



AN EPISTEMIC VIRTUE-THEORETIC SOLUTION TO THE PROBLEM OF AI AUTHORSHIP *

Uma Solução Baseada na Epistemologia da Virtude para o Problema da Autoria de IA

Tiegue Vieira Rodrigues **

Abstract: Luciano Floridi's concept of *distant writing*, which is the textual production mediated by generative AI, has presented us with a challenge to traditional accounts of authorship and epistemic agency. In this paper, we analyze the epistemological consequences of this new paradigm. Our argument is that while distant writing seems to reduce the author's role, it reframes it, shifting the locus of epistemic competence from direct textual composition to the second-order acts of goal setting, prompting, and curation. By comparing this phenomenon to Ernest Sosa's virtue-theoretic account of epistemic agency, we develop an epistemology of authorship for the AI era. According to the view we advocate, the *Apt Curation Model*, the human author remains the central epistemic agent, not as the originator of sentences, but as the virtuous curator whose skill makes the final text an apt epistemic performance. I conclude by considering and rebutting the objection that the opacity of AI models reduces authorial success to mere epistemic luck.

Keywords: Generative AI. Authorship. Virtue epistemology. Epistemic Agency, Luciano Floridi. Epistemic Responsibility.

* Artigo recebido em 03.10.2025 e aprovado para publicação em 13.11.2025.

** Doctor of Philosophy from the Pontifical Catholic University (PUCRS), Adjunct Professor of Philosophy in the Department of Philosophy at the Federal University of Santa Maria (UFSM), Coordinator of the Postgraduate Program at UFSM, Santa Maria, RS, Brazil. ORCID: <https://orcid.org/0000-0001-6011-4788>. E-mail tiegue.rodrigues@gmail.com

Resumo: O conceito de escrita distante (*distant writing*) de Luciano Floridi, que é a produção textual mediada por IA generativa, apresenta um desafio às concepções tradicionais de autoria e agência epistêmica. Neste artigo, analisamos as consequências epistemológicas desse novo paradigma. Nosso argumento é que, embora a escrita distante pareça reduzir o papel do autor, ela o reformula, deslocando o locus da competência epistêmica da composição textual direta para os atos de segunda ordem de estabelecimento de metas, elaboração de *prompts* (comandos) e curadoria. Ao comparar esse fenômeno com a concepção da epistemologia da virtude de Ernest Sosa sobre a agência epistêmica, desenvolvemos uma epistemologia da autoria para a era da IA. Segundo a visão que defendemos, o Modelo da Curadoria Apta (*Apt Curation Model*), o autor humano permanece o agente epistêmico central, não como o originador das sentenças, mas como o curador virtuoso cuja habilidade torna o texto final uma performance epistêmica apta. Concluo considerando e refutando a objeção de que a opacidade dos modelos de IA reduz o sucesso autoral à mera sorte epistêmica.

Palavras-chave: IA Generativa. Autoria. Epistemologia da Virtude. Agência Epistêmica. Luciano Floridi. Responsabilidade Epistêmica.

Introduction

The traditional view of authors shows them working alone as intellectual workers who develop concepts through detailed sentence construction. The image combines Romantic genius with Enlightenment reason to show authors who achieve a perfect mental-textual connection during their work. The emergence of advanced generative artificial intelligence systems has made the traditional authorial image completely outdated. The modern author across scientific to academic to student fields now works with large language models (LLMs) which generate, combine, and summarize texts at an exceptional level of speed. The technological transformation introduces more than operational changes because it challenges our basic ideas about authorship, originality, and epistemic control which threatens to diminish human experiences.¹

Luciano Floridi² identifies *distant writing* as the fundamental concept which defines this emerging writing method. The process of distant writing involves human operators who direct semi-autonomous generative systems through composition tasks while keeping themselves separate from *direct writing* activities. AI operates as a generative partner which produces semantic content through innovative processes while forming a cognitive system that works with the user.³ The collaborative process

¹ THORSTAD, D. The epistemic threat of lazy expertise, in *Episteme*, v. 20, n. 4, 2023, p. 812-828.

² FLORIDI, L. *Distant Writing: Literary Production in the Age of Artificial Intelligence*. Revised version 5. [S.l.: s.n.], 2025.

³ CLARK, A.; CHALMERS, D. The extended mind, in *Analysis*, v. 58, n. 1, 1998, p. 7-19.

between humans and AI systems produces an untraceable link between original intentions and final written output which breaks down the conventional author-text relationship. The current situation creates an intellectual and ethical dilemma about how we should distribute intellectual and epistemic authority within our established systems. The collaborative textual production process makes it difficult to determine which person should receive credit for ideas and bear responsibility for mistakes.

We investigate, in this paper, the fundamental ethical problem which emerges from this crisis. The main issue requires a solution to create a normative framework for authorship which maintains human control and knowledge-based accountability during distant writing operations. A proper solution to this question requires moving past basic technology explanations to develop an established epistemological structure which handles this emerging intellectual work distribution system. The author's role undergoes transformation through distant writing even though it seems to diminish their overall involvement in the process. The main focus of epistemic competence shifts from creating the text directly to managing three essential functions which include goal setting, suggestion, and content curation.

The transformation of direct composition work into curation practice enables us to maintain authorship according to our proposed solution. The AI era requires authors to execute four essential tasks of content selection, refinement, and integration and validation of AI-generated material according to a virtue-theoretical framework.⁴ The author deserves recognition for their intellectual work in distant writing because their epistemic success stems from curatorial virtues which lead to the creation of authentic intellectual products.⁵

The paper follows this structure to support our main argument. The first section of this paper explains distant writing as a control problem which disrupts traditional methods of authorship attribution. The Socratic interlude in Section 3 uses Sosa's virtue-theoretical framework to show how the problem generates an unsolvable challenge for traditional compositional agency models. The Apt Curation Model serves as the paper's main solution which consists of four distinct curatorial virtues. The following sections establish the defense of this model. The opacity objection in Section 5 represents the most powerful counterargument which argues

⁴ SOSA, E. *A virtue epistemology: Apt belief and reflective knowledge*. Oxford: Oxford University Press, 2007, v. 1. And, SOSA, E. *Reflective knowledge: Apt belief and reflective knowledge*. Oxford: Oxford University Press, 2009, v. 2.

⁵ GRECO, J. *Achieving knowledge: A virtue-theoretic account of epistemic normativity*. Cambridge: Cambridge University Press, 2010.

that curation amounts to untrustworthy submission to an unreliable information source. The two-part defense in Section 6 uses Floridi's Levels of Abstraction (LoA) to show how curatorial virtues enable responsible subordination. The paper investigates three main areas where this model creates changes in our understanding of originality and educational requirements, intellectual property, and law regulations. Concluding statements are presented in section 8.

1. Distant Writing and the Problem of Locus of Control

The particular form of human-AI collaboration underlying this exploration is referred to by Floridi as *distant writing*.⁶ The philosophical utility of the concept lies not in its technological specificity, but in the conceptual shift it imposes on our understanding of cognitive artifacts.⁷ For centuries, our writing tools have been extensions of the author's direct will, designed to convey their thoughts more quickly, clearly, and sustainably. From the pen to the word processor, the role of technology has been fundamentally prosthetic: it executes but does not subordinate. This process by which technologies co-shape human action and experience is a form of *technological mediation*.⁸ *Distant writing*⁹ represents a categorical break with this tradition. The novelty of this practice and the source of its epistemological challenge lie in the nature of AI's contribution as a

⁶ Floridi's (The method of levels of abstraction, in *Minds and Machines*, v. 18, n. 3, 2008, p. 303-329) framing of distant writing draws a deliberate and illuminating analogy to Franco Moretti's (Conjectures on world literature, in *New Left Review*, n. 1, 2000, p. 54-68) 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. Moretti's approach was interpretative and retrospective: it sought to understand and map the literary past by aggregating and analyzing extant texts.

⁷ Cognitive artifacts are usually understood as artificial devices that maintain, display, or operate on information in order to serve a representational function and that affect human cognitive performance. See HUTCHINS, E. *Cognition in the wild*. Cambridge, MA: MIT Press, 1995.

