdicchio, 2017).

Procedural knowledge—what Gilbert Ryle (1949) termed "knowing how"—refers here to the practical competence involved in using large language models (LLMs) effectively. This includes knowing how to construct prompts that elicit relevant, stylistically consistent, and thematically appropriate outputs, as well as understanding how to iteratively refine these prompts in response to generated content. This is not a trivial skill; it often involves trial and error, sensitivity to the model's training corpus and behavior, and an implicit grasp of its probabilistic tendencies. For instance, prompting a model to generate a dialogue in the style of Platonic dialectic, or to continue a narrative in a postmodern literary register, requires more than superficial familiarity with those forms—it demands the ability to encode structural and thematic cues into the prompt language itself. In this sense, effective prompting becomes a form of epistemic design: a strategic orchestration of input conditions to navigate a vast latent space of possible outputs (Bommasani et al., 2021; Hayles, 1999).

Evaluative knowledge—the capacity to assess the coherence, originality, and quality of the outputs—is equally indispensable. Unlike procedural knowledge, which governs the interaction with the model, evaluative knowledge determines what counts as success within that interaction. It involves aesthetic sensibility, critical reasoning, domain expertise, and often a form of intellectual taste. For example, a philosopher using an LLM to draft a passage on epistemic injustice must not only recognize whether the terminology is used correctly but also whether the argument presented aligns with current debates, avoids fallacies, and contributes meaningfully to the discourse. The model may produce text that is stylistically smooth or superficially accurate, yet miss the deeper conceptual coherence required in academic writing. Here, the author must function as an epistemic gatekeeper—distinguishing between mere surface plausibility and genuine argumentative integrity (Battaglia, 2022; Grasso, 2023).

Semantic knowledge—knowing that certain statements are true or false, or understanding the meaning and implications of what is being discussed—remains foundational. This includes background knowledge of the subject matter, conceptual frameworks, and the historical and cultural context in which certain ideas are situated. In traditional authorship, semantic knowledge informs the generation of content directly; in distant writing, it becomes the criterion by which generated content is evaluated and revised. This shift has profound implications. As Bender and Koller (2020) emphasize, LLMs do not possess semantic knowledge in any robust sense. They generate plausible continuations of text based on statistical regularities in their training data, not on a grounded understanding of meaning. This distinction is epistemically decisive: while the model can simulate informed discourse, the justification for any claim it makes can only be provided by a human who understands both the claim and its evidentiary basis (Strzelecki, 2023; Bender & Koller, 2020).

A helpful analogy can be drawn from scientific modeling. Consider a climate model that produces probabilistic forecasts based on vast datasets and sophisticated simulations. The model's output has instrumental value, but it lacks intrinsic epistemic authority. It

[28] Calvino's (1986) "Cybernetics and Ghosts" thoughtfully considers the relationship between the creative human spirit and the logical, combinatorial possibilities of "literary machines," anticipating many of the tensions explored in distant writing. His letters and later reflections also offer insights into his views on literature and the changing world (Calvino, 2013, 2023).

is the scientist who interprets those results, assesses their plausibility, compares them with empirical observations, and integrates them into theoretical or policy frameworks. Without this layer of human interpretation, the model's outputs remain raw and uninterpreted. Similarly, in distant writing, the LLM opens a space of textual possibilities, but it is the author who imposes conceptual structure, discerns narrative coherence, and assumes responsibility for the final product's epistemic legitimacy (Nyholm, 2023; Floridi, 2023).

The human author, therefore, is not simply a user of the tool but the bearer of epistemic responsibility. The LLM can be a powerful epistemic instrument—amplifying creativity, facilitating exploration, and enabling novel forms of expression—but it cannot autonomously generate knowledge in the analytic sense. It lacks belief, justification, and understanding—the core ingredients of knowledge as traditionally conceived in epistemology (Floridi, 2025b). What it provides are candidate expressions of thought that must be vetted, interpreted, and, if necessary, rejected or revised by an epistemic agent (Gillespie, 2020; Vee, 2022).

In this light, distant writing should not be viewed as a displacement of human creativity or cognition, but as a transformation of their structure and site (Floridi, 2025a). Knowledge generation in this paradigm is distributed but asymmetrical: the machine provides syntactic and stylistic variation, while the human retains epistemic authority and bears ethical responsibility. The novelty of distant writing lies not in automating authorship, but in reconfiguring the relationship between tool and thinker, between output and understanding, and ultimately, between knowledge and its expression (Creative Commons, 2023; Floridi, 2014).

