Araştırma Makalesi
Başvuru: 17.03.2026
Kabul: 24.04.2026
Atıf: Eşsiz, Dursun. "Academic AI Ethics: The Normative Limits of Generative Artificial Intelligence in Higher Education," *Temaşa Felsefe Dergisi* sayı: 25 (Haziran 2026): 200-227. https://doi.org/10.55256/temasa.1911850

Academic AI Ethics: The Normative Limits of Generative Artificial Intelligence in Higher Education

Dursun Eşsiz[1]

ORCID: 0000-0002-9767-2780

DOI: 10.55256/TEMASA.1911850

Abstract

This study treats the use of generative artificial intelligence in higher education not merely as a problem of technical efficiency, convenience, or regulation, but as a normative issue concerning the academic subject's capacity to think, the quality of intellectual labor, its relation to truth, and institutional responsibility. The article's central claim is that the legitimacy of AI in academia should not be assessed through reactive responses to isolated crisis topics, but within an integrated philosophical model. To this end, the study first examines the current literature on AI in higher education along the axes of academic integrity, authorship, assessment, pedagogical design, research, and governance; it then discusses the philosophical gap that becomes visible within this literature. It subsequently reinterprets the normative possibilities offered by Kant, Aristotle, Plato, and Descartes as the classical philosophical foundations of debates on academic AI, and extends this framework onto a contemporary plane through phenomenology, Heidegger, and approaches to digital habitus. In the final stage, the theoretical model developed is applied to areas of academic practice to show under what conditions generative AI becomes a legitimate tool that supports thinking and under what conditions it becomes an ethical problem that erodes autonomy, intellectual character, and an orientation toward truth. In this way, the article proposes an integrated normative model for academic AI ethics.

Keywords: Generative Artificial Intelligence, Academic Ethics, Higher Education, Artificial Intelligence Ethics, Intellectual Autonomy, Normative Ethics, Philosophy of Technology.

Akademik Yapay Zekâ Etiği: Yükseköğretimde Üretken Yapay Zekânın Normatif Sınırları

Öz

Bu çalışma, üretken yapay zekânın yükseköğretimdeki kullanımını yalnızca teknik verimlilik, kolaylık veya düzenleme sorunu olarak değil, akademik öznenin düşünme yetisi, entelektüel emeğin niteliği, hakikatle ilişkisi ve kurumsal sorumluluğu açısından normatif bir mesele olarak ele almaktadır. Makalenin temel iddiası, yapay zekânın akademideki meşruiyetinin tek tek kriz başlıklarına verilen tepkilerle değil, bütünlüklü bir felsefi model içinde değerlendirilmesi gerektiğidir. Bu amaçla çalışma, öncelikle yükseköğretimde yapay zekâya ilişkin güncel literatürü akademik dürüstlük, yazarlık, ölçme-değerlendirme, pedagojik tasarım, araştırma ve yönetişim eksenlerinde incelemekte; ardından bu literatürde belirginleşen felsefi boşluğu tartışmaktadır. Devamında Kant, Aristoteles, Platon ve Descartes'ın sunduğu normatif imkânlar, akademik yapay zekâ tartışmasının klasik felsefi temelleri olarak yeniden yorumlanmakta; fenomenoloji, Heidegger ve dijital habitus yaklaşımlarıyla bu çerçeve çağdaş bir düzlemde genişletilmektedir. Son aşamada geliştirilen teorik model, akademik uygulama alanlarına uyarlanarak üretken yapay zekânın hangi koşullarda düşünmeyi destekleyen meşru bir araç, hangi koşullarda ise özerkliği, entelektüel karakteri ve hakikat yönelimini aşındıran etik bir sorun haline geldiği ortaya konmaktadır. Böylece makale, akademik yapay zekâ etiği için birleşik bir normatif model önermektedir.

Anahtar Kelimeler: Üretken Yapay Zekâ, Akademik Etik, Yükseköğretim, Yapay Zekâ Etiği, Entelektüel Özerklik, Normatif Etik, Teknoloji Felsefesi.

[1] Assistant Professor Dr., Erciyes University, Faculty of Letters, Department of Korean Language and Literature. dessiz@erciyes.edu.tr

Introduction

The rapid and deep penetration of generative artificial intelligence systems, especially large language models, into higher education has reopened debate about how universities produce knowledge, teach, assess, and build public trust. At first, this transformation was largely treated as a technical innovation, and benefits such as faster text production, personalized learning support, assistance in the writing process, research efficiency, and administrative convenience were emphasized. Yet it quickly became clear that the problem is not simply that a new tool has entered the campus. The real issue is how this tool relates to the academic subject's labor, autonomy, claim to truth, and the university's educational telos. For this reason, the debate on AI extends beyond a calculation of instrumental benefits and risks and becomes tied to more fundamental questions about what academic life is.[2]

This is precisely the point of departure of the present article. Although the existing educational and ethical literature on AI makes extremely important observations, it often employs a scattered normative language. For example, when one says that "the student should develop their own thinking," one is in fact speaking of autonomy; when one says that "we must cultivate good persons and good researchers," one is speaking of character and virtue; when one says that "a text's appearing correct is not the same as its being grounded in genuine knowledge," one is speaking of the distinction between truth and appearance; and when one says that "the student should use their own reason," one is speaking of the independence of critical judgment. In other words, philosophy is not absent from this literature; on the contrary, it is often present in unnamed, fragmentary, and implicit form. The problem is that these latent philosophical foundations have not been systematized.[3]

Within this framework, the article's central thesis may be stated as follows: the legitimacy of generative AI in academia cannot be assessed by asking how fast, persuasive, or useful it is. It must be assessed by asking four prior questions. Does a given form of AI use preserve the academic subject's autonomous judgment, or does it quietly displace it? Does it strengthen the intellectual character - the habits, dispositions, and practical wisdom - through which genuine scholarly formation occurs, or does it erode them? Does it keep the difference between grounded knowledge and its fluent appearance visible, or does it normalize that confusion? And does it leave methodical self-possession - the subject's ownership of its own cognitive process - intact, or does it progressively assume that process on the subject's behalf?