⁸ VERBEEK, P. P. *What things do: Philosophical reflections on technology, agency, and design*. University Park, PA: Penn State Press, 2005. Verbeek's *theory of mediation* argues that technologies are not neutral intermediaries but actively co-shape the relationship between humans and their world, influencing perception, action, and moral experience.

⁹ 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. 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, X.E., PÉREZ-VERDUGO, M. Generative midtended cognition and Artificial Intelligence: thinging with thinging things, in *Synthese* 205, 137 (2025). <https://doi.org/10.1007/s11229-025-04961-4>

generative partner,¹⁰ a difference that becomes clear when compared to other cognitive technologies.

Consider a calculator. It is a perfect example of a prosthetic cognitive tool. It performs complex mathematical operations with the utmost precision, far exceeding the capabilities of the human mind. Nonetheless, we can see that its function is solely determined by the explicit commands of the user. The user inserts data as input and specifies the kind of operation (e.g. multiplication). Then, the device conveys a deterministic result. The user retains full control over the high- and low-level epistemic processes; the calculator expands their cognitive horizons but does not provide new semantic content or make autonomous decisions. The intellectual responsibility for the final calculation and its application lies solely with the user.

A search engine represents a more complex case. It is an information retrieval tool whose workings remain hidden from most users. The system retrieves extensive human-created data through queries to generate ordered results that might match the search criteria. The function operates as a reference tool because it shows users content instead of producing new material. The entire process of evaluating information reliability and extracting data and statement synthesis and argument construction remains under human control. Search engines are able to provide the raw material, but the user continues to be the architect and builder.

Generative AI, as used in distant writing, operates on a completely different level. It is not merely prosthetic, referential, or corrective (like a grammar checker that refines an existing text). It is a generative partner. When a scholar commands an LLM: 'Write a three-paragraph summary of the Gettier problem¹¹ and explain its relevance to justified and true beliefs', the AI doesn't simply retrieve a relevant article. It creates a new artifact. It selects the relevant concepts, structures the sentences, determines the logical flow between them, and presents the whole in coherent prose. The human author is at a cognitive and causal *distance*.¹² They are responsible for initiating and guiding the process but are not in any relevant sense the direct cause of the words on the page. The control over the fundamental details of composition (the choice of vocabulary, the syntactical arrangements, the rhetorical flourishes) has been delegated to a non-human system.

¹⁰ The concept of a 'generative partner' should be distinguished from the classic 'extended mind' thesis, which argues that external tools can, under certain conditions, become literal parts of a cognitive process. The generative partner model, in contrast, emphasizes a collaborative division of labor between two distinct (though coupled) systems: one human, one artificial. For reference see footnote 3.

¹¹ See GETTIER, E. L. Is justified true belief knowledge?, in *Analysis*, v. 23, n. 6, 1963, p. 121-123.

¹² See footnote 2.

This delegation of control is the source of the central challenge that distant writing poses for epistemology, a challenge best formulated as a locus of control problem. The locus of control problem is the fundamental ambiguity introduced by the joint creation of a text by a human and a generative AI regarding the origin and attribution of the text's central properties. This reflects classic problems of distributed agency, where it becomes difficult to attribute responsibility for an outcome to a single agent within a collective.¹³ We can separate these properties into two interconnected categories. First, there are the *epistemic properties*: truth, falsity, justification, and warrant. If an AI-assisted text makes an assertion that "Kant's categorical imperative can be derived from basic utilitarian principles," who should we attribute the epistemic status of this assertion to? Is the assertion justified by the human's prior knowledge, which might underlie the suggestion, or is it an unjustified *confabulation*¹⁴ of the model's statistical correlations? Secondly, semantic properties include meaning, reference, style, and argumentative structure. Whose "voice" animates the text?

The locus of control is not lost but distributed in a way that is both *novel* and *opaque*. The human retains overall architectural control (determining topic, thesis, and constraints), while the AI exercises subordinate, compositional control (generating the specific sentences that embody the human plan). This distributed control makes unambiguous attribution of intellectual property nearly impossible under our current conceptual frameworks.

To illustrate the challenges, we can take a vivid example. Consider a bioethicist, Dr. Chen, who holds a master's degree in law and is writing a thesis on the ethics of genetic optimization. She struggles to formulate a concrete objection to her own view. She challenges the AI: 'Develop a strong, philosophically sophisticated counterargument to the position that germline genetic optimization is morally permissible if it aims to improve the common good'. The model presents a highly convincing argument, synthesizing concepts from deontology and communitarianism in a way Dr. Chen had not previously considered. It even includes a particularly apt, but ultimately fabricated, quote attributed to a real philosopher. Dr. Chen recognizes the genius of the argument, corrects the fabricated quote, and integrates the argument into her work.

The locus of control problem can now be seen as manifested in several ways. The central idea of the counterargument (the new synthesis of

¹³ PETTIT, P. Responsibility incorporated, in *Ethics*, v. 117, n. 2, 2007, p. 171-201.

¹⁴ The term 'confabulation' is borrowed from neuropsychology, where it describes the production of fabricated or distorted memories without the conscious intention to deceive. Its application to LLMs captures the phenomenon of generating plausible-sounding falsehoods based on statistical patterns rather than factual recall. For a technical overview of this phenomenon, often termed 'hallucination', see JI, Z. *et al.* Survey of hallucination in natural language generation, in *ACM Computing Surveys*, v. 55, n. 12, 2023, p. 1-38.

deontological and communitarian ideas) does not originate from Chen's expertise. It was generated by the AI. Although it had the expertise to recognize its value, it cannot claim to have *authored*¹⁵ it in the traditional sense. The epistemic recognition of this part of her work is now deeply ambivalent. On the other hand, the factual error (the fabricated quote) also originates with the AI. While Dr. Chen is undoubtedly epistemically responsible for not first checking and then correcting the error, the source of the untruth lies with her generating partner. This is not a simple *typo* of her own making; it is new misinformation introduced by the tool itself. Her control over the factuality of the text was incomplete.

This philosophical problem has immediate practical consequences for the normative structures that govern intellectual life. First, it destabilizes *academic integrity*. The concept of *plagiarism*, for example, is based on a clear distinction between using someone else's work without citing it and creating one's own work. Distant writing exists in the murky space between the two. Is Chen's use of the AI argument a bright contribution from a research assistant or an act of academic dishonesty? Our current definitions cannot properly answer this question. Second, it threatens the economy of *epistemic recognition* (credit).¹⁶ Scientific and humanistic progress depends on a system in which intellectual contributions are attributed to those who make them. This recognition system encourages research and provides a map of the intellectual history of a field. If the source of a groundbreaking hypothesis or a crucial counterexample comes from an LLM, the entire system of attribution is thrown into disarray.

Lastly, we can see that the problem of locus of control is also a problem of *accountability*. If an AI-based legal opinion includes a fabricated case citation that leads to *adverse consequences* in a legal case, or if a medical information bot produces dangerously false advice, who carries the legal and ethical responsibility? The user who entered the prompt? The company that developed and trained the AI? The *opacity* of the generative process makes assigning responsibility a daunting task. The problem, then, is not just a philosophical curiosity. The basic norms that ensure accountability, integrity, and progress in our knowledge production practices are directly challenged and require a rethinking of *epistemic intellectual responsibility* in the age of AI.¹⁷ However, before we can decide how to adapt these norms,

¹⁵ The concept of the 'author' as a specific function, rather than a simple descriptor of the person who writes, has a rich philosophical history. See FOUCAULT, Michel. What is an author? In: FOUCAULT, Michel. *Language, counter-memory, practice: Selected essays and interviews*. Ithaca: Cornell University Press, 1977, p. 113-138.

¹⁶ The idea that knowledge is an achievement for which an agent deserves credit is a central tenet of virtue epistemology. See Greco, J. *Achieving knowledge: A virtue-theoretic account of epistemic normativity*. Cambridge: Cambridge University Press, 2010.

¹⁷ MILLER, L. C. AI and the epistemic responsibilities of the intellectual, in *Philosophy & Technology*, v. 37, n. 1, 2024, Art. 20.

we need to clearly understand why this distributed form of control is so contrary to our traditional theories of knowledge and action, a task to which we now turn.