Within the traditionally epistemological canon, creativity has not been foregrounded as a core component of epistemic evaluation. The dominant focus has been on belief, justification, and truth—on what it means to know, and under what conditions knowledge can be said to exist. Creativity, by contrast, has often been relegated to aesthetics, psychology, or the philosophy of mind (Boden, 2004). However, the emergence of distant writing as a literary methodology compels a reassessment of creativity's epistemic function. It invites us to consider whether acts of creative synthesis, recombination, and imaginative exploration might themselves be forms of epistemic labor—activities that contribute not just to the generation of aesthetic products, but to the advancement of understanding and the expansion of conceptual repertoires (Kvanvig, 2003; Elgammal et al., 2017).[29]

John Kvanvig (2003) has provided one of the most compelling arguments for this expanded view. He contends that the epistemic value of cognitive practices should not be reduced to the narrow metric of truth acquisition. Instead, practices that promote understanding, foster conceptual clarity, or generate fruitful questions may be epistemically valuable even when they do not immediately yield justified true beliefs. Distant writing exemplifies this claim. When an author collaborates with a large language model to generate unexpected narrative configurations, juxtapose divergent stylistic registers, or explore counterfactual scenarios, the result may not always be factually informative in the traditional sense. However, these outputs can generate novel perspectives, surface previously unarticulated tensions, or reframe familiar problems in new ways. Such cognitive outcomes—though not reducible to propositional knowledge—can play a crucial role in the cultivation of epistemic insight (Gaut, 2010; Milan, 2024).

Consider, for example, a novelist working on a philosophical novel about identity and memory. In a conventional writing process, the author might remain constrained by their own stylistic habits or conceptual biases. But by using a language model to generate speculative dialogues or alternative narrative arcs, the author gains access to a repertoire of textual possibilities that can provoke reflection, illuminate blind spots, or reveal thematic connections not previously envisaged. In this sense, creativity becomes an exploratory epistemic method—an engine for expanding the modal space of literary thought (Hayles, 1999).[30]

Importantly, this conception of creativity as epistemically generative challenges the longstanding assumption of epistemic individualism that underlies much of analytic philosophy. Under the individualist model, knowledge production is assumed to reside within the cognitive capacities of single agents—agents who reason, observe, and infer in relative isolation. Yet distant writing is irreducibly collaborative and systemic (Floridi, 2025a). The human author is not operating in a vacuum, but is embedded in a socio-technical system that includes not only LLMs, but also digital interfaces, prompt repositories, training datasets, and feedback mechanisms.[31] The epistemic trajectory of the writing process is thus co-shaped by technological affordances and design decisions that structure how information is accessed, generated, and curated (Battaglia, 2022; Vee, 2022).

This shift in perspective aligns with broader developments in social and virtue epistemology, particularly the growing recognition of the situated and distributed nature of knowledge. Scholars like Alvin Goldman (1999; 2011) and Miranda Fricker (2007) have emphasized the social embeddedness of epistemic practices, pointing to how testimonial exchange, institutional structures, and power dynamics influence who is recognized as a knower and what counts as credible knowledge. Distant writing, viewed through

[29]The exploration of creativity via AI also touches on themes found in Italo Calvino's *Invisible Cities* (1974), where Marco Polo generates descriptions of countless cities for Kublai Khan, blurring the lines between observation, imagination, and systematic generation.

[30] The process can be likened to a writer navigating a "garden of forking paths," a concept famously explored in literature (e.g., by Borges), but here the paths are partially generated by an algorithmic muse. Calvino's *If on a Winter's Night a Traveler* (1981) itself plays with multiple narrative beginnings and the reader's role in constructing meaning, which resonates with the interactive nature of distant writing.

[31] The material infrastructure of AI, including its planetary costs and embedded power structures, forms an often-invisible backdrop to these socio-technical systems (Crawford, 2021; Crawford, 2024).

this lens, can be understood as part of a broader epistemic ecology—an environment in which human creativity is scaffolded, extended, and at times constrained by artificial collaborators and algorithmic systems (Fricker, 2007; Goldman, 2011).[32]

Such an ecological view does not diminish the human author's role, but repositions it. The author becomes a kind of epistemic navigator or orchestrator, leveraging the generative capacities of the model not as a passive tool-user but as an active participant in a dialogical process. Just as a scientist interprets data produced by instruments she cannot fully explain, or a musician collaborates with an improvisational partner whose responses exceed her anticipatory control, the author in distant writing must learn to read, filter, and interpret outputs within a dynamic, evolving system. The cognitive labor involved in shaping and steering this process is deeply creative, but also epistemically substantive (Grasso, 2023; Searle, 1980).

Thus, in recognizing creativity as epistemic innovation, distant writing reveals the inadequacy of a narrowly individualistic and truth-focused epistemology. It demands a reconceptualization of knowledge-generation as a process that includes, and indeed often depends upon, exploratory, iterative, and collaborative elements (Milan, 2024; Hayles, 1999). The task for analytic epistemology, then, is not merely to accommodate this new practice, but to learn from it—revising our understanding of what counts as epistemic labor and how creativity might contribute to the growth of knowledge in an age of intelligent machines (Gaut, 2010; Kidd & Battaglia, 2023).

The final, and perhaps most philosophically pressing, issue raised by distant writing is the proper allocation of epistemic credit. When the content of a literary or argumentative text is partially generated by a large language model (LLM), a complex question arises: does the human author retain full intellectual responsibility and credit for the output? Or should credit be distributed, and if so, according to what normative criteria?