These four questions are not derived from the current policy debate; they are derived from the philosophical tradition. The first is Kant's question, the second Aristotle's, the third Plato's, and the fourth Descartes's. Together, they constitute a normative framework that the article develops across four classical axes and extends, through phenomenology and Heidegger, to address the structural conditions that shape the academic subject before any individual choice is made. The AI debate cannot be reduced to the binary of permission or prohibition. The real issue is which forms of use are compatible with the university's irreducible functions: the cultivation of thinking persons,

[2] N. J. Francis, S. Jones and D. P. Smith, "Generative AI in Higher Education: Balancing Innovation and Integrity," *British Journal of Biomedical Science* (2025). https://doi.org/10.1080/09674845.2025.2474477; R. Michel-Villarreal, E. L. Vilalta-Perdomo, D. E. Salinas-Navarro, R. Thierry-Aguilera and F. S. Gerardou, "Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT," *Education Sciences* 13, (2023): 856. https://doi.org/10.3390/educsci13090856; J. Dempere, K. P. Modugu, A. Hesham and L. K. Ramasamy, "The Impact of ChatGPT on Higher Education," *Frontiers in Education* (2023). https://doi.org/10.3389/feduc.2023.1206936

[3] S. M. Qadhi, A. Alduais, Y. Chaaban and M. Khraisheh, "Generative AI, Research Ethics, and Higher Education Research: Insights from a Scientometric Analysis," *Information* 15, (2024): 325. https://doi.org/10.3390/info15060325; K. Bittle and O. El-Gayar, "Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda," *Information* 16, (2025): 296. https://doi.org/10.3390/info16040296; O. P. Barus, A. N. Hidayanto and I. Eitiveni, "Mapping Generative AI's Ethical Issues in Higher Education: A Felt-Guided Systematic Review," *Polyglot: Jurnal Ilmiah* (2025). https://doi.org/10.19166/pji.v21i2.10020

the production of genuine knowledge, and the assumption of public epistemic responsibility.[4]

The literature on generative artificial intelligence in higher education has expanded at an extraordinary pace since 2023 and, within a short time, has separated into several dominant clusters. The first cluster has formed around opportunities and potential use cases. In this literature, the functions of AI, such as learning support, personalization, linguistic assistance, idea generation, draft preparation, feedback provision, and the acceleration of research processes, are emphasized. Particularly in early discussions, AI was often depicted as an innovative tool that increases university efficiency and offers accessible support to students. The strength of this approach lies in its recognition of the pedagogical possibilities of technology; its weakness lies in its often failing to ask sufficiently what conception of the human and of the university these possibilities serve.[5]

The second cluster has taken shape around academic integrity and plagiarism. Studies in this area show that generative AI's effectiveness in producing homework assignments, essays, reports, code, and exam answers radically challenges the classical understanding of the "student product". A substantial portion of instructors emphasize that the core difficulty here is not merely the increased risk of cheating, but also the growing difficulty of recognizing student labor, originality, and assessment validity. Systematic reviews likewise show that this concern is not incidental; the problem of academic integrity has become an almost constitutive theme in the literature. Yet the same studies also show that an approach based solely on prohibition and detection strategies is both practically inadequate and pedagogically indefensible.[6]

The third cluster is organized around authorship, contribution, disclosure obligations, and academic publishing. With the spread of AI-assisted text generation, the questions "who is the author?", "Who bears responsibility for the text?" and "Where does the boundary between assistance and substitution begin?" have become central. Although a significant portion of debates in this area agree that AI cannot be listed as an author, the more difficult issue is how to define the degree of the human author's contribution. As AI's role in research design, literature review, data interpretation, text construction, and source selection grows, the legal and ethical meaning of authorship must be reconsidered. This literature draws attention not only to the risks of fake references and hallucinations, but also to the danger that intellectual labor itself may become invisible.[7]

---

[4] M. I. G. de Pinho, A. P. D. Costa and C. G. de Pinho, "Ethical and Responsible Use of GenAI in Research Context," *Práxis Educativa* 20, (2025): e25388. https://doi.org/10.5212/PraxEduc.v.20.25388.057; L. W Roberts, "Addressing the Novel Implications of Generative AI for Academic Publishing, Education, and Research," *Academic Medicine* (2024). https://doi.org/10.1097/ACM.0000000000005667; J. Crawford, M. Cowling and K.-A. Allen, "Leadership Is Needed for Ethical ChatGPT: Character, Assessment, and Learning Using Artificial Intelligence (AI)," *Journal of University Teaching & Learning Practice* 20, no. 3 (2023): Article 02. https://doi.org/10.53761/1.20.3.02

[5] T. R. Noviandy, A. Maulana, G. M. Idroes, Z. Zahriah and M. Paristiowati, "Embrace, Don't Avoid: Reimagining Higher Education with Generative Artificial Intelligence," *Journal of Social Research* (2024). https://doi.org/10.60084/jeml.v2i2.233; A. Vieira and A. Mesquita, "Generative Artificial Intelligence in Higher Education: Challenges, Opportunities and Pedagogical Implications," *Journal of Technologies Information and Communication* 5, no. 1 (2025): 36578. https://doi.org/10.55267/rtic/16675; Michel-Villarreal et al., "Challenges and opportunities of generative AI for higher education as explained by ChatGPT".; H. Gimpel, K. Hall, S. Decker, T. Eymann and L. Lämmermann, "Unlocking the Power of Generative AI Models and Systems Such as GPT-4 and ChatGPT for Higher Education: A Guide for Students and Lecturers," *Hohenheim Discussion Papers in Business, Economics and Social Sciences* 02-2023.

[6] R. E Guamán Chávez, "Ética e Integridad Académica en el Uso de la Inteligencia Artificial Generativa en la Educación Superior," *Revista Científica Multidisciplinar G-NER@NDO* (2025). https://doi.org/10.60100/rcmg.v6i1.392; M. S. Kadwa, "Written Assignments and the Ethical Considerations of Artificial Intelligence in Higher Education," *International Journal of Linguistics, Literature and Translation* (2025). https://doi.org/10.32996/ijllt.2025.8.9.27; C. Gallent-Torres, A. Zapata-González and J. L. Ortego-Hernando, "The Impact of Generative Artificial Intelligence in Higher Education: A Focus on Ethics and Academic Integrity," *RELIEVE* 29, no. 2 (2023): Article M5. https://doi.org/10.30827/relieve.v29i2.29134; Bittle and El-Gayar, "Generative AI and academic integrity in higher education: A systematic review and research agenda".

[7] M. D. A. G. Xames and A. A. Grado, "ChatGPT for Research and Publication: Opportunities and Challenges," *Journal of Applied Learning & Teaching* 6, no. 1 (2023). https://doi.org/10.37074/jalt.2023.6.1.20; K. T Kotsis, "The Role of ChatGPT in Academic Writing: Pedagogical and Ethical Dimensions," *EthAIca* 4, (2025): 420. https://doi.org/10.56294/ai2025420; Roberts, "Addressing the novel implications of generative AI for academic publishing, education, and research".; M Askari, "Reclaiming Authorship in the Age of Generative AI: From Panic to Possibility," *AI Magazine* (2025). https://doi.org/10.1002/aaai.70022; J. G. Meyer, R. J. Urbanowicz, P. C. N. Martin, K. O'Connor and R. Li, "ChatGPT and Large Language Models in Academia: Opportunities and Challenges," *Biodata Mining* 16, (2023): 20. https://doi.org/10.1186/s13040-023-00339-9

The "philosophical gap" in the literature is therefore not the absence of ethical concerns, but the fact that these concerns have often not been given theoretical clarity. Existing studies are extremely successful at making the crisis visible, but they often fail to adequately systematize the language of value that explains why it is a crisis. What must be done from this point forward is to approach technology neither through romantic optimism nor through cultural panic, but with normative clarity. The place of AI in higher education can be understood neither through the discourse of innovation alone nor through the discourse of prohibition alone; it must be thought through the question of which human and academic good it serves.[8]

1. Classical Philosophical Foundations: The Normative Axes of AI Ethics

The ethical legitimacy of generative artificial intelligence in academia cannot be grasped solely through rules of use or institutional precautions; at a deeper level, it rests on questions of what the academic subject is, which human and epistemic goods the university protects, and under what conditions thinking remains authentic. For this reason, the debate on AI in higher education must be reconsidered through the foundational normative axes offered by classical philosophy. Kant foregrounds autonomy, duty, and the obligation to use one's own reason; Aristotle foregrounds character, habit, and the telos of education; Plato foregrounds the distinction between truth and appearance and the criteria of authentic learning; Descartes foregrounds methodical doubt, intellectual self-possession, and the risk of cognitive externalization. Read together, these lines show that, in the face of generative AI, the real issue is not merely use, but the transformation of the human relation to knowledge, labor, and thought.