2. Interlude: The failing of Traditional Models of Action

The problem of locus of control uncovers a central incongruity between the practice of distant writing and our traditional understanding of authorship. To properly provide a diagnostic of such discrepancy with analytical precision, we need a robust normative framework able to evaluate epistemic achievements. We believe that Ernest Sosa's virtue epistemology, particularly his distinction between accurate, adroit, and apt achievements, offers an extraordinarily powerful diagnostic tool for this purpose.¹⁸

Applying this framework to a typical case of distant writing, it becomes clear why traditional models of action that focus on the direct act of writing inexorably fail. This interlude will demonstrate that such models force us into a dilemma that compels us to conclude that the work of the human author is neither *genuinely action-related*¹⁹ nor *epistemically creditable*.

For Sosa, every achievement, from athletic performance to intellectual statement, can be assessed along three dimensions. An action is:

Accurate if it achieves its immediate goal. For an archer's shot, this means hitting the target; for a belief or proposition, it means being true.

Adroit if it manifests a relevant competence or skill of the agent. An archer's shot is adroit if it is well aimed and skillfully executed, relying on the agent's archery skills.²⁰

Apt if it is accurate because it was adroit. This is the crucial normative criterion. The success of the achievement must be due to the agent's exercise of adroitness.

¹⁸ For virtue epistemology accounts see footnotes 4 and 5; additionally see ZAGZEBSKI, L. T. *Epistemic authority: A theory of trust, authority, and autonomy in belief*. Oxford: Oxford University Press, 2012.

¹⁹ The question of what constitutes an 'action' as opposed to a mere 'happening' is a foundational issue in the philosophy of action. An *action* is typically understood as behavior that is *intentional* or done for a reason. See ANSCOMBE, G. E. M. *Intention*. Oxford: Basil Blackwell, 1957.

²⁰ The archer is Sosa's paradigmatic example used throughout his work to illustrate the tripartite structure of performances. He introduces it to distinguish success that is merely attributable to the agent from that which is fully creditable. See Sosa, E. (2007), *A virtue epistemology*, especially Chapter 2.

Aptness is what distinguishes a *genuine* achievement from a *lucky* success. This is a distinction that has been essential to an epistemological analysis since its presentation by Gettier, and the analysis of *epistemic luck* also remains a central task in this field.²¹ An archer might fire a clumsy, misguided shot (a clumsy performance) that is suddenly blown directly onto the target by a gust of wind. The shot might be considered as accurate, but it certainly is not apt. Its success is attributable to the wind, not the archer's skill. For Sosa, knowledge is an *apt* belief: a belief that is true because it was competently formed.²² An achievement worthy of epistemic recognition must be apt.

Now let's apply this AAA framework to a case of distant writing. Consider Dr. Evans, a political scientist researching the concept of legitimacy. He *prompts* an LLM: 'Produce a critical analysis of Rawlsian public reason, focusing on the objections of deliberative democracy theorists, and structure it as a literature review'. The model produces an extremely accurate text that faithfully summarizes the key debates and even makes a subtle connection that Evans hadn't considered. After completing his revision of the text, Evans meticulously cross-checks all the statements against known sources, seamlessly integrating the text into his manuscript. The final result, as many have noted, stands out as an exceptional work of scholarship.

Is Dr. Evans' performance (his literary act) apt? In order to give an answer to this question, we must identify the *adroitness* that leads to the text's *accuracy*. This is where the dilemma emerges.

Horn 1: AI is the adroit actor.

On a simple analysis, the adroitness in this scenario (the ability to compose fluent prose, synthesize complex philosophical arguments, and structure a literature review) is manifested by the AI. The system architecture and training data are the source of the competence that leads to the well-formed text. The AI's performance is adroit. However, if we attribute the adroitness to the AI, Dr. Evans' performance cannot be apt. The text's accuracy is explained by the AI's competence, not his own.

In this view, Dr. Evans' role is comparable to that of a project manager who hires a brilliant but unsupervised subcontractor. The final product may be excellent, but its success is attributable to the subcontractor's skill. Dr.

²¹ PRITCHARD, D. *Epistemic luck*. Oxford: Oxford University Press, 2005.

²² While Sosa defines knowledge as apt belief, other prominent virtue epistemologists offer different formulations. Linda Zagzebski, for example, defines knowledge as a "state of true belief arising out of acts of intellectual virtue", a more motivational account. Our paper's framework, however, relies on Sosa's performance-based model. See ZAGZEBSKI, L. T. *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge: Cambridge University Press, 1996, p. 271.

Evans initiated the process but is not the author in any meaningful sense, and thus the finished text cannot be considered his epistemic achievement. His situation resembles that of Sosa's archer, whose shot is corrected in mid-flight by a technologically advanced "smart arrow" that guides itself to the target. In such cases of cognitive enhancement,²³ attributing success to the agent becomes deeply problematic from a virtue theory perspective.²⁴ The shot is accurate, the arrow's performance is adroit, but the archer's performance is not apt. He gets credit for choosing the target, perhaps, but not for the successful shot itself. If the AI is the adroit agent, then the human author is stripped of the very status we seek to understand.

Horn 2: The human is the adroit agent.

Let's reject the first horn of the dilemma and contend that Dr. Evans is the relevant agent. To do this, we need to locate his adroitness. The best recognizable contender for his skill is *prompt generation*. Perhaps a *virtuous prompter*²⁵ is a kind of author. However, there seems to be a problem with this move, namely, that the causal and explanatory connection between the prompter's adroitness and the accuracy of the lengthy, nuanced text is extremely weak.

LLMs are notoriously unpredictable. A very ordinary prompt may produce a very good result, while a highly specified one may produce a mediocre or hallucinatory result. If the AI makes a novel connection that Evans himself was not able to anticipate, that specific and valuable accuracy cannot plausibly be attributed to the skill manifested in his prompt. In this case, the success of the result far outweighs the skill applied to the input.

This discrepancy between the supposed skill (the prompt) and the successful outcome means that the performance fails the aptness condition. The accuracy of the text is not attributable to the adroit prompter; from the perspective of the agent's competence, this is a matter of *epistemic luck*.²⁶ The situation Dr. Evans is in is comparable to that of a novice chess player

²³ For Carter and Pritchard "cognitive enhancement" refers to the use of technology to augment or extend cognitive abilities (see CARTER, J. A.; PRITCHARD, D. The epistemology of cognitive enhancement. In: CLAUSEN, J.; LEVY, N. (Eds.). *The Routledge handbook of the philosophy of the extended mind*. London: Routledge, 2017, p. 316-326). The "smart arrow" and the LLM are both functional enhancements that take over a component of a task that would traditionally be performed by the agent.

²⁴ CARTER; PRITCHARD, The epistemology of cognitive enhancement.

²⁵ While we ultimately argue that mere prompting is an insufficient locus for authorial competence, the study of *prompt engineering* as a distinct skill is a rapidly emerging field in computer science and human-computer interaction.

²⁶ The specific form of luck at play here appears to be what Pritchard calls *intervening luck* (see PRITCHARD, *Epistemic luck*), where a lucky event intervenes between the exercise of an agent's competence and the successful outcome, thereby breaking the causal link and undermining the achievement.

that is playing against a grandmaster. The novice asks a spectator, who happens to be the world champion, ‘What’s a good move here?’ The master whispers a bright move, the novice executes it, and ultimately wins the match. This illustrates a classic problem in the *epistemology of testimony*: the norms that govern our trust in others as sources of information.²⁷ The move was adroit (it stemmed from the master’s competence), and accurate (it led to victory), but the novice’s performance in asking does not seem to count as apt. The success is attributable to the master’s advice, not to the novice’s ability to solicit it. Likewise, Dr. Evans’ success seems borrowed; it is not an achievement attributable to his competence.

This interlude therefore presents us with an unacceptable choice. The AAA model, applied to the compositional act of *distant writing*, suggests that either the human is not the author in the proper sense (Horn 1), or their authorship fails to constitute a reliable epistemic achievement (Horn 2). The problem is that both conclusions are deeply unsatisfying and undermines the intuition that a conscientious scientist like Dr. Evans actually accomplishes something epistemically valuable. The fault, then, lies not in the AAA model itself, but in our application. We attempt to locate the relevant competence in the wrong place: in direct sentence formation. This dilemma shows that this focus leads to a dead end. To safeguard a normative account of authorship, we must shift the focus of human compositional capacities to another set of higher-order epistemic practices. We now turn to the development of this alternative model.