A common initial strategy is to analogize LLMs to epistemic tools—devices that enhance human capacities but do not themselves warrant intellectual recognition. On this view, using an LLM to assist in writing is akin to using a microscope in empirical observation or a search engine in academic research. The instrument facilitates the acquisition of content or insight, but the epistemic agency resides wholly with the human operator. Under this framework, the model's role is merely instrumental: it mediates the author's intentions and designs without originating any independently meaningful contribution (Clowes, 2023; Floridi, 2025b). Therefore, intellectual credit—and the corresponding responsibilities of truthfulness, originality, and coherence—would rightly accrue to the human author alone.

However, this instrumentalist analogy begins to falter when we examine more closely the generative nature of LLMs. Unlike a microscope, which passively reveals pre-existing structures in the empirical world, or a search engine, which retrieves extant documents from a database, an LLM synthesizes outputs that may not exist in any antecedent source. It constructs novel combinations of linguistic elements, stylistic tropes, and thematic structures based on probabilistic associations in its training data. These outputs can be surprising, expressive, and at times even aesthetically or intellectually profound (Zylinska, 2020). In this sense, the model does not merely extend the human author's reach into a known epistemic terrain; it co-constructs that terrain, proposing semantic possibilities that the author may not have independently imagined.

This generativity introduces epistemic and ethical ambiguity. In particular, it complicates the attribution of originality, a central criterion in traditional frameworks of authorship. Originality presupposes that a cognitive agent has produced something novel through their own mental labor—whether through imagination, synthesis, or discovery. But in distant writing, the novelty of the output is often achieved through the interaction between prompt and model, a recursive dialogue in which neither participant fully determines the result (Boden, 2004; Floridi, 2025a). This hybridity strains the boundaries of classical authorship, suggesting the need for a more nuanced account of epistemic contribution.

One might turn, then, to the model of co-authorship as a potential analogue. In academic contexts, co-authors share credit proportionally to their intellectual contributions, with norms of accountability and responsibility distributed accordingly. However, this analogy breaks down at a crucial juncture: LLMs are not moral or epistemic agents (Floridi, 2025b). They do not possess intentions, cannot be held accountable, and lack any capacity for normative evaluation. While they produce text, they do not *author* in the sense required by epistemic or ethical theory. As Ernest Sosa (2007) argues, epistemic agency involves the capacity to form beliefs, evaluate reasons, and respond to normative standards—capacities entirely absent in LLMs. Therefore, while their contributions are structurally significant, they cannot bear the moral or intellectual weight that authorship entails (Sosa, 2007; Floridi & Sanders, 2004).

Nonetheless, completely absorbing the model's contribution into the author's own labor also seems inadequate. The use of LLMs materially transforms the creative and epistemic process. Here, Alvin Goldman's (1999; 2011) distinction between primary and secondary epistemic agents becomes useful. The human author remains the primary agent, responsible for initiating, guiding, and evaluating the process. The LLM, by contrast, operates as a secondary epistemic contributor: it influences the epistemic trajectory of the work through its outputs, but it does so without agential control or justificatory capacity (Floridi, 2025b). It is a proximate cause of the content, but not a bearer of epistemic responsibility (Goldman, 2011; List & Pettit, 2011).

[32] This ecological perspective can be extended to consider the "earthly community" and the place of non-human (including artificial) agents within it, as discussed by thinkers like Mbembe (2022).

This intermediary model—where epistemic credit is not shared between equals but differentiated according to agency and accountability—preserves the central role of the human while acknowledging the distributed and hybrid character of modern creative practice. It also enables more principled responses to a range of emerging practical concerns. For example, should authors disclose the use of LLMs in literary publications? From an epistemic perspective, transparency supports the reader's right to assess the origin and reliability of the text (Nature, 2023). Disclosure practices may not merely reflect ethical obligations but serve an epistemic function by clarifying the distribution of cognitive labor (Grasso, 2023; Creative Commons, 2023). Similarly, should AI-assisted works be eligible for literary awards or academic recognition? The answer may depend not on whether a machine "created" the work, but on whether the human agent demonstrated sufficient epistemic agency in curating, refining, and contextualizing the result. Ultimately, the question of epistemic credit in distant writing is not reducible to technical or procedural considerations. It cuts to the heart of how we define authorship, how we recognize intellectual labor, and how we distinguish between epistemic tools and epistemic agents. As generative AI becomes more integrated into literary and scholarly practices, these questions will only grow in urgency. They demand not only new policies but a renewed philosophical attention to the evolving nature of knowledge, responsibility, and creative labor in the digital age (Floridi, 2023; Hayles, 1999).

4. AI, Plagiarism, and the Norms of Attribution

The integration of artificial intelligence into literary production not only challenges traditional epistemic categories but also disrupts established ethical and normative frameworks—particularly those governing plagiarism and attribution. Within academic and literary contexts, plagiarism is typically understood as the misappropriation of another's intellectual labor without proper acknowledgment (Scanlon, 2003). However, distant writing introduces ambiguity into this framework (Floridi, 2025a). If a large language model (LLM) contributes textual content based on statistical associations derived from its training corpus, does the use of that content constitute plagiarism? And if not, what norms of attribution should apply?