1.1. Kant: Autonomy, Heteronomy, and the Duty to Think for Oneself

The Kantian axis opens not with a prohibition but with a demand. In the essay "What is Enlightenment?" (1784), Kant defines the unenlightened condition as one of self-incurred tutelage - the failure to use one's own understanding without guidance from another - and he locates its cause not in a deficiency of reason but in a deficiency of resolve: "Laziness and cowardice are the reasons why such a large proportion of men, even when nature has long emancipated them from alien guidance, nevertheless gladly remain immature for life."[9] The imperative that follows - "Sapere aude!", "Have courage to use your own understanding!" - names for Kant the foundational obligation of rational beings, an obligation that is not contingent on external conditions but follows from the very nature of rational agency.[10]

The contemporary relevance of this formulation is striking. What Kant identifies as the structural temptation of the unenlightened condition - willingness to surrender the burden of judgment to another in exchange for comfort and efficiency - describes with unusual precision the risk that generative AI poses to the academic subject. The risk is not that AI will prohibit independent thought; it is that the student, researcher, or faculty member may willingly allow AI to relieve them of the work of thinking, because AI makes that relief so fluent and so immediately available. The form of tutelage Kant diagnosed in the eighteenth century has acquired a new technological infrastructure.

---

[8] C. S. Sibilin, "Education and the Epistemological Crisis in the Age of ChatGPT," *Critical Review* 35, no. 4 (2023): 414-25. https://doi.org/10.1080/08913811.2023.2284042; V. Strunk and J. Willis, "Generative Artificial Intelligence and Education: A Brief Ethical Reflection on Autonomy," *EDUCAUSE Review* 13, (January 2025).; Francis et al., "Generative AI in higher education: Balancing innovation and integrity".; T Hyland, "Educational Responses to Artificial Intelligence (AI) Applications: Problems and Promise," *Qeios* (2023). https://doi.org/10.32388/08UCQU; A Watanabe, "Have Courage to Use Your Own Mind, with or without AI: The Relevance of Kant's Enlightenment to Higher Education in the Age of Artificial Intelligence," *Electronic Journal of e-Learning* 22, no. 2 (2024): 46-58. https://doi.org/10.34190/ejel.21.5.3229

[9] Immanuel Kant, "An Answer to the Question: What is Enlightenment?" (1784) in *Kant: Political Writings*, ed. Hans Reiss, 2nd ed. (Cambridge: Cambridge University Press, 1991), 54.

[10] Kant, "What is Enlightenment?," 54. The Latin phrase "Sapere aude" is taken from Horace, "Epistles" I.2.40; Kant adopts it as the motto of the Enlightenment.

To understand the full force of this diagnosis, it is necessary to grasp what Kant means by autonomy and its counterpart, heteronomy. In the Groundwork of the Metaphysics of Morals (1785), Kant identifies the autonomy of the will as the supreme principle of morality: the will is autonomous when it gives the law to itself rather than receiving it from outside. As he writes: "The principle of every human will as a will legislating universally through all its maxims... grounds itself on no interest."[11] Heteronomy, by contrast, arises when the will is governed by something external to reason - by inclination, by the commands of another, or, crucially for our purposes, by the operations of a system that performs the cognitive work of judgment on the subject's behalf. Kant is precise on this point: a will that acts from heteronomous grounds "would need yet another law, which limited the interest of its self-love to the condition of a validity for the universal law."[12] In other words, acting from heteronomy is not merely an imperfection; it is a structural abdication of rational agency.

Applied to the academic context, the Kantian distinction between autonomy and heteronomy yields a diagnostic criterion of considerable specificity. The question is not whether a student or researcher uses AI - assistance, Kant consistently acknowledges, is not incompatible with rational self-governance. The question is whether the subject remains the ultimate author of their own judgment. Kant identifies the three maxims of enlightened thinking as: first, to think for oneself; second, to think from the standpoint of everyone else; and third, to think consistently. [13] Each of these maxims is at stake in the AI debate. The first is violated when the student allows AI to formulate the research question, construct the argument, and determine the conceptual framework, retaining for themselves only the work of submission. The second is undermined when AI-generated outputs, trained on existing data distributions, reproduce dominant perspectives while marginalizing the critical distance that genuine scholarly inquiry requires. The third is compromised when the student accepts AI outputs without subjecting them to the kind of sustained rational scrutiny that alone, on Kantian grounds, constitutes responsible epistemic practice.

The deontological dimension of the Kantian framework introduces a further consideration that the AI debate frequently underestimates. Kant's categorical imperative requires, in one of its central formulations, that one act only according to that maxim which one could at the same time will to become universal law.[14] Applied to academic AI use, this formulation is illuminating: a student who presents AI-generated text as their own work is acting on a maxim - "I may misrepresent the source of my intellectual labor when it is advantageous to do so" - which cannot be universalized without contradiction. If every student adopted this maxim, the very practice of academic assessment, which presupposes a truthful relation between the produced text and the intellectual process it is meant to represent, would be destroyed. The student's particular act of concealment parasitically depends on the general norm of honest representation that its universalization would abolish. This is not merely an argument about rule-following; it is an argument about the rational structure of the academic enterprise itself.

Kant's second formulation of the categorical imperative - to treat humanity, whether in one's own person or that of another, always as an end and never merely as a means - adds a further dimension.[15] The obligation of disclosure is not merely procedural; it is an expression of respect. When a researcher conceals AI's role in producing their work, they do not only mislead their audience; they treat those who receive and build upon the work as means

[11] Immanuel Kant, *Groundwork of the Metaphysics of Morals (1785)*, trans. And ed. Mary Gregor and Jens Timmermann. (Cambridge: Cambridge University Press, 2012), Ak 4:432.

[12] Kant, *Groundwork of the Metaphysics of Morals (1785)*, Ak 4:432–433.

[13] Allen W. Wood, "Kant's Ethical Thought" in *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, ed. Paul Guyer (Lanham: Rowman & Littlefield, 1998), 170. Wood reconstructs the three maxims of enlightened thinking from Kant's Critique of Practical Reason (Ak 5:294–295) and Anthropology (Ak 7:200, 228–229). The maxims are: (1) think for yourself; (2) think from the standpoint of everyone else; (3) think consistently.

[14] Kant, *Groundwork of the Metaphysics of Morals (1785)*, Ak 4:421: "Act only in accordance with that maxim through which you can at the same time will that it should become a universal law."