3. From Composition to Precise Curation

The dilemma outlined in the previous section highlights the inadequacy of any model that locates the author’s competence solely in the act of composition. Such models inevitably fail because distant writing delegates precisely the *low-level* control of sentence formation to AI. The solution is not to abandon human agency, but to *redistribute* it. The primary epistemic work of the distant writer is not compositional, but *curatorial*. We propose the *Apt Curation Model*, which views authorship in the age of AI as the proper steering of a generative process. According to this account, a text produced through distant writing is a genuine epistemic achievement worthy of the human author if its success is due to the author’s adroit exercise of a set of higher-order curatorial virtues.

²⁷ See GOLDBERG, S. C. *Relying on others: An essay in epistemology*. Oxford: Oxford University Press, 2010. And, LACKEY, J. *Learning from words: Testimony as a source of knowledge*. Oxford: Oxford University Press, 2008.

These virtues are not monolithic, but rather a set of distinct but interconnected competencies that, successfully exercised, form the core of responsible and effective authorship. They are the central components of a *normative epistemology* for the AI age,²⁸ focused on the apt guidance of our cognitive behavior.²⁹ We can distinguish at least four central curatorial virtues: architectural, dialogic, evaluative, and integrative.

The first is the *architectural virtue*. This competence can be precisely described as being able to design everything relevant to compose the structure of the epistemic project's intellectual process. This includes preparing the research question,³⁰ carefully formulating the thesis, outlining the supporting argumentative steps, anticipating objections, and so forth. This is a typically human act of intellectual design and a form of *exploratory creativity*.³¹ The author acts as an architect, creating the blueprint for an argument. In this analogy, AI is a team of sophisticated but irrational construction workers. They can lay bricks, assemble components, and even suggest minor structural improvements, but they cannot design the building. The form, function, and coherence of the final structure are attributable to the architect's skill. Likewise, the conceptual coherence and argumentative integrity of an AI-assisted text are due to the author's architectural competence in defining the project's goals and structure.

Secondly, there is the *dialogical virtue*. This competence involves skillful interaction with the generative model, viewing the engagement not as a single instruction but as an iterative and critical conversation. This reflects the fact that human action is fundamentally situated and improvisational.³² Effective use of LLMs requires iterative feedback loops rather than one-shot prompting. An expert user does not simply accept the first response uncritically. Instead, they recognize the model's strengths and weaknesses, identify which aspects of the output require refinement, and strategically guide subsequent iterations toward improvement. This process resembles a conductor's work: the artistry lies not merely in assembling the orchestra and initiating performance, but in the detailed, iterative dialogue, refining

²⁸ A regulative epistemology, as defined by Roberts and Wood (2007), is concerned not just with analyzing concepts like knowledge or justification, but with providing guidance for the proper conduct of our cognitive lives.

²⁹ ROBERTS, R. C.; WOOD, W. J. *Intellectual virtues: An essay in regulative epistemology*. Oxford: Oxford University Press, 2007.

³⁰ The idea that scientific and philosophical progress begins with the clear formulation of problems or questions is a classic theme in the philosophy of science. See POPPER, K. R. *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge, 1963.

³¹ BODEN, M. A. *The creative mind: Myths and mechanisms*. 2. ed. London: Routledge, 2004.

³² Suchman's influential work critiques the model of human cognition as simple plan-following, arguing instead that our actions are fundamentally responsive to the immediate social and physical context in which they unfold. See SUCHMAN, L. A. *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press, 1987.

specific passages, adjusting dynamics, and shaping the collective output through skilled interaction. The dialogical virtue thus consists in the ability to engage in this productive, iterative refinement process.³³

Third, and perhaps most important, is the *evaluative virtue*. This virtue is essential for critical *epistemic filtering*. The results generated by a large language model (LLM) can be viewed as intellectual raw material, but their quality remains uncertain. This necessitates a careful approach characterized by what cognitive scientists refer to as *epistemic vigilance*.³⁴ AI output is often a mixture of precise summarization, surprising insights, subtle bias, and obvious invention (*hallucination*).³⁵ The evaluative virtue is the application of the author's specific expertise as a rigorous filter, a skill of utmost importance in the modern *age of misinformation*.³⁶ This involves carefully verifying claims, identifying and eliminating fabricated content, assessing the relevance of the arguments put forward to the overarching argument, and recognizing new insights. This is the skill of a museum curator examining a vast collection of artifacts. The quality and historical accuracy of the final display are paramount in this act. Similarly, is the epistemic integrity of an AI-assisted text attributable to the author's evaluative competence.

Finally, there is the *integrative virtue*, the virtue of synthesis. It involves merging AI-generated content with the author's own writing and ideas in a way that creates a cohesive and coherent whole. This integrative capacity is a core cognitive ability, central to our capacity to combine information from disparate sources into unified expressions.³⁷ Like a sculptor shaping raw stone into a statue, the author transforms rough AI output into a polished work through skilled refinement. The final text is not merely a compilation of ingredients but a new intellectual artifact that reflects the author's unique perspective and voice. Just as a chef's dish transcends its individual components through careful selection, combination, and presentation, an aptly curated text becomes genuinely attributable to its author through the exercise of integrative competence. The whole exceeds the sum of its parts, embodying the author's intellectual signature.

³³ This interaction can be understood through the Vygotskian concept of 'scaffolding', where a more competent agent provides the structure that enables a less competent agent (or a tool) to accomplish a task it could not perform alone. See WOOD, D.; BRUNER, J. S.; ROSS, G. The role of tutoring in problem solving, in *Journal of Child Psychology and Psychiatry*, v. 17, n. 2, 1976, p. 89-100.

³⁴ SPERBER, D. *et al.* Epistemic vigilance, in *Mind & Language*, v. 25, n. 4, 2010, p. 359-393.

³⁵ 'Hallucination' is the term of art used in AI research to describe instances where an LLM generates text that is nonsensical, factually incorrect, or untethered to the provided source material, yet presents it with a high degree of confidence. For a comprehensive overview, see JI, Z. *et al.* Survey of hallucination in natural language generation, in *ACM Computing Surveys*.

³⁶ O'CONNOR, C.; WEATHERALL, J. O. *The misinformation age: How false beliefs spread*. New Haven: Yale University Press, 2019.

³⁷ CARRUTHERS, P. The cognitive functions of language, in *Behavioral and Brain Sciences*, v. 25, n. 6, 2002, p. 657-674.

Under the *Apt Curation Model*, the dilemma from our previous interlude dissolves. Let us return to Dr. Evans and his paper on public reasons. The final text is accurate. It is also adroit, but the adroitness resides in the exercise of curatorial virtues. Its success is therefore apt, given that the text's accurateness and argumentative strength are *due to* its adroit curation. The project's success is attributable to its architectural, dialogical, evaluative, and integrative skill. The element of luck is not eliminated, but is duly managed and subordinated to competence, as in any complex intellectual endeavor. The author, distanced from the composition, reemerges as the central and creditable epistemic agent.

4. The Opacity Objection and the Problem of Reference

The *Apt Curation Model* appears to successfully reallocate the author's agency, grounding it in a set of higher-order managerial virtues. By doing so, it seems to save authorship from the dilemma of being either non-agential or merely lucky. However, this solution can be challenged by an objection that attacks the very nature of the raw material being curated. At the heart of this objection is the *opacity* of generative AI.³⁸ This is not simply due to technical illiteracy on the part of the user, but often an intrinsic characteristic of machine learning models themselves.³⁹ By focusing on the author's actions, the *Apt Curation Model* can ignore the epistemically suspect origin of the AI output. The problem, when viewed through this lens, is not simply one of luck, but of engaging in a form of *epistemically irresponsible deference* to an unreliable, non-rational source.

The term "opacity" in this context encompasses more than just the user's difficulty in grasping the intricate coding or neural structure of the model. It highlights a fundamental lack of clarity regarding how the model connects its inputs to its outputs. The method through which a large language model (LLM) produces text does not resemble logical reasoning nor does it employ a transparent causal framework. Instead, it is a probabilistic process focused on predicting the next likely token (which may be a word or part of a word) based on patterns learned from an extensive training dataset. The model lacks understanding of the statements it produces, possesses no beliefs or intentions, and has no access to factual knowledge as such.