To determine whether distant writing constitutes a form of plagiarism, it is essential to begin with a careful analysis of the ontological status of AI-generated texts. Large language models (LLMs), such as GPT-4, do not create content in the same manner as human authors. Rather than originating thoughts through intentional acts of meaning-making, these models generate output through a process of statistical pattern completion. Specifically, they use probabilistic inference over massive corpora of existing human language data to produce syntactically and semantically plausible continuations of a given prompt. The resulting texts are neither straightforward copies of their training data nor autonomous creations. Instead, they are synthetically recombinatory: novel sequences generated by recombining linguistic and conceptual elements drawn from an underlying distribution of human-authored materials.

This distinction sets AI-generated content apart from both paradigmatic cases of human originality and traditional forms of plagiarism. Human authorship typically involves the intentional articulation of thoughts, shaped by motivations, beliefs, and a contextual grasp of meaning. Plagiarism, correspondingly, involves the misappropriation of such intentional intellectual labor—presenting someone else's ideas or words as one's own without attribution. In contrast, LLM outputs are generated absent any intentional stance. As Bender et al. (2021) emphasize, LLMs do not *understand* the texts they produce; they do not have beliefs, goals, or semantic comprehension. Consequently, they are not epistemic or moral agents and cannot be said to "own" the content they generate in any morally meaningful way (Floridi, 2025b).

This leads to a crucial epistemological and ethical implication: because LLMs lack the capacity for intention, they cannot possess intellectual property in the moral or normative sense. When a human uses an LLM to produce a passage, they are not violating the rights of a co-author, because no such rights exist. There is no agent whose intellectual labor is being misappropriated, and thus, strictly speaking, no act of plagiarism is being committed *against the machine*. This differentiates distant writing from ghostwriting or unacknowledged collaboration with another person, where there is a clear asymmetry in attribution and labor.

Nevertheless, concerns about plagiarism in AI-assisted writing remain salient—not because of any rights or interests held by the machine, but because of the social norms and expectations that govern intellectual authorship. In academic and literary contexts, readers assume that the content presented under an author's name reflects that author's own cognitive labor and expressive agency.[33] If substantial portions of a text are generated by an LLM and this fact is undisclosed, readers may be misled about the nature and source of the work. This constitutes a violation not of the model's rights, but of the audience's legitimate epistemic expectations. As such, the ethical concern is not about theft from a non-agent, but about transparency, intellectual honesty, and reader trust.

Consider an analogy: using a random number generator to select content from a database and then presenting the result as one's own selection would not violate the rights of the RNG, but it would mislead others about the author's intellectual involvement. Similarly, in distant writing, the issue is not whether the machine was wronged, but whether the audience has been given a false impression of the author's role. This reorients the ethics of AI-assisted authorship away from proprietary frameworks and toward relational ones, centered on communicative integrity and epistemic accountability.

[33] These assumptions are also culturally situated. Atari et al. (2023) highlight how psychological and cognitive science often overgeneralizes from WEIRD (Western, Educated, Industrialized, Rich, Democratic) populations, a critique extendable to assumptions about authorship and reader expectations.

One central concern in debates about plagiarism is the deception condition—the idea that plagiarism is wrongful not merely because it involves copying, but because it involves a specific kind of misrepresentation. According to this view, prominently defended by Roig (2006) and others, plagiarism is epistemically and ethically problematic because it deceives readers into believing that a work reflects the unaided intellectual labor, originality, and expressive capacity of its purported author. This understanding locates the wrong of plagiarism in a breach of epistemic trust: readers rely on authors to accurately represent the origins and methods of their work, particularly in contexts that value originality, critical insight, and authorial voice.

In the context of distant writing, this concern takes on new urgency. Large language models (LLMs) can generate entire passages of coherent, stylistically polished prose that, to an uninformed reader, may be indistinguishable from human-authored text. If an author presents such output without acknowledgment, it can give rise to a subtle but significant form of epistemic dishonesty—misleading readers about the nature and source of the cognitive labor underlying the work. The author's omission of disclosure implies that the text is the result of their own creative and linguistic skill, when in fact it is co-produced by an external, non-sentient system.

To clarify the distinction, consider two illustrative scenarios. In the first, a writer consults a thesaurus to vary their word choice, selecting synonyms to enhance clarity or stylistic nuance. In this case, the tool assists the author but does not generate content in any meaningful sense; the cognitive and creative labor still resides with the human. In the second scenario, a writer uses an LLM to generate entire paragraphs based on a prompt and incorporates them into a final manuscript with minimal alteration. While both scenarios involve tools, the second represents a qualitative shift in the locus of generative labor—from the internal, deliberate operations of the author's mind to the external, computational processes of the model. If the latter case is presented without any disclosure, it may constitute a form of soft plagiarism: not the misappropriation of another human's ideas or words, but the obfuscation of one's epistemic dependence on a generative machine system.