[15] Kant, *Groundwork of the Metaphysics of Morals (1785)*, Ak 4:429: "Act so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only."

for the advancement of their own reputation, rather than as partners in a shared epistemic enterprise. The norm of disclosure is thus grounded not in institutional rule but in the Kantian principle of respect for persons as rational ends.

Two qualifications must be registered before drawing practical conclusions. First, the Kantian framework, as developed here, operates at the level of the individual rational agent and the maxims on which they act. It is well-suited to diagnose the failure of autonomous judgment and the ethics of disclosure, but it does not, by itself, address the structural conditions under which such failure becomes normalized - the institutional pressures, the incentive structures, the design of assessment regimes - that systematically encourage AI-dependent behavior. This limitation will be taken up when the analysis turns to Heidegger's account of technological enframing. Second, Kant's conception of the rational subject - the autonomous will that legislates universally, independent of sensible inclination - represents a specifically modern philosophical configuration, one that, as the transition to Aristotle will show, differs fundamentally from the ancient understanding of the relation between reason, character, and practice. The Kantian subject is, in principle, prior to its social world and its habits; the Aristotelian subject is constituted by them. Both frameworks are necessary; neither is sufficient on its own.

What the Kantian analysis contributes to the normative model developed in this article is, then, threefold. It establishes the courage to think for oneself as a non-negotiable condition of academic life, grounded not in institutional convention but in the nature of rational agency. It identifies the gradual surrender of judgment to AI - not through coercion but through the comfort of efficiency - as a form of self-incurred tutelage directly analogous to what Kant diagnosed as the structural obstacle to enlightenment. And it grounds the obligation of disclosure in the categorical imperative itself, showing that the honest representation of one's intellectual labor is not merely a procedural requirement but a condition of treating others and oneself as rational ends. The principle that follows is direct: AI use that keeps the rational subject in the position of final author, that subjects every output to the scrutiny of one's own reason, and that represents the epistemic process honestly, is defensible on Kantian grounds. AI use that silently displaces judgment, that accepts outputs without critical scrutiny, and that misrepresents the epistemic process is, in Kant's terms, not merely imprudent but morally impermissible.

1.2. Aristotle: Virtue, Habit, and the Telos of Academic Formation

The Kantian analysis establishes a demanding criterion: legitimate AI use is use that preserves the rational subject's role as the final author of its own judgment, acts in accordance with principles it could universalize, and represents its epistemic process honestly. Yet this criterion, precise as it is, operates at a specific level. It evaluates individual acts and the maxims that govern them; it diagnoses failures of autonomous judgment as they occur. What it cannot do, by itself, is account for the conditions under which the capacity for such judgment is formed in the first place - or eroded, before the student ever faces the question of what maxim to follow. That dimension of the problem belongs to a different philosophical tradition.

Where the Kantian framework asks whether the rational subject has preserved its autonomy, the Aristotelian framework asks a prior and, in some respects, more fundamental question: what kind of person is being formed? This shift in register marks a genuine philosophical difference, not merely a difference of emphasis. For Kant, the moral worth of an act is independent of the character from which it issues; what matters is whether it is done from duty, in accordance with the categorical imperative. For Aristotle, by contrast, the act and the agent cannot be separated in this way: a just act performed accidentally or under compulsion is not yet an expression of justice in any meaningful sense. What matters is that the act issues from a stable disposition - a "hexis" - that has been formed through practice and that reliably inclines the agent toward the good. This is why Aristotelian ethics is irreducible to Kantian deontology and why both are needed: the Kantian framework diagnoses the failure of autonomous

judgment; the Aristotelian framework diagnoses the failure of character formation that makes such judgment possible in the first place.

The foundational text for this analysis is "Nicomachean Ethics" Book II, where Aristotle states the principle that governs the entire argument: "Moral virtue comes about as a result of habit, whence also its name (ethike) is one that is formed by a slight variation from the word ethos (habit)... neither by nature, then, nor contrary to nature do the virtues arise in us; rather we are adapted by nature to receive them, and are made perfect by habit."[16] The significance of this passage for the AI debate is direct. Virtues — including the intellectual virtues that constitute good academic practice — are not capacities that can be acquired by observation, declaration, or the possession of polished outputs. They are dispositions that form through repeated first-personal exercise. A student does not become capable of careful argumentation by reading careful arguments or by receiving AI-generated careful arguments; they become capable of it by repeatedly attempting to construct arguments themselves, failing, reflecting on the failure, and trying again. This is what Aristotle means when he says, in a formulation whose pedagogical implications are precise: "We become just by doing just acts, temperate by doing temperate acts, brave by doing brave acts."[17]

Generative AI, understood in this Aristotelian frame, introduces a specific kind of risk that differs from the Kantian concern about the displacement of autonomous judgment. The Aristotelian risk is subtler and longer-term: it is the risk that the student forms the wrong habits - not through dramatic violations of academic norms, but through the gradual normalization of patterns of cognitive behavior that, repeated over time, shape character in particular directions. If a student consistently turns to AI at the point of difficulty - at the moment when formulating a question becomes hard, when constructing an argument requires sustained effort, when sustained attention to a problem is required - they are not merely receiving external assistance with a particular task. They are, in the Aristotelian sense, practicing a mode of responding to intellectual difficulty. And practice, for Aristotle, is constitutive: repeated patterns of behavior become dispositions, and dispositions become character. The student who habitually offloads difficulty may be forming habits of cognitive impatience, dependence on external resolution, and intolerance of the friction that genuine inquiry requires - precisely the dispositions that are antithetical to intellectual virtue.

The concept of "phronesis" - practical wisdom - is essential to the Aristotelian analysis at this point, and it is one that has no equivalent in the Kantian framework. Aristotle defines "phronesis" as "a true and reasoned state of capacity to act with regard to the things that are good or bad for man."[18] It is not a capacity to apply universal rules to particular cases; it is the capacity to perceive what is required in a particular situation, to judge what matters here and now, in a way that cannot be reduced to rule- following or algorithmic procedure. Aristotle is explicit that "phronesis" cannot be merely theoretical: "while there is such a thing as excellence in art, there is no such thing as excellence in practical wisdom; and in art he who errs willingly is preferable, but in practical wisdom, as in the virtues, he is the reverse."[19] The practically wise person does not produce correct outputs by following a procedure; they see clearly, judge well, and act rightly - and they can do this because character and understanding have been formed together through a sustained practice of engagement with real situations.

Generative AI cannot have "phronesis", and this is not an accidental limitation that future development might overcome. "Phronesis" presupposes a being for whom things matter, a being embedded in a form of life with particular commitments and vulnerabilities, a being whose understanding of what is good has been formed through the

[16] Aristotle, "Nicomachean Ethics" in *The Complete Works of Aristotle*, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984), II.1, 1103a14–25.
[17] Aristotle, "Nicomachean Ethics," II.1, 1103b1–2.
[18] Aristotle, "Nicomachean Ethics," VI.5, 1140b4–6.
[19] Aristotle, "Nicomachean Ethics," VI.5, 1140b22–25.

experience of success and failure in real situations. What AI can produce is an extraordinarily sophisticated simulation of the outputs that "phronesis" characteristically generates - contextually sensitive formulations, apparently calibrated judgments, plausible arguments. But the simulation of the output is not the possession of the capacity. A student who routinely substitutes AI-generated outputs for their own attempts at judgment is not merely bypassing "phronesis"; they are depriving themselves of the conditions under which "phronesis" forms. Academic writing, in this perspective, is not merely a communicative act; it is a practice of judgment, and its value as a pedagogical form lies precisely in requiring the student to exercise the capacities that constitute intellectual character.