³⁸ The *problem of opacity* is a central theme in the ethics and philosophy of AI, often discussed under the heading of *explainability* or *interpretability*. See WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual explanations without opening the black box: Automated decisions and the GDPR, in *Harvard Journal of Law & Technology*, v. 31, n. 2, 2017, p. 841-887.

³⁹ BURRELL, J. How the machine 'thinks': Understanding opacity in machine learning algorithms, in *Big Data & Society*, v. 3, n. 1, 2016, p. 1-13.

It remains a stochastic parrot⁴⁰ replicating patterns of human reasoning without grasping the underlying concepts.⁴¹

To illustrate, consider scientists' use of an electron microscope. It exemplifies a complex instrument based on *reliable* physical principles, which reinforces the confidence in the accuracy of the results it produces.⁴² In contrast, users of large language models (LLM) cannot have the same confidence in the accuracy of their content. Such language models are not able to prioritize the discovery of factual truths, but rather the creation of coherent and credible texts. Thus, relying on a large language model doesn't necessarily feel like trusting on a reliable scientific tool; it may be more like relying on the testimony of a source whose reliability is fundamentally questionable. In determining how one is justified in relying on testimony as a source of knowledge, we can identify multiple criteria. Recognizing the difference between the various possible sources is key to addressing the problem of epistemic standards.⁴³

Given this situation, we can restate the essential problem in the *epistemology of testimony*. A pertinent inquiry emerges: under what circumstances is it rational to adopt a belief based on the assertions of another source?⁴⁴ Such deference generally requires good reasons to *trust* the testimony.⁴⁵ The problem is that an *AI-generated output* meets none of the standard criteria. AIs are not competent experts, they cannot be truthful, and their reliability is notoriously inconsistent, producing confident "*hallucinations*" with no apparent logic.⁴⁶

⁴⁰ BENDER, E. M. *et al.* On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *Proceedings Of The 2021 Acm Conference On Fairness, Accountability, And Transparency*. [S.l.]: ACM, 2021, p. 610-623.

⁴¹ While the *stochastic parrot* metaphor is incisive, some philosophers of language argue that meaning and understanding can be emergent properties of complex systems, even if the system's components do not *understand* in a human-like way. In this paper, however, we proceed from the widely held view that LLMs currently lack genuine semantic understanding. See SEARLE, J. R. Minds, brains, and programs, in *Behavioral and Brain Sciences*, v. 3, n. 3, 1980, p. 417-424.

⁴² This view of scientific instruments aligns with the tradition of New Experimentalism, which argues that we can have rational belief in the entities revealed by instruments because we can manipulate them and understand their causal properties, even if their underlying theory is incomplete. See HACKING, I. *Representing and intervening*: Introductory topics in the philosophy of natural science. Cambridge: Cambridge University Press, 1983.

⁴³ ADLER, J. E. *Belief's own ethics*. Cambridge, MA: MIT Press, 2002. And, AUDI, R. The place of testimony in the fabric of knowledge and justification, in *American Philosophical Quarterly*, v. 34, n. 4, 1997, p. 405-422.

⁴⁴ COADY, C. A. J. *Testimony: A philosophical study*. Oxford: Clarendon Press, 1992.

⁴⁵ FAULKNER, P. *Knowledge on trust*. Oxford: Oxford University Press, 2011. Trust, in this context, is more than mere reliance on an object's regular behavior; it involves a normative expectation of goodwill or competence directed toward another agent. For a foundational account of trust as a three-part relation (A trusts B to do X), see BAIER, A. Trust and anti-trust, in *Ethics*, v. 96, n. 2, 1986, p. 231-260.

⁴⁶ JI, Z. *et al.* Survey of hallucination in natural language generation.

Consider the analogy of a *Confabulating Oracle*. This oracle possesses the ability to convey assertions that reflect profound, well-founded truths; yet it is equally proficient in crafting intricate and convincingly plausible fabrications. Importantly, there is no method to predict beforehand which category of statement it will emit. An agent who consults such an oracle and then presents its pronouncements as their own reasoned conclusions would be acting with *epistemic recklessness*.⁴⁷ The agent's actions would not be considered a form of virtuous curation, but rather an abdication of their epistemic responsibility.

The opacity objection claims that the distant writer finds himself in precisely this position. This charge directly attacks the curatorial virtues. First, it undermines the *evaluative virtue*, arguing that genuine evaluation is impossible. The author can only effectively test the AI's output against their preexisting knowledge. When the AI produces a genuinely novel claim, the author lacks independent grounds for its verification. Accepting such novel claims requires an act of *faith*. Furthermore, the criticism reframes the integrative virtue as a kind of *epistemic laundering*,⁴⁸ a mechanism that creates an unjustified facade of credibility for the AI's unfounded assertions.

At its core, the criticism posits that the *Apt Curation Model* is inadequate, as no amount of adept *post-hoc filtering* can redeem a source that is inherently epistemically flawed. The inherent opacity of the LLM suggests that relying on it is not a collaborative endeavor but rather a risky intellectual gamble. By emphasizing the author's curatorial skills, the model shifts the focus away from the core issue that the initial AI-generated content is constructed on unstable statistical foundations rather than on solid epistemic bases. The allegation thus shifts from accusations of mere chance to charges of an even more profound breach of epistemic responsibility.⁴⁹

5. Levels of Abstraction and Responsible Subordination

The opacity objection presents a meaningful dilemma, as it perceives distant writing not as a skillful compilation, but rather as an epistemically

⁴⁷ This charge of recklessness echoes the classic deontological position in the ethics of belief, which holds that it is "wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence". See CLIFFORD, W. K. The ethics of belief, in *The Contemporary Review*, 1977, v. 29, p. 289–309, here p. 309.

⁴⁸ 'Epistemic laundering', a term this paper introduces, is conceptually related to the problem of *responsibility laundering* in complex organizations, where the distribution of tasks across a hierarchy makes it difficult to assign blame or accountability to any single individual for a negative outcome. See BOVENS, M. *The quest for responsibility: Accountability and citizenship in complex organisations*. Cambridge: Cambridge University Press, 1998.

⁴⁹ This *profound breach* can be understood as a failure to adhere to the fundamental norms of assertion. An *assertion* is not merely a statement of a proposition but a social act that commits

negligent submission to a deceptive oracle. The accusation suggests that no degree of post-processing refinement can redeem a method whose roots are fundamentally detached from both reason and reality. Despite the validity of this objection, it hinges on a nuanced categorization mistake⁵⁰ concerning AI, leading to a flawed understanding of the essential role of curatorial virtues. A strong rebuttal can be established from two perspectives: firstly, by utilizing Floridi's analytical framework of *Levels of Abstraction (LoA)*⁵¹ to appropriately classify AI as a *tool* instead of a *source of testimony*; and secondly, by illustrating that curatorial virtues embody precisely the practice of responsible epistemic deference in this framework.

The first move in countering the objection involves assessing the system at the appropriate Level of Abstraction (LoA). An LoA, in Floridi's sense, is a specific, well-defined model of a system, consisting of a set of observables or variables chosen to analyze the system for a particular purpose.⁵² The choice of LoA is not arbitrary; it is determined by the nature of the study. The opacity objection commits a category error by analyzing the LLM at an inappropriate LoA. This is analogous to adopting the wrong *stance* toward a system; the objection takes an *intentional stance* and treats the AI as an agent with beliefs,⁵³ when it should take a *design stance* and treat it as a complex artifact.⁵⁴

The objection analyzes the LLM at the LoA of cognitive science and asks, 'Does it make reliable statements?' At this LoA, the LLM fails. It is a stochastic parrot, not a rational agent. However, this is the wrong level to analyze the practice of distant writing. The author interacts with the AI as a complex instrument. The corresponding LoA is therefore the

the speaker to the truth of that proposition, for which they can be held accountable. See WILLIAMSON, T. *Knowledge and its limits*. Oxford: Oxford University Press, 2000.

⁵⁰ The charge that the objection makes a *categorization mistake* echoes Gilbert Ryle's famous philosophical diagnosis. A category mistake is the error of assigning a property to a thing that could not possibly have that property (e.g., describing a university as a separate building alongside its colleges and libraries). See RYLE, G. *The concept of mind*. Chicago: University of Chicago Press, 1949.