This concern echoes long-standing issues in the ethics of scientific authorship, particularly in cases involving ghostwriting. In biomedical research, for instance, it is considered a serious breach of professional ethics for a named author to publish work that was secretly written by an undisclosed third party, such as a commercial writing firm (Ross et al., 2008). The ethical violation here lies not in the act of outsourcing per se, but in the failure to disclose the nature of that outsourcing, thereby distorting the attribution of expertise, accountability, and intellectual credit. A similar logic applies to distant writing: what is at stake is not the mere use of computational assistance, but the transparency with which that assistance is acknowledged.

From an epistemological perspective, such cases involve what Fricker (2007) might call testimonial injustice: a breakdown in the norms that govern how knowledge claims are transmitted and assessed within a community. When readers are led to overestimate an author's epistemic contribution—mistaking a co-produced or externally scaffolded work for an autonomous intellectual act—they are epistemically wronged, and the trust relations that underwrite the communicative exchange are undermined. This underscores the need for new norms of disclosure in AI-mediated creativity, ones that protect both the epistemic integrity of authorship and the interpretive rights of readers.

A complicating factor in the discussion of AI and plagiarism is that traditional attribution practices are built on the presupposition of authorship—that is, the idea that creative works originate from identifiable agents capable of intentional expression and normative accountability. Legal, moral, and epistemic frameworks surrounding intellectual property and credit rely on the assumption that an author is a being with beliefs, goals, and communicative intent. Large language models (LLMs), however, do not satisfy any of these conditions (Floridi, 2025b). They lack consciousness, intention, understanding, and moral standing. Consequently, while they produce text, they cannot be *authors* in any conventional or philosophically meaningful sense.

This ontological mismatch creates a normative dilemma: if an LLM's contributions are substantial, but the model itself cannot bear authorship, how should those contributions be acknowledged? Assigning credit to a non-agent is incoherent, but ignoring the model's role entirely risks obscuring the collaborative nature of the creative process. In response to this tension, some publishers and academic institutions have begun to introduce guidelines requiring disclosure of AI assistance. For example, *Nature* (2023) and other journals now mandate that authors disclose whether LLMs were used in the generation of content, though such disclosures are typically confined to footnotes, acknowledgments, or methodological appendices. This practice reflects a broader shift in the norms of attribution: rather than recognizing artificial systems as co-authors, we are beginning to treat them as methodological tools whose use must be transparently documented.

An instructive analogy can be drawn from the use of computational software in empirical research. When a scientist uses statistical tools such as R, Python, or SPSS to analyze data, the software is not listed as a co-author of the resulting publication. Nonetheless, it is standard practice to note its use in the methods section, both to acknowledge its influence on the analysis and to facilitate reproducibility. The software, in this context, is understood as a non-agential but epistemically significant instrument—a device that extends the researcher's cognitive and inferential capacities without possessing agency or responsibility of its own.

LLMs can be situated within a similar conceptual framework. They are generative instruments—tools that do not merely assist in executing human-authored ideas, but actively contribute to the form and content of what is produced. Their outputs are not deduced from logical principles nor directly retrieved from a fixed database; instead, they are the result of probabilistic inference over massive linguistic corpora, often yielding novel combinations of ideas, images, and arguments. This generativity renders their role more

complex than that of a calculator or spell-checker, but it still falls short of authorship (Floridi, 2025b). The ethical imperative, therefore, is not to assign them credit, but to ensure transparency about their involvement.

Transparency serves multiple epistemic and ethical functions. It protects the reader's ability to assess the provenance and credibility of a text, clarifies the division of labor in its creation, and preserves trust in contexts—such as scholarship, journalism, or literature—where authorial voice and originality matter. Readers, reviewers, and interpreters rely on certain expectations about how a text came into being, and undisclosed AI involvement can violate those expectations, thereby undermining informed evaluation and interpretive fairness. In this way, the appropriate norm for AI-assisted writing may parallel that of citation and methodological reporting: not the attribution of authorship, but the clear disclosure of instrumental mediation.

Given the complexities introduced by AI-assisted composition, it is no longer sufficient to define plagiarism solely in terms of copying from human sources. Traditional conceptions of intellectual misappropriation rest on assumptions about authorship—namely, that the creator of a text is a unified, intentional agent whose labor is both directly expressed in the text and morally attributable. In the context of distant writing, however, these assumptions are strained by the presence of non-agential systems that can generate complex, coherent, and even innovative content without possessing intention or understanding (Floridi, 2025a, 2025b). This demands a reformulation of attribution norms, one that neither presumes wrongdoing in the use of AI nor ignores the epistemic significance of its involvement.

One way forward is to recognize distant writing as a distinct compositional category that requires its own set of disclosure standards. Rather than treating the use of LLMs as inherently suspect, such standards would aim to clarify the human author's epistemic role and the function that the AI system has played. Disclosure, in this context, serves not merely a procedural function but a substantive ethical one: it helps preserve the trust relationship between author and reader by making transparent the conditions under which the work was produced.