A further dimension of the Aristotelian analysis concerns the relation between "phronesis" and moral virtue. Aristotle insists that the two are inseparable: "it is impossible to be practically wise without being good."[20] The practically wise person is not merely technically skilled; they are oriented toward the right ends. This connection matters in the academic context because it means that the AI question cannot be assessed as a purely technical or instrumental matter - as a question about which tools are useful and which are not. The question of how AI is used in academic work is also a question about the ends toward which that work is oriented. If the student's end is the production of a text that satisfies formal requirements, AI is a highly effective instrument. If the student's end is the formation of the kind of intellectual character that genuine education aims to cultivate, the calculus changes: what matters is not the quality of the output but the quality of the engagement that produced it.

This Aristotelian perspective also casts light on the question of what legitimate AI use looks like in practice. The criterion is not the quantity of AI assistance but the disposition it serves and the habits it reinforces. AI used to generate a complete draft that the student then accepts with minimal engagement is an exercise in the wrong habit: it trains the student to treat intellectual difficulty as an obstacle to be removed rather than as the productive friction through which understanding develops. AI used as an object of critique - where the student reads an AI-generated argument, identifies its weaknesses, reconstructs its assumptions, and produces a response that engages its best version - is a different matter. Here, AI functions not as a replacement for intellectual engagement but as a provocation to it. The Aristotelian criterion, therefore, is whether the use strengthens or weakens the dispositions constitutive of intellectual character: the capacity to sustain attention, to tolerate uncertainty, to engage with difficulty, and to form and test judgments through one's own cognitive effort.[21]

There is, finally, a dimension of the Aristotelian framework that must be recognized as both a limitation and a resource. The "telos" toward which Aristotelian education is directed - the cultivation of "eudaimonia", the flourishing of the human being in accordance with its highest capacities - is embedded in a particular conception of the human being's relation to its community, its practices, and its world. The Aristotelian subject is not the self-grounding rational legislator of Kantian ethics; it is a being whose capacities for excellent activity are realized only within particular social and institutional conditions. This means that the Aristotelian framework, unlike the Kantian, cannot assess AI use solely at the level of the individual agent. The habits that AI use reinforce or undermine are shaped by the institutional environments in which that use occurs - by course design, assessment structures, the norms of academic communities, and the incentive structures that govern them. This is a point to which the discussion will return when examining institutional governance. For now, the fundamental Aristotelian principle stands: the legitimate use of AI in academic settings is use that serves the formation of intellectual virtue - patient inquiry, careful judgment, honest self-assessment, and the courage to engage with difficulty rather than outsource it.

[20] Aristotle, "Nicomachean Ethics," VI.13, 1144b31–32.

[21] Petricini, "What Would Aristotle Do?" Petricini applies the Aristotelian virtue framework directly to the navigation of generative AI in higher education, arguing that the central pedagogical question is not whether AI is used but whether its use reinforces or undermines the intellectual dispositions — attentiveness, tolerance of difficulty, sustained engagement — that constitute genuine scholarly formation. The analysis converges with the criterion developed here.

1.3. Plato: Knowledge, Appearance, and the Technology of Forgetting

The Aristotelian analysis has shown that the AI question is, at its core, a question about formation: whether the patterns of engagement that AI use reinforce over time cultivate or undermine the intellectual character that genuine academic work requires. This shifts the focus from the discrete act of judgment - Kant's concern - to the long arc of habituation through which the capacity for judgment develops or fails to develop. Yet both frameworks, Kantian and Aristotelian alike, share a common presupposition: that the issue can be framed in terms of what the student or researcher does with AI. There is a prior question that neither framework poses directly, a question about what the products of AI-assisted work actually are - whether they constitute knowledge in any philosophically meaningful sense, or whether they present, with increasing fluency, only the appearance of it.

The Platonic axis introduces a dimension that neither the Kantian nor the Aristotelian framework addresses directly: the question of whether what is produced and circulated in the academic world constitutes knowledge at all, or merely its appearance. For Plato, this is not an incidental concern but the central problem of education and of the philosophical life. The distinction between "episteme" — genuine, grounded knowledge — and "doxa" — true opinion that cannot give an account of itself — runs through the dialogues as a persistent diagnostic instrument. What makes the Platonic analysis particularly valuable in the present context is not only the "episteme/doxa" distinction itself but a specific text in which Plato engages directly with the epistemological consequences of a new technology of knowledge production. That text is the "Phaedrus", and the parallel it offers to generative AI is remarkably precise.

In the "Phaedrus" (274c–275b), Socrates recounts the Egyptian myth of Theuth and Thamus. Theuth, the inventor of letters, presents his invention to King Thamus with the following claim: "This branch of learning, my king, will make the Egyptians wiser and will improve their memory. The drug for memory and wisdom has been discovered!" Thamus's response is Plato's philosophical verdict on writing as a technology: "You have not discovered a drug for memory, but for reminding. You offer your students an apparent, not true wisdom. For they have heard much from you without real teaching, and they will appear rich in knowledge when for the most part there's an absence of knowledge, and they will be difficult to be with since they appear wise rather than really being wise."[22]

The structure of Thamus's objection is precise and demands careful attention, because it is not a rejection of writing as such but an analysis of what writing does to the relation between the learner and knowledge. Writing, Thamus argues, produces a forgetting in those who rely upon it: they "recollect from the outside with foreign signs, rather than themselves recollecting from within by themselves."[23] The result is a specific kind of epistemic deception — not the deception of false information, but something more insidious: the appearance of knowledge in the absence of knowledge. The person who has read widely but not understood deeply presents the outward marks of learning while lacking its substance.

Socrates elaborates the point through the contrast between written speech and living dialogue. Written words, he observes, "stand there as if alive, but if you question them, they remain in complete and solemn silence. The same for words written down. You might suspect that they would speak as if they understand something, but if you ask them about anything in the text in the hopes of learning something, the words signify only one thing, and always the same thing."[24] The defect of writing, for Plato, is not that it conveys false content but that it cannot respond,

[22] Plato, *Phaedrus*, trans. S. Scully. (Newburyport, MA: Focus Publishing, 2003), 274e–275a. The passage runs from 274c to 275b in the standard Stephanus pagination.; Loos and Radicke, "Using ChatGPT-3 as a Writing Tool." Loos and Radicke apply Plato's critique of writing in the Phaedrus directly to ChatGPT, arguing that contemporary media representations of AI writing tools reproduce precisely the structure Thamus warns against: the appearance of wisdom in the absence of genuine understanding. Their analysis provides a contemporary elaboration of the Platonic argument developed here.

[23] Plato, *Phaedrus*, 275a.

[24] Plato, *Phaedrus*, 275d.

cannot adapt, cannot engage with the particular questioner. It is a static deposit of the results of thinking, severed from the living process that produced it.