⁵¹ Floridi's method of LoA is a powerful tool for analyzing complex informational systems. It bears a family resemblance to other multi-level analytical frameworks, such as David Marr's influential tripartite model for analyzing cognitive systems, which distinguishes between the computational, algorithmic/representational, and implementational levels. See MARR, D. *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press, 1982.

⁵² FLORIDI, L. The method of levels of abstraction, in *Minds and Machines*, v. 18, n. 3, 2008, p. 303-329.

⁵³ While Dennett argues for an instrumentalist interpretation of the *intentional stance* (i.e., it is a useful predictive tool), other philosophers have argued for a more realist interpretation. In this paper, our argument is compatible with Dennett's instrumentalism, as it treats the ascription of beliefs to the AI as a methodological error for the task at hand. For a classic realist critique of Dennett, see FODOR, J. Fodor's guide to mental representation: The intelligent auntie's vade-mecum, in *Mind*, v. 94, n. 373, 1985, p. 76-100.

⁵⁴ See DENNETT, D. C. *The intentional stance*. Cambridge, MA: MIT Press, 1987.

instrumental or functional level. At this LoA, the key question is not, ‘Is the source reliable?’ but ‘What procedures can be used to reliably produce epistemically valuable results with this instrument?’

To illustrate, let’s consider the analogy of a seismologist who is employing an advanced computer simulation in order to predict earthquake probabilities. This simulation is informed by a set of complex algorithms and extensive datasets. Analyzing this simulation through the lens of cognitive science, particularly at the *Level of Abstraction* (LoA), and questioning whether it “believes” an earthquake is imminent would be a misguided inquiry. Instead, the seismologist operates at an instrumental LoA, where her proficiency lies in her adeptness at utilizing and interpreting this tool, representing a form of ‘thing knowledge’ distinct from mere *propositional knowledge*.⁵⁵ Her expertise encompasses the design of the model’s parameters, the interpretation of its probabilistic output, and the validation of this information against other geological evidence. Thus, her conclusions regarding seismic risk do not stem from mere reliance on the simulation’s output, but rather reflect a well-informed judgment derived from her skilled application of this sophisticated instrument.

This parallel also applies to the distant writer; the opacity of the underlying mechanisms of the language model does not negate the author’s capability, just as the seismologist’s reliance on the *opaque* code of the simulation does not diminish her agency. In both scenarios, competence is demonstrated through the effective management of an instrumental process. If AI is properly viewed as a tool rather than as faulty testimony, we can turn to the second line of defense. The objection claims that trust in AI’s results is an act of irresponsible deference. This is only true if the deference is blind. However, the *Apt Curation Model* explains precisely what makes such deference responsible and transforms it into a form of *earned trust*.⁵⁶ *Curatorial virtues* are the practices that transform a raw, unreliable result into a responsibly validated claim.

Blindly copying and pasting text from an LLM is *epistemically vicious*.⁵⁷ The *evaluative virtue*, on the other hand, is the mechanism of critical deference.

⁵⁵ BAIRD, D. *Thing knowledge: A philosophy of scientific instruments*. Berkeley: University of California Press, 2004. The emphasis on knowledge gained through active, instrumental engagement has deep roots in the American pragmatist tradition, which views knowledge not as a passive representation of reality but as a product of problem-solving activity. See DEWEY, J. *The quest for certainty: A study of the relation of knowledge and action*. New York: Minton, Balch & Company, 1929.

⁵⁶ FRICKER, M. *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press, 2007.

⁵⁷ An epistemic vice, in contrast to an intellectual virtue, is a cognitive character trait, attitude, or way of thinking that systematically obstructs the gaining, keeping, or sharing of knowledge. Blind deference to an unreliable source would be a paradigmatic example. For a taxonomy of such vices, see BATTALY, H. Varieties of epistemic vice. In: MATHESON, J.; VITZ, R. (Eds.). *The ethics of belief: Individual and social*. Oxford: Oxford University Press, 2014, p. 51-76.

The objection mistakenly assumes that evaluation is limited to testing claims against pre-existing knowledge. When faced with a novel claim, evaluative virtue *requires* a rigorous process of independent verification,⁵⁸ thus fulfilling the epistemic duties of the listener in the exchange of testimony.⁵⁹ The author has to be careful; they must treat the new claim not as testimony to be accepted, but as a hypothesis to be tested. Only when a claim survives this external validation can it be responsibly endorsed. Deference becomes a practice guided by *epistemic vigilance*.⁶⁰

Furthermore, the *integrative virtue* is not, as the objection suggests, a form of epistemic laundering. It is the final act of assuming full epistemic responsibility. By rewriting, reframing, and embedding vetted claims into their own argumentative structure, the author explicitly puts their own epistemic authority⁶¹ behind it. This process establishes their accountability. The act of integration is the public declaration that the author, not the AI, is the ultimate warrant for the text's assertions.

The opacity objection thus fails. It mischaracterizes the LLM by analyzing it at the wrong LoA, and it underestimates the normative force of the curatorial practices. The distant writer is not a passive recipient of testimony from a Confabulating Oracle. They are the active, responsible director of an epistemic process, using a powerful generative tool to produce a text for which they, and they alone, are ultimately accountable.

6. Implications and Further Questions

The defense of the *Apt Curation Model* is not merely a solution to a philosophical dilemma; it provides a normative framework with significant implications for how we understand and regulate intellectual production in an era of generative AI. By shifting the focus from authorial writing skills to content selection and organization, this method changes our understanding of originality and responsibility. It asks for a fundamental

⁵⁸ This process of independent verification is analogous to the concept of 'severe testing' in the philosophy of science, where a hypothesis is only said to be well-corroborated if it survives a test that it would have been very likely to fail if it were false. The author's evaluative virtue functions as a severe test for the AI's novel claims. See MAYO, D. G. *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press, 1996.

⁵⁹ MCMYLER, B. *Testimony, truth, and authority*. Oxford: Oxford University Press, 2011.

⁶⁰ ORIGGI, G. Epistemic vigilance and epistemic deference: The role of trust in the production and in the transmission of knowledge, in *Episteme*, v. 9, n. 1, 2012, p. 1-13.

⁶¹ Epistemic authority is the legitimate power to transmit knowledge or justified belief to others through testimony or other forms of address. By integrating a claim, the author exercises their own authority, inviting the reader to trust them, not the AI. For a comprehensive account, see ZAGZEBSKI, *Epistemic authority*.

shift in our teaching methods. Furthermore, it potentially opens a world of new perspectives in philosophical debate.

One of the main concerns surrounding distant writing is that it signals the end of human originality. The *Apt Curation Model* responds to this by shifting the locus of originality from the Romantic ideal of *creatio ex nihilo*,⁶² a conception of the author as a solitary genius that has been strongly contested for decades.⁶³ Originality is not lost; it is relocated toward a more architectural and integrative conception. From this perspective, originality manifests itself in several key areas, such as:

Architectural Originality: The novelty of the conceptual design.

Dialogic Originality: The skillfulness demonstrated in iterative engagement with AI.

Integrative Originality: The unique authorial voice and vision that emerge from the synthesis of AI-generated content with the author's own insights.

This is analogous to the originality of the film director, conceptualized in *auteur theory*⁶⁴ as the primary author of a film.⁶⁵ A director is usually seen as the *author* of a movie, even though she did not write every line of dialogue. It is thought that her originality exists in her unique vision and the masterful integration of others' contributions into a unified work of art. The distant writer, like the film director or a virtuous curator, is original in exactly the same way.

The model thus provides a clear basis for assigning full epistemic responsibility⁶⁶ to the human author. The conscious act of curation is simultaneously an act of taking responsibility. This framework offers a clear path to maintaining academic integrity. The operational question should not

⁶² Latin for 'creation out of nothing'. This phrase is associated with a theological view of divine creation but was adopted by Romantic-era thinkers to describe the artist as a quasi-divine, solitary genius whose work originates entirely from their own imagination, independent of tradition or influence. For a critical history of this conception of authorship, see WOODMANSEE, M. *The author, art, and the market: Rereading the history of aesthetics*. New York: Columbia University Press, 1994.