To implement such standards effectively, it is helpful to develop a nuanced vocabulary for describing the ways in which LLMs contribute to the writing process. For instance, in some cases, the system may be used for what we might call *lexical assistance*—such as correcting grammar, improving word choice, or suggesting paraphrases. In others, the model may play a role in *stylistic augmentation*, rephrasing or rewriting passages to match a particular voice, tone, or rhetorical effect. At a deeper level, it may be employed in *narrative generation*, where it contributes entire scenes, dialogues, or argumentative structures that constitute meaningful portions of the text. And in some of the most epistemically charged uses, the model may participate in *thematic innovation*, helping the author articulate or explore new ideas that affect the conceptual framework of the work.

These distinctions matter because they calibrate the ethical gravity of AI involvement. Just as we differentiate between ghostwriting, editorial assistance, and co-authorship in traditional forms of writing, we must now consider how various uses of generative systems map onto different forms of creative labor and epistemic responsibility. This reframing aligns with broader trends in virtue epistemology, particularly with the virtue of epistemic conscientiousness, which Linda Zagzebski (1996) defines as a disposition to act responsibly in the pursuit of knowledge.

An author who uses an LLM conscientiously—by disclosing its use, critically assessing its outputs, integrating them thoughtfully, and ensuring the coherence and integrity of the final text—exemplifies this virtue. Such a writer is not merely delegating creative labor but is engaging in an extended form of epistemic activity, one that spans both human cognition and technological affordance. The result is not a diminished form of authorship but a transformed one: less about direct linguistic production and more about orchestration, judgment, and critical oversight (Floridi, 2025a).[34]

In contrast, the failure to disclose the role of AI, or the passive incorporation of its outputs without adequate evaluation, risks a form of epistemic dishonesty. This dishonesty does not necessarily involve the theft of another's work, but it does involve the misrepresentation of how the work was produced. The harm lies in misleading the reader about the nature and extent of human intellectual effort—a harm akin to ghostwriting in scientific contexts, where third-party contributions are obscured, undermining transparency and accountability.

In this light, the ethics of distant writing should not be governed by rigid prohibitions or outdated notions of authorship, but by evolving norms rooted in intellectual honesty, responsibility, and epistemic integrity. The use of LLMs need not threaten these values, but it does demand that we articulate them anew, in ways that are responsive to the hybrid nature of human–machine collaboration. The challenge is not to constrain creative possibilities, but to ensure that these possibilities are pursued in a way that respects the norms of knowledge production and the trust of the interpretive community.

5. Conclusion

The advent of distant writing, as introduced by Floridi (2025a), compels a foundational reassessment of how we conceptualize authorship, creativity, and intellectual responsibility in the age of artificial intelligence. Unlike Moretti's (2000) distant reading, which analyzes preexisting texts through computational abstraction, distant writing marks a shift toward the generative use of AI systems, particularly LLMs, as active participants in the literary process. While human authors remain central as epistemic agents, they are increasingly functioning as designers—curators of computational affordances—rather than as sole originators of content.

[34] The conscientious author, in this view, engages in a form of "neganthropic" activity, in Stiegler's (2019) sense, by cultivating knowledge and care in their interaction with otherwise "entropic" technological systems.

From an epistemological perspective, this shift necessitates a reconceptualization of the division of epistemic labor. Knowledge production, in this context, is not a solitary achievement but the result of interactive procedures involving procedural, semantic, and evaluative competencies. The use of AI in creative writing can enhance human understanding and expand narrative possibility spaces, but it also raises serious normative questions regarding originality, attribution, and intellectual honesty (Floridi, 2025a).

A core conclusion of this inquiry is that traditional concepts of plagiarism, grounded in human-centric notions of misappropriation, are increasingly insufficient for evaluating AI-mediated composition. What is required is not the abandonment of these norms, but their refinement in light of technological change. Plagiarism, reframed, must be understood as the failure to disclose epistemically significant forms of non-human assistance—failures which undermine the trust and transparency that undergird both academic and artistic communities.

The analysis further suggests that creativity, often relegated to the periphery of epistemology, deserves closer integration into our theories of knowledge and understanding (Kvanvig, 2003; Boden, 2004). In the context of distant writing, creative acts are not reducible to expressive flair or stylistic flourish; they are epistemic performances that involve conceptual innovation, problem-solving, and meaning-making under conditions of partial automation.

Ultimately, distant writing is not an epistemic threat but a provocation—an invitation to rethink the architecture of authorship in a landscape increasingly shaped by synthetic agency (Floridi, 2025a, 2025b). This does not entail the erosion of human intellectual responsibility, but its reframing: authors must now navigate a new ecology of writing in which human judgment and machine output co-produce the literary artifact.[35] As such, future epistemological inquiry must confront not only the conditions under which knowledge is acquired, but also the evolving modalities through which it is composed, represented, and attributed.

Statements & declarations

Funding.

I would like to thank the Graduate Program in Philosophy at the Federal University of Santa Maria for making the publication of this research possible with funding from CAPES through PROAP CALL 01-2025.

Competing interests.