The application of this analysis to generative AI is not an exercise in loose analogy; it is a direct extension of Plato's argument. Generative AI is, in the relevant sense, a technology of production that is structurally analogous to writing as Thamus describes it — and in crucial respects more extreme. Writing at least preserves the traces of a human act of understanding; the written text is the "eikon", the image, of living speech that was once "written with knowledge in the soul of one who understands."[25] A generative AI output is not even this: it is produced by a system that has processed vast corpora of human discourse without any act of understanding, without any soul in which knowledge was present. What the student receives from AI is not the image of someone else's understanding but a statistically coherent arrangement of linguistic patterns that presents the form of knowledge while being epistemically empty in the relevant sense. The Platonic warning - that apparent wisdom is more dangerous than acknowledged ignorance, because it forecloses the inquiry that genuine ignorance motivates - applies here with heightened force.

The "Meno" provides the complementary conceptual instrument. In that dialogue, Socrates draws the distinction between "episteme" and "doxa" through the figure of Daedalus's statues: "True opinions, as long as they remain, are a fine thing and all they do is good, but they are not willing to remain long, and they escape from a man's mind, so that they are not worth much until one ties them down by giving an account of the reason why... After they are tied down, in the first place they become knowledge, and then they remain in place. That is why knowledge is prized higher than correct opinion, and knowledge differs from correct opinion in being tied down."[26] The difference between true opinion and knowledge is not a difference in the propositional content of what is believed; a person with true opinion may believe exactly the same thing as a person with knowledge. The difference lies in the account (logos) that the knower can give - the ability to explain why what is believed is true, to trace its grounds, to connect it to other truths, and to recognize the limits of what one knows.

This distinction cuts directly to the core of the problem posed by generative AI in academic settings. AI can produce outputs that are propositionally true, formally coherent, and contextually plausible — outputs that, to an external observer, present all the marks of genuine understanding. What AI cannot produce is the "logos" that transforms opinion into knowledge: the internal account that the knower can give because they have traced the reasoning themselves, engaged with the difficulty, and arrived at understanding through their own epistemic effort. A student who presents an AI-generated text has, in Platonic terms, traded the slow and demanding process of tying down opinions into knowledge for the immediate availability of a text that looks like the product of that process without being it. The Platonic analysis names precisely what is lost in this exchange: not information, not even correct opinion, but the "logos" - the capacity for self-grounded rational account — that constitutes the difference between knowing and merely seeming to know.

The "Republic"'s allegory of the cave extends this analysis to its most fundamental level.[27] Education, for Plato, is not the transmission of content into a passive recipient; it is a turning of the whole soul - "periagoge" - away from shadows toward the light of truth. The prisoners in the cave mistake images for realities, not because they are unintelligent, but because they have never been compelled to turn and look at what casts the shadows. Their condition is not ignorance of particular facts; it is a structural orientation toward appearances that precludes genuine

[25] Plato, *Phaedrus*, 276a. The contrast between the living speech "written with knowledge in the soul" and its written image runs from 275c to 277a.

[26] Plato, "Meno" in *Plato: Complete Works*, ed. John M. Cooper (Indianapolis: Hackett, 1997), 97e–98a.

[27] Plato, "Republic" in *Plato: Complete Works*, ed. John M. Cooper (Indianapolis: Hackett, 1997), 514a–521b. The allegory of the cave occupies Book VII, 514a–521b; the concept of "periagoge" (turning of the soul) is stated at 518c–d.

cognition. The pedagogical implication is direct: education is not a matter of providing better or more accurate images. It is a matter of transforming the orientation of the learner - developing the capacity and the disposition to turn toward truth rather than settle for its semblance.

From this perspective, the biggest risk of generative AI in academic settings is not the risk of individual deception — though that is real and serious — but the risk of normalizing the cave. If students are systematically trained, through the design of their educational environments, to treat the polished outputs of AI as substitutes for the hard work of understanding, they are not merely bypassing a particular cognitive operation. They are being habituated to an orientation toward appearance — toward the semblance of understanding rather than its substance - in a way that, if left unchallenged, may become the default mode of their intellectual lives. The Platonic framework insists that genuine education must resist this normalization, not through prohibition alone but through the design of learning environments that keep the difference between knowledge and its appearance continuously visible and pedagogically salient.[28]

A qualification must be registered here, as it was in the Kantian and Aristotelian sections. Plato's critique of writing in the "Phaedrus" is not ultimately a rejection of writing; it is a reminder that writing must serve living inquiry rather than replace it. Socrates himself writes nothing, yet the dialogues he inspired remain among the most powerful instruments of philosophical education ever produced. The Platonic point, applied to AI, is not that AI should be excluded from academic life but that its relation to genuine learning must be kept clear and continuously interrogated. AI is a tool that can provide information, generate frameworks, and present plausible formulations - but it cannot perform the "periagoge", the turning of the soul toward truth, that is the substance of education. That turning can only be performed by the learner, through the sustained, first-personal effort of confronting difficulty, questioning appearances, and demanding of oneself the "logos" that transforms opinion into knowledge.

1.4. Descartes: Method, the Cogito, and the Risk of Epistemic Displacement

The Platonic analysis has identified the deepest risk of AI-assisted academic work as an epistemic one: the normalization of the appearance of knowledge in the place of knowledge itself, the substitution of fluent outputs for the first-personal cognitive labor through which genuine understanding forms. This risk operates at the level of what the learner receives and circulates - the text, the output, the product. The Cartesian framework shifts the question one level deeper: not what is received, but what is performed. If Plato asks whether the product is knowledge or its simulation, Descartes asks who is actually thinking when the product is made - and whether the subject remains the author of their own cognitive process, or whether that process has been progressively assumed by the system.

Of the four classical philosophical axes examined in this article, the Cartesian framework is perhaps the most internally complex in its bearing on generative AI. Descartes is often invoked in AI discussions as a source of warnings about the externalization of thought - and there is genuine philosophical warrant for this reading. Yet to treat Descartes merely as a cautionary voice is to miss the specificity of what his philosophy contributes. The productive question is not simply whether AI threatens human thinking, but how the Cartesian account of what thinking is, and what it requires, allows us to diagnose the particular form of that threat.

The methodological foundation of Descartes' epistemology is laid in the Discourse on the Method, where he states his governing principle with characteristic directness: "I think, therefore I am" - but the force of this principle

[28] Huang et al., "A Case Study for Educators with ChatGPT and Plato's Allegory of the Cave." Huang and colleagues demonstrate, through a concrete case study, how Plato's allegory of the cave can function as a productive pedagogical framework for educators navigating the age of ChatGPT — showing that the structural risk of mistaking AI-generated outputs for genuine understanding directly instantiates what Plato diagnosed as the condition of the prisoners.

depends on the procedure that precedes it.[29] The "cogito" is not a logical deduction; it is the one certitude that survives the method of systematic doubt. Descartes' method requires that the thinker personally perform the work of withholding assent from everything that can be doubted, until the one thing that cannot be doubted - the very act of doubting, and thus of thinking - reveals itself. As he writes in the "Meditations on First Philosophy": "I must withhold my assent from these former beliefs just as carefully as I would from obvious falsehoods, if I want to discover any certainty in the sciences."[30] The point to hold firmly is that this certitude cannot be borrowed, received, or delegated. The "cogito" is not a conclusion that an authority can hand to the thinker. It must be arrived at through a first-personal act of inquiry that the thinker performs for themselves.