⁶³ BARTHES, R. The death of the author. In: BARTHES, R. *Image-music-text*. New York: Hill and Wang, 1977. (Obra original publicada em 1967).

⁶⁴ It is worth noting that the auteur theory, while influential, is not without its critics. Most famously, the film critic Pauline Kael argued that it often ignores the deeply collaborative nature of filmmaking and can lead to the overvaluation of a director's stylistic tics at the expense of a film's overall quality. The analogy is used here to emphasize the director's integrative vision, not to dismiss the contributions of other participants. See KAEI, P. Circles and squares, in *Film Quarterly*, v. 16, n. 3, 1963, p. 12-26.

⁶⁵ SARRIS, A. Notes on the Auteur Theory in 1962, in *Film Culture*, n. 27, 1962, p. 1-8.

⁶⁶ Epistemic responsibility, in this context, refers to the duty of an agent to govern their cognitive conduct in a way that is truth-conducive and justifiable to others. It is a central concept in normative epistemology. For a foundational, book-length treatment, see CODE, L. *Epistemic responsibility*. Providence, RI: Brown University Press, 1987.

be ‘Did you use AI?’ but rather ‘Did you perform your curatorial duties responsibly?’

This has profound pedagogical implications. Current educational strategies focused solely on AI detection are ultimately unsustainable. If the *Apt Curation Model* is correct, the goal of education should be to explicitly teach the virtues of curation as a form of advanced *digital literacy*.⁶⁷ The 21st century curriculum should focus on developing students’ architectural, dialogic, evaluative, and integrative skills, while critically engaging with the challenges and limitations of educational technology by promoting responsible use rather than avoidance.⁶⁸

However, this approach raises further questions. First, there is the question of *dynamic development*. What happens as AI systems become more reliable? Will the weight of the evaluative virtue diminish, once again transforming the nature of authorship? Second, the question of domain specificity remains. Does the model apply uniformly to scientific and creative writing? The relative importance of the curatorial virtues may vary significantly across domains.

Finally, there are complex legal ramifications. Legal concepts like *copyright*⁶⁹ are customarily associated with original works of authorship.⁷⁰ The Apt Curation Model distinguishes the architectural agent from the compositional process. How this distinction applies to current copyright law remains uncertain, a situation recently described by some scholars as a copyright crisis.⁷¹ Can a curator claim copyright over sentences they did not write? This question demands exploration as distant writing becomes commonplace.

Conclusion

The rise of generative artificial intelligence and its facilitation of *distant writing* has created a significant conceptual challenge to traditional notions of authorship. Throughout the paper, we have explored how the

⁶⁷ BAWDEN, D. Origins and concepts of digital literacy. In: LANKSHEAR, C.; KNOBEL, M. (Eds.). *Digital literacies: Concepts, policies and practices*. New York: Peter Lang, 2008, p. 17-32.

⁶⁸ SELWYN, N. *Is technology good for education?* Cambridge: Polity Press, 2016.

⁶⁹ U.S. copyright law, for example, protects “original works of authorship fixed in any tangible medium of expression” but explicitly states that it does not extend to “any idea, procedure, process, system, method of operation, concept, principle, or discovery”. The core of the legal challenge is whether the curatorial virtues, which deal with ideas, processes, and concepts, can meet the threshold for copyright protection, which is traditionally focused on the final “expression”. See 17 U.S.C. § 102.

⁷⁰ GINSBURG, J. C.; BUDIARDJO, L. A. Authors and machines, in *Berkeley Technology Law Journal*, v. 34, n. 1, 2019, p. 343-410.

⁷¹ LEMLEY, M. A. How to think about the copyright crisis, in *Journal of the Copyright Society of the U.S.A.*, v. 71, 2024, p. 1-28.

introduction of a semi-autonomous generative partner gives rise to a complex control dilemma, obscuring attribution and jeopardizing the idea of a human author as the primary epistemic agent accountable for a text. The Socratic discussion presented initially highlighted the depth of this issue. When evaluated through Sosa's *virtue epistemology*, efforts to pinpoint authorial competence in the traditional composition process led to an intractable conflict: we either assign the key skill of composition (the adroitness) to the machine, which renders the human contribution *non-apt*, or we diminish the human's role to mere *epistemic luck*, failing to establish a credible claim to achievement. This analysis demonstrated that existing models, which emphasize direct control over sentence generation, are fundamentally unsuitable for addressing the challenges posed by this new technological landscape.

In light of this, the paper proposes and defends the *Apt Curation Model*, a normative framework designed to navigate this dilemma by shifting the focus of human intellectual labor. According to this model, authorship is not eliminated but rather redefined; it evolves from a craft of composition into an art of management. The author's expertise is now seen in the adept facilitation of various higher-order *curatorial skills*. The author functions as an architect, creating the conceptual foundation and argumentative framework; as a *dialogical partner*, directing the interaction between human and machine through an iterative and critical process; as a curator, applying expert judgment and epistemic diligence to evaluate the raw outputs for accuracy and relevance; and finally, as an integrator, integrating the vetted content with personal insights to create a coherent and unified intellectual work. This combination of virtues defines the author's competence. Consequently, the final text is deemed an apt accomplishment, as its epistemic success can be clearly attributed to this diverse array of human skills.

Moreover, this model effectively addresses the significant counterargument that reliance on opaque AI constitutes irresponsible epistemic deference. The defense proceeds on two fronts. First, by analyzing AI at the appropriate *Level of Abstraction*, we reconceptualize it not as a testimonial source requiring trust, but as a sophisticated instrument requiring skilled operation. The seismologist does not trust her simulation's beliefs; she competently interprets its outputs. Similarly, the distant writer does not defer to the AI's testimony but rather skillfully directs and validates its generative processes. Second, far from being mere epistemic window-dressing, the *curatorial virtues* constitute precisely the responsible practices that transform reliance on AI from *blind deference* into *justified trust*. The rigorous processes of verification, evaluation, and integration ensure that the author's acceptance of AI-generated content is neither careless nor opportunistic. Rather than passive acceptance, the model demands critical scrutiny at every stage, transforming potentially unreliable raw material

into responsibly validated knowledge claims. It is this combination of instrumental competence and epistemic vigilance that distinguishes apt curation from mere confabulation laundering.

In conclusion, the *Apt Curation Model* serves as more than a theoretical answer; it charts a path forward. It equips us with terminology to differentiate between responsible users and those who are negligent, thereby establishing a new basis for academic integrity standards and advocating for a pedagogical shift that emphasizes the development of these critical *intellectual virtues* over mere prohibition. These reflections sketch a portrait not of a diminished author, but rather of a transformed one, whose responsibilities multiplied. The author would become the ultimate validator of the text's identity and the primary responsible agent. Applying this reasoning, we conclude that fears of replacement are misguided: this represents not a threat but a transformation. Although our tools may evolve, the essence of human *intellectual agency* persists through adaptation. There is no need to consider generative AIs as marking the end of authorship; Instead, they have always relied on distinctively human qualities: purposeful intent, deep understanding, thoughtful synthesis, and the willingness to accept responsibility for our claims. This concept underscores that our most important contributions to knowledge creation have always gone beyond mere automated text generation. Instead, they rely on typically human qualities such as purposeful intent, deep understanding, thoughtful synthesis, and the determined willingness to take responsibility for our statements.

References

- ABRAMS, M. H. *The mirror and the lamp: Romantic theory and the critical tradition*. Oxford: Oxford University Press, 1953.
- ADLER, J. E. *Belief's own ethics*. Cambridge, MA: MIT Press, 2002.
- ANSCOMBE, G. E. M. *Intention*. Oxford: Basil Blackwell, 1957.
- AUDI, R. The place of testimony in the fabric of knowledge and justification, in *American Philosophical Quarterly*, v. 34, n. 4, 1997, p. 405-422.
- AXTELL, G. Recent work on virtue epistemology, in *Philosophical Quarterly*, v. 34, n. 1, 1997, p. 1-26.
- BAIER, A. Trust and antitrust, in *Ethics*, v. 96, n. 2, 1986, p. 231-260.
- BAIRD, D. *Thing knowledge: A philosophy of scientific instruments*. Berkeley: University of California Press, 2004.
- BARTHES, R. The death of the author. In: _____. *Image-music-text*. Tradução de S. Heath. New York: Hill and Wang, 1977. (Obra original publicada em 1967).