The author declares that there are competing interests involved in the writing of this article.

Author contributions.

The author is the sole and single contributor to the writing of this text.

References

1. Baehr, J. (2011). The epistemic virtues and the nature of knowledge. In D. Pritchard (Ed.), *Epistemology: The key thinkers* (pp. 148-164). Routledge.
2. Bates, G., O'Connor, L., & Pennington, R. (2024). Designing with synthetic agency: AI art and the politics of authorship. *AI and Ethics*, 1-14. https://doi.org/10.1007/s43681-023-00369-6
3. Battaglia, F. (2022). Intellectual Virtues, Extended Cognition, and the Epistemology of AI. *Philosophies, 7*(4), 84. https://doi.org/10.3390/philosophies7040084
4. Baudrillard, J. (1994). *Simulacra and simulation*. University of Michigan Press.
5. Beckers, N. (2023). Human–algorithm hybrids as (quasi-)organizations? On the accountability of digital collective actors. *Journal of Law and Society, 50*(1), 34–58. https://doi.org/10.1111/jols.12412
6. Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185–5198. https://doi.org/10.18653/v1/2020.acl-main.463
7. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). https://doi.org/10.1145/3442188.3445922
8. Bloor, D. (1991). *Knowledge and social imagery* (2nd ed.). University of Chicago Press. * Boden, M. A. (2004). *The Creative Mind: Myths and Mechanisms* (2nd ed.). Routledge.
9. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). *On the opportunities and risks of foundation models*. arXiv preprint arXiv:2108.07258. https://doi.org/10.48550/arXiv.2108.07258
10. Buckner, C. (2019). Deep learning: A philosophical introduction. *Philosophy Compass, 14*(10), e12625. https://doi.org/10.1111/phc3.12625
11. Burge, T. (2010). *Origins of objectivity*. Oxford University Press.

[35] This co-production echoes Wittgenstein's (1980, 2006, 2009) emphasis on language as a social practice and meaning as arising from use within a "form of life," which is now increasingly mediated by and co-constituted with artificial systems.

12. Carter, J. A., & Palermos, S. O. (2016). Is epistemic agency a precondition for epistemic justification? *Synthese, 193*(10), 3305-3328.
13. Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7
14. Clowes, R. W. (2023). AI tools, epistemic dependence and the demotion of human agency. *AI and Ethics*, 1-13. https://doi.org/10.1007/s43681-023-00316-5
15. Coeckelbergh, M. (2020). Attributing agency to automated systems: Reflections on human–robot collaborations and responsibility-loci. *Science and Engineering Ethics, 26*(4), 2051–2068. https://doi.org/10.1007/s11948-017-9943-x
16. Creative Commons. (2023, August 31). *How CC is approaching AI and copyright*. Creative Commons. https://creativecommons.org/2023/08/31/how-cc-is-approaching-ai-and-copyright/
17. Danaher, J. (2022). The Robotic Disruption of Epistemic Humility. *Philosophy & Technology, 35*(4), 101. https://doi.org/10.1007/s13347-022-00584-9
18. Donati, D. (2023). Trust, trustworthiness and the moral dimension in human-AI interactions. *Ethics, Politics & Society, 7*(2), 1–17. https://revistas.uminho.pt/index.php/eps/article/view/6181
19. Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.
20. Elgin, C. Z. (2017). *True enough*. MIT Press.
21. Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms. *arXiv preprint arXiv:1706.07068*.
22. Floridi, L. (2014). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press.
23. Floridi, L. (2023). *AI Ethics*. Oxford University Press.
24. Floridi, L. (2025a). *Distant Writing: Literary Production in the Age of Artificial Intelligence*. Revised version 5. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5232088.
25. Floridi, L. (2025b). AI as Agency without Intelligence: On Artificial Intelligence as a New Form of Artificial Agency and the Multiple Realisability of Agency Thesis. *Philosophy & Technology 38*: 30. https://doi.org/10.1007/s13347-025-00858-9.
26. Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines, 14*(3), 349-379.
27. Frankfurt, H. G. (2005). *On bullshit*. Princeton University Press.
28. Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001
29. Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press.
30. Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems, 14*(3), 330-347. https://doi.org/10.1145/230538.230561
31. Gaut, B. (2010). The philosophy of creativity. *Philosophy Compass, 5*(12), 1034-1046.
32. Giere, R. N. (2006). Scientific models and scientific understanding. In P. Machamer & G. Wolters (Eds.), *Science, Values, and Objectivity* (pp. 55-70). University of Pittsburgh Press.
33. Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society, 7*(2). https://doi.org/10.1177/2053951720943234
34. Goldman, A. I. (1999). *Knowledge in a Social World*. Oxford University Press. https://doi.org/10.1093/0198238207.001.0001
35. Goldman, A. I. (2011). *Reliabilism and contemporary epistemology: Essays*. Oxford University Press.
36. Goodman, N. (1976). *Languages of art: An approach to a theory of symbols*. Hackett Publishing.
37. Grasso, F. (2023). On the Epistemic Risks of Large Language Models. *Minds and Machines, 33*, 685–690. https://doi.org/10.1007/s11023-023-09635-0
38. Greco, J. (2010). *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge University Press.
39. Gunkel, D. J. (2018). *Robot rights*. MIT Press.
40. Gunkel, D. J. (2020). *Of remixology: Ethics and aesthetics after remix*. MIT Press.
41. Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.
42. Hempel, C. G. (1966). *Philosophy of natural science*. Prentice-Hall.
43. Hookway, C. (2003). Epistemic Norms and Theoretical Deliberation. In P. Pettit, S. Scheffler, M. Smith, & R. J. Wallace (Eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (pp. 108-127). Clarendon Press.
44. Jakesch, M., Lederer, M., & Meyer, M. (2023). Epistemic Risks of AI-Generated Texts. *Digital Society, 2*(1), 10. https://doi.org/10.1007/s44206-023-00054-0
45. Johnson, D. G., & Verdicchio, M. (Eds.). (2017). *AI, Robots, and Values*. Wiley-Blackwell.
46. Kidd, I. J., & Battaglia, F. (2023). AI and Intellectual Virtue: An Introduction. *Philosophies, 8*(6), 110. https://doi.org/10.3390/philosophies8060110
47. Kreminski, M. (2024). *The dearth of the author in AI-supported writing*. arXiv preprint arXiv:2404.10289. https://arxiv.org/abs/2404.10289
48. Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