This structural feature of the Cartesian method is what gives it its contemporary relevance. The question it raises for generative AI is not whether AI can produce outputs that look like thought - plainly it can, with extraordinary fluency - but whether those outputs can substitute for the first-personal performance of inquiry that Descartes regards as the condition of genuine knowledge. The answer, on Cartesian grounds, is clearly no. When a student delegates to AI the work of formulating a research question, constructing an argument, identifying the weaknesses in a position, or synthesizing sources, they are not merely receiving assistance; they are bypassing precisely the cognitive operations that the Cartesian tradition identifies as constitutive of intellectual agency. The result may be a well-constructed text, but it is not, in any epistemologically meaningful sense, the student's knowledge. The product has been separated from the process of inquiry that alone - for Descartes - grounds a claim to understanding.

The distinction between thinking and processing is indispensable here. Generative AI produces outputs through the statistical modeling of vast corpora. The process is one of pattern recognition and probabilistic generation; it involves no act of attention, no withholding of assent, no encounter with doubt. As Dukor and Izeji observe in their epistemological analysis of the relation between Descartes' thinking mind and AI processing, the Cartesian "cogito" names something categorically different from computational operation: it is thought's self-presence to itself, the first-personal awareness that one is thinking, which has no equivalent in any system that processes without experiencing.[31] The machine operates; it does not doubt. And because it does not doubt, it cannot, in Descartes' sense, think.

This distinction matters practically, not only philosophically. The Cartesian framework makes visible a specific kind of risk that the AI debate often underestimates: not the sudden replacement of human thinking, but its gradual displacement. The risk is that if enough cognitive operations - question formulation, source selection, conceptual synthesis, argumentative structure - are systematically offloaded to AI, the student's capacity to perform those operations independently may atrophy without the student noticing, because the outputs continue to appear adequate. The student may believe they are thinking; what is actually happening is that the infrastructure of their thinking has been progressively taken over by the system. Descartes' term for the failure to use one's own reason is "prejudice" in the technical sense: the uncritical acceptance of received opinion without subjecting it to personal scrutiny.[32] In the age of generative AI, the form this prejudice takes is subtler than the deference to authority Descartes criticized in the Scholastic tradition. It is a deference to fluency - the willingness to accept as understanding what is in fact only well-organized output.

---

[29] René Descartes, *Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth in the Sciences (1637*), Part IV. The standard translation is by Donald A., *Cress: Discourse on Method and Meditations on First Philosophy*, 4th ed. (Indianapolis: Hackett, 1998), 17-18.

[30] René Descartes, *Meditations on First Philosophy* (1641), *First Meditation, in Discourse on Method and Meditations on First Philosophy*, trans. Donald A. Cress. 4th ed. (Indianapolis: Hackett, 1998), 59.

[31] M. Dukor and E. Izeji, "Investigating Epistemological Relations Between Descartes' Thinking Mind and Artificial Intelligence Processing," *SSRN preprint*, n.d., 4–7.

[32] Descartes, *Discourse on the Method*, Part I, 3–4.

Fırat's recent reading of Descartes' philosophy of autodidacticism in postdigital education offers a productive elaboration of this danger. Fırat argues that the Cartesian commitment to inquiry-led learning - learning in which the individual constructs their own method rather than receiving pre-packaged conclusions - is not a historical curiosity but a pedagogical norm with enduring force. In postdigital environments, the risk is not that students will be formally prevented from learning, but that the infrastructure of learning will offer such efficient shortcuts to resolution that the experience of genuine inquiry - of remaining with a problem, of forming and testing hypotheses, of discovering the inadequacy of one's initial understanding - will be systematically bypassed.[33] The student may, on this view, graduate with a credential but without having undergone the cognitive formation that Descartes regarded as the substance of education.

The pedagogical consequence for academic practice follows directly from this analysis. In Cartesian terms, the legitimate use of AI is use that keeps the student's own act of judgment - their own withholding and granting of assent - at the center of the intellectual process. AI that presents a student with a range of perspectives on a problem and requires them to assess and choose between those perspectives is different in kind from AI that presents a student with a conclusion and invites them to accept it. The first supports the methodical process; the second bypasses it. This is the Cartesian criterion for distinguishing productive from problematic uses of AI: not the quantity of assistance, but whether the locus of epistemic agency remains with the student or has been displaced to the system.

There is, however, a dimension of the Cartesian framework that must be held alongside this analysis and that complicates any straightforwardly celebratory reading of methodical doubt as the answer to the AI problem. Descartes' project is itself a form of subjectivism: the establishment of the thinking subject as the ground of all certainty. This move, as the transition to the Heideggerian analysis in the following section will show, is part of the historical trajectory that makes modern technology's transformation of the world possible. The emphasis on the autonomous rational subject - the self-grounding individual who clears away received authority and constructs knowledge on the basis of their own reason - is not an innocent epistemological position. It participates in the same metaphysical configuration that, in its late-modern radicalization, dissolves the subject itself into a resource to be optimized. This does not undermine the Cartesian diagnosis of AI-induced epistemic displacement; it means, rather, that the Cartesian framework must be extended and critically situated rather than taken as sufficient on its own.

What the Cartesian analysis contributes to the integrated normative model developed in this article is, then, twofold. It identifies methodical doubt - the first-personal performance of inquiry, the withholding of assent, the awareness that one is the author of one's own cognitive process - as a non-delegable condition of genuine academic work. And it identifies the gradual displacement of this condition by systematically efficient AI assistance as the specific form of epistemic risk that the Cartesian tradition is best placed to diagnose. The principle that follows is direct: AI use that keeps methodical doubt alive, that preserves the student's rational responsibility for their conclusions, and that facilitates thought without performing it on the student's behalf, is defensible within the Cartesian framework. AI use that short-circuits the first-personal encounter with difficulty and uncertainty - however fluent its outputs - is, on Cartesian grounds, an obstacle to intellectual formation rather than a support for it.

2. Contemporary Extensions: Phenomenology, Heidegger, and Digital Habitus

The four classical frameworks developed above illuminate, from different angles, what is at stake when generative AI enters the academic world: the displacement of autonomous judgment, the erosion of intellectual character, the substitution of knowledge's appearance for its substance, and the externalization of the cognitive process

[33] M. Fırat, "Reimagining Descartes' Philosophy of Autodidacticism in the Postdigital Education," *Educational Philosophy and Theory*, advance online publication (2026). https://doi.org/10.1080/00131857.2025.2611000

itself. Yet each of these frameworks shares a common presupposition — that the problem can be assessed at the level of what the individual subject does. What they cannot, by structural design, address is the prior question: what kind of epistemic world is being built around the subject, and how does that world shape what counts as thinking before any individual decision is made?