- BATTALY, H. Varieties of epistemic vice. In: MATHESON, J.; VITZ, R. (Eds.). *The ethics of belief: Individual and social*. Oxford: Oxford University Press, 2014, p. 51-76.
- BAWDEN, D. Origins and concepts of digital literacy. In: LANKSHEAR, C.; KNOBEL, M. (Eds.). *Digital literacies: Concepts, policies and practices*. New York: Peter Lang, 2008, p. 17-32.
- BENDER, E. M. *et al.* On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *Proceedings Of The 2021 Acm Conference On Fairness, Accountability, And Transparency*. [S.l.]: ACM, 2021, p. 610-623.
- BODEN, M. A. *The creative mind: Myths and mechanisms*. 2. ed. London: Routledge, 2004.
- BOVENS, M. *The quest for responsibility: Accountability and citizenship in complex organisations*. Cambridge: Cambridge University Press, 1998.
- BURRELL, J. How the machine ‘thinks’: Understanding opacity in machine learning algorithms, in *Big Data & Society*, v. 3, n. 1, 2016, p. 1-13.
- CARRUTHERS, P. The cognitive functions of language, in *Behavioral and Brain Sciences*, v. 25, n. 6, 2002, p. 657-674.
- CARTER, J. A.; PRITCHARD, D. The epistemology of cognitive enhancement. In: CLAUSEN, J.; LEVY, N. (Eds.). *The Routledge handbook of the philosophy of the extended mind*. London: Routledge, 2017, p. 316-326.
- CLARK, A.; CHALMERS, D. The extended mind, in *Analysis*, v. 58, n. 1, 1998, p. 7-19.
- CLIFFORD, W. K. The ethics of belief, in *The Contemporary Review*, 1977, v. 29, p. 289-309.
- COADY, C. A. J. *Testimony: A philosophical study*. Oxford: Clarendon Press, 1992.
- CODE, L. *Epistemic responsibility*. Providence, RI: Brown University Press, 1987.
- DENNETT, D. C. *The intentional stance*. Cambridge, MA: MIT Press, 1987.
- DEWEY, J. *The quest for certainty: A study of the relation of knowledge and action*. New York: Minton, Balch & Company, 1929.
- FAULKNER, P. *Knowledge on trust*. Oxford: Oxford University Press, 2011.
- FLORIDI, L. The method of levels of abstraction, in *Minds and Machines*, v. 18, n. 3, 2008, p. 303-329.
- FLORIDI, L. *Distant Writing: Literary Production in the Age of Artificial Intelligence*. Revised version 5. [S.l.: s.n.], 2025.
- FODOR, J. Fodor’s guide to mental representation: The intelligent auntie’s vade-mecum, in *Mind*, v. 94, n. 373, 1985, p. 76-100.
- FRICKER, M. *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press, 2007.
- GETTIER, E. L. Is justified true belief knowledge?, in *Analysis*, v. 23, n. 6, 1963 p. 121-123.
- GINSBURG, J. C.; BUDIARDJO, L. A. Authors and machines, in *Berkeley Technology Law Journal*, v. 34, n. 1, 2019, p. 343-410.

- GOLDBERG, S. C. *Relying on others: An essay in epistemology*. Oxford: Oxford University Press, 2010.
- GOLDMAN, A. I. *Knowledge in a social world*. Oxford: Clarendon Press, 1999.
- GOODMAN, N. *Ways of worldmaking*. Indianapolis: Hackett Publishing, 1978.
- GRECO, J. *Achieving knowledge: A virtue-theoretic account of epistemic normativity*. Cambridge: Cambridge University Press, 2010.
- HACKING, I. *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press, 1983.
- HUTCHINS, E. *Cognition in the wild*. Cambridge, MA: MIT Press, 1995.
- Ji, Z. *et al.* Survey of hallucination in natural language generation, in *ACM Computing Surveys*, v. 55, n. 12, 2023, p. 1-38.
- KAEL, P. Circles and squares, in *Film Quarterly*, v. 16, n. 3, 1963, p. 12-26.
- KALLESTRUP, J. Extended knowledge and the value of cognitive integration, in *Philosophical Issues*, v. 30, n. 1, 2020, p. 169-185.
- KANT, I. An answer to the question: What is enlightenment? In: _____. *Practical philosophy*. Editado e Traduzido por M. J. Gregor. Cambridge: Cambridge University Press, 1999, p. 11-22. (Obra original publicada em 1784).
- KELP, C. *Good thinking: A defense of virtue epistemology*. Oxford: Oxford University Press, 2021.
- LACKEY, J. *Learning from words: Testimony as a source of knowledge*. Oxford: Oxford University Press, 2008.
- LEMLEY, M. A. How to think about the copyright crisis, in *Journal of the Copyright Society of the U.S.A.*, v. 71, 2024, p. 1-28.
- MARR, D. *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press, 1982.
- MAYO, D. G. *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press, 1996.
- MCMYLER, B. *Testimony, truth, and authority*. Oxford: Oxford University Press, 2011.
- MILLER, L. C. AI and the epistemic responsibilities of the intellectual, in *Philosophy & Technology*, v. 37, n. 1, 2024, Art. 20.
- MORETTI, F. Conjectures on world literature, in *New Left Review*, n. 1, 2000, p. 54-68.
- NYHOLM, S. *The ethics of AI ethics: An introduction and critique*. Oxford: Oxford University Press, 2023.
- O'CONNOR, C.; WEATHERALL, J. O. *The misinformation age: How false beliefs spread*. New Haven: Yale University Press, 2019.
- ORIGGI, G. Epistemic vigilance and epistemic deference: The role of trust in the production and in the transmission of knowledge, in *Episteme*, v. 9, n. 1, 2012, p. 1-13.
- PASQUALE, F. *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press, 2015.

- PETTIT, P. Responsibility incorporated, in *Ethics*, v. 117, n. 2, 2007, p. 171-201.
- POPPER, K. R. *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge, 1963.
- PRITCHARD, D. *Epistemic luck*. Oxford: Oxford University Press, 2005.
- ROBERTS, R. C.; WOOD, W. J. *Intellectual virtues: An essay in regulative epistemology*. Oxford: Oxford University Press, 2007.
- RYLE, G. *The concept of mind*. Chicago: University of Chicago Press, 1949.
- SARRIS, A. Notes on the Auteur Theory in 1962, in *Film Culture*, n. 27, 1962, p. 1-8.
- SEARLE, J. R. Minds, brains, and programs, in *Behavioral and Brain Sciences*, v. 3, n. 3, 1980, p. 417-424.
- SELWYN, N. *Is technology good for education?* Cambridge: Polity Press, 2016.
- SOSA, E. *A virtue epistemology: Apt belief and reflective knowledge*. Oxford: Oxford University Press, 2007. v. 1.
- SOSA, E. *Reflective knowledge: Apt belief and reflective knowledge*. Oxford: Oxford University Press, 2009. v. 2.
- SPERBER, D. *et al.* Epistemic vigilance, in *Mind & Language*, v. 25, n. 4, 2010, p. 359-393.
- SUCHMAN, L. A. *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press, 1987.
- THORSTAD, D. The epistemic threat of lazy expertise, in *Episteme*, v. 20, n. 4, 2023, p. 812-828.
- VERBEEK, P. P. *What things do: Philosophical reflections on technology, agency, and design*. University Park, PA: Penn State Press, 2005.
- WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual explanations without opening the black box: Automated decisions and the GDPR, in *Harvard Journal of Law & Technology*, v. 31, n. 2, 2017, p. 841-887.
- WILLIAMSON, T. *Knowledge and its limits*. Oxford: Oxford University Press, 2000.
- WOOD, D.; BRUNER, J. S.; ROSS, G. The role of tutoring in problem solving, in *Journal of Child Psychology and Psychiatry*, v. 17, n. 2, 1976, p. 89-100.
- WOODMANSEE, M. *The author, art, and the market: Rereading the history of aesthetics*. New York: Columbia University Press, 1994.
- ZAGZEBSKI, L. T. *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge: Cambridge University Press, 1996.
- ZAGZEBSKI, L. T. *Epistemic authority: A theory of trust, authority, and autonomy in belief*. Oxford: Oxford University Press, 2012.