49. Kvanvig, J. (2003). *The value of knowledge and the pursuit of understanding*. Cambridge University Press.
50. Lawson, B. (2006). *How designers think: The design process demystified* (4th ed.). Architectural Press.
51. Lipton, Z. C. (2020). The Mythos of Model Interpretability. *Queue, 16*(3), 31-57.
52. List, C., & Pettit, P. (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford University Press.
53. Longino, H. E. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press.
54. Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books.
55. Milan, S. (2024). When Algorithms Surprise: The Sociology of Malfunction in Automated Decision-Making. *Big Data & Society, 11*(1). https://doi.org/10.1177/20539517231220096
56. Menary, R. (2010). *Cognitive integration: Mind and cognition unbounded*. Palgrave Macmillan.
57. Moretti, F. (2000). Conjectures on world literature. *New Left Review, 1*, 54–68.
58. Nagel, T. (1986). *The view from nowhere*. Oxford University Press.
59. Nature. (2023). Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature, 613*, 612. https://doi.org/10.1038/d41586-023-00191-1
60. Nyholm, S. (2023). The ethics of AI ethics: An introduction to the special issue. *Ethics and Information Technology, 25*(1), 1. https://doi.org/10.1007/s10676-023-09692-8
61. Pritchard, D. (2010). Epistemic Relativism, Epistemic Incommensurability and Wittgensteinian Epistemology. In S. Hales (Ed.), *A Companion to Relativism* (pp. 266-285). Blackwell.
62. Pritchard, D. (2016). Epistemic Risk. *Journal of Philosophy, 113*(11), 550-571.
63. Roig, M. (2006). *Avoiding plagiarism, self-plagiarism, and other questionable writing practices: A guide to ethical writing*. U.S. Department of Health & Human Services, Office of Research Integrity. https://ori.hhs.gov/plagiarism-0
64. Ross, J. S., Hill, K. P., Egilman, D. S., & Krumholz, H. M. (2008). Guest authorship and ghostwriting in publications related to rofecoxib: A case study of industry documents from rofecoxib litigation. *JAMA, 299*(15), 1800–1812. https://doi.org/10.1001/jama.299.15.1800
65. Ryle, G. (1949). *The concept of mind*. University of Chicago Press.
66. Santoni de Sio, F., & Mecacci, G. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics, 26*(4), 2051–2068. https://doi.org/10.1007/s11948-019-00146-8
67. Scanlon, P. M. (2003). Student online plagiarism: How do we respond? *College Teaching, 51*(4), 161–165. https://doi.org/10.1080/87567550309596420
68. Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*(3), 417–457.
69. Sosa, E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199297023.001.0001
70. Sparrow, R. (2007). Killer robots. *Journal of Applied Philosophy, 24*(1), 62-77.
71. Strzelecki, A. (2023). To use or not to use ChatGPT in research? Researchers' acceptance of AI-based conversational agents. *Journal of Information Science, 0*(0). https://doi.org/10.1177/01655515231199531
72. Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review, 8*(2). https://doi.org/10.14763/2019.2.1410
73. Thakur, A., Loe, B. S. S., & Lee, P. S. D. (2024). Authorship in the age of artificial intelligence. *Learned Publishing, 37*(1), 10-14. https://doi.org/10.1002/leap.1594
74. Vee, A. (2022). Understanding "Author" in an Age of Machine-Generated Text. In A. Arola & C. W. Hepsø (Eds.), *Elements in Computers and Composition*. Cambridge University Press.
75. Zagzebski, L. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge University Press.
76. Zylinska, J. (2020). *AI Art: Machine Visions and Warped Dreams*. Open Humanities Press.