Generative AI is not merely a neutral tool used externally by the subject; it is an environmental technology that settles into regimes of attention, habits of writing, practices of decision-making, and self-understanding.[34] For this reason, the discussion must be expanded through contemporary lines of thought such as phenomenology, Heidegger, and digital habitus. This expansion continues where classical philosophy leaves off, shifting the focus from normative principles alone to the structure of experience and the way the world is disclosed.

The basic contribution of the phenomenological perspective is to show that meaning is not exhausted by symbolic processing - and that this is not a contingent limitation of current AI systems but a structural consequence of what meaning is. The founding insight of phenomenology, established by Husserl and developed by his successors, is that consciousness is essentially intentional: every act of awareness is directed toward something, it is always consciousness of an object, situation, or state of affairs. As Husserl's analysis of intentionality shows, this directedness is not a relation between a mind and a world standing over against it, but the very structure through which a world becomes present at all.[35] Intentionality is not a property that minds have in addition to processing information; it is the condition under which anything can show up as meaningful in the first place. A generative AI system processes tokens and generates statistically coherent sequences. It operates, in Husserl's terms, entirely at the level of sign-manipulation - without the intentional acts through which signs acquire meaning in the first place.

Merleau-Ponty's phenomenology of embodiment deepens this analysis at a level that is especially pertinent to academic work. For Merleau-Ponty, the body is not an instrument the mind uses to interact with the world; it is the very medium through which the world becomes intelligible. As he writes: "The body is our general medium for having a world."[36] Meaning does not arise first in consciousness and then get expressed through the body; it arises in the body's ongoing, pre-reflective engagement with its environment. This is what Merleau-Ponty calls operative intentionality - an intentionality "already at work before any positing or any judgement," a form of understanding embedded in skilled comportment, perceptual familiarity, and the habituated engagement with a domain of practice.[37] Writing, reading, and inquiry - the core activities of academic life - are forms of this embodied engagement. The student who works through a problem is not merely manipulating symbols according to rules; they are bringing to bear a whole sedimented history of bodily and cognitive engagement with a discipline, a history that shapes what counts as a question, what registers as a difficulty, and what feels like understanding. No AI system has or can have this history, because it has no body through which a world of practice is first disclosed.

Dreyfus's critique of classical AI extends the phenomenological argument into the specific terrain of artificial cognition. Dreyfus argued, drawing on both Heidegger and Merleau-Ponty, that the foundational assumption of classical AI - that intelligent behavior can be reproduced by formalizing the rules implicit in skilled human

[34] Bekalp, "Fenomenolojik Perspektiften Yapay Zekâ: Yönelimsellik, Beden ve Çerçeve Problemi".; N. Durmaz, "Heidegger'in Felsefesi Bağlamında Yapay Zekâ," *MetaZihin* 8, no. 2 (2025): 181. https://doi.org/10.51404/metazihin.1767253

[35] The account of intentionality given here draws on Husserl's development of the concept across the Logical Investigations and Ideas I. For a concise account of Husserl's breakthrough recognition of intentionality as "the essential correlation between thinking and its object," see Dermot Moran, *Introduction to Phenomenology* (London: Routledge, 2000), xxxi–xxxii. The key text is Edmund Husserl, *Ideas About a Pure Phenomenology and to a Phenomenological Philosophy, First Book*, trans. F. Kersten. (Dordrecht: Kluwer, 1983), §84: intentionality as the defining characteristic of consciousness "in the pregnant sense."

[36] Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. Colin Smith. (London: Routledge, 2002), 169.

[37] Merleau-Ponty, *Phenomenology of Perception*, 498. The phrase "operative intentionality" (fungierende Intentionalität) is Merleau-Ponty's retrieval of a concept from Husserl's later manuscripts; Merleau-Ponty uses it to designate the pre-reflective bodily engagement with the world that precedes and conditions all explicit acts of consciousness.

performance - rests on a philosophical error: it presupposes that human beings understand their worlds through explicit representations and rule-following, when in fact the background of practical understanding through which we navigate the world is non-representational, holistic, and irreducibly bodily.[38] The same diagnosis applies, with different emphases, to large language models. These systems are extraordinarily effective at generating outputs that resemble the products of understanding; but resemblance to the outputs of understanding is not understanding. What is missing is what Dreyfus called the background: the pre-thematic, embodied, situationally sensitive orientation to a domain that makes it possible to recognize what matters, to perceive relevance, and to sustain the kind of engaged attention through which genuine inquiry proceeds. Academic writing and learning, for precisely this reason, are not reducible to linguistic regularity; they involve embodied attention, lived intensity, historical context, and the first-personal orientation toward a problem that no system without a world can possess.

The Heideggerian line carries this ontological inquiry to its deepest register. What distinguishes Heidegger's approach from the classical philosophical axes examined earlier is not merely that it operates at a more fundamental level, but that it questions the very framework within which the AI debate is ordinarily conducted. For Heidegger, technology is not a neutral aggregate of instruments, nor is it simply a domain of human activity that can be regulated from the outside. Technology is, rather, a particular historical mode in which being itself is disclosed - a way of revealing that progressively transforms everything it touches, including the human being and the act of thinking itself.

Heidegger announces this orientation with a deliberately paradoxical opening: "The essence of technology is by no means anything technological."[39] The point is not rhetorical provocation but philosophical precision: to understand what technology does, one cannot remain at the level of particular machines or devices. The essence of modern technology - what Heidegger calls "Gestell", rendered in English as "Enframing" - operates before any specific apparatus comes into view. Enframing designates "that way of revealing which holds sway in the essence of modern technology and which is itself nothing technological."[40] More precisely, Enframing is "the gathering together of that setting-upon which sets upon man, i.e., challenges him forth, to reveal the real, in the mode of ordering, as standing-reserve."[41]

The concept of standing-reserve (Bestand) is indispensable here. Under the dominion of Enframing, entities no longer stand over against the human being as objects in the classical sense; they come to presence exclusively as resources ordered and held in readiness for further ordering. As Heidegger writes, "everywhere everything is ordered to stand by, to be immediately at hand, indeed to stand there just so that it may be on call for a further ordering."[42] The standing-reserve is not merely a stock of things; it is a mode of being - the manner in which everything, including human beings themselves, presences when Enframing prevails. What is most disquieting, Heidegger observes, is not that man orders things as standing- reserve but that he too risks being drawn into the same ordering: man "comes to the very brink of a precipitous fall; that is, he comes to the point where he himself will have to be taken as standing-reserve."[43]

[38] Hubert L. Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian," *Philosophical Psychology* 20, no. 2 (2007): 247–268, here 251–253. Dreyfus argues that skillful human coping with the world is grounded in "a background of practices, the significance of which is built into our bodily responses," and that this background is precisely what cannot be captured in explicit rules or representations.

[39] Martin Heidegger, "The Question Concerning Technology" in *The Question Concerning Technology and Other Essays*, trans. William Lovitt (New York: Harper & Row, 1977), 4.

[40] Heidegger, "The Question Concerning Technology," 20.

[41] Heidegger, "The Question Concerning Technology," 20.

[42] Heidegger, "The Question Concerning Technology," 17.

[43] Heidegger, "The Question Concerning Technology," 27.

Transposed into the academic context, this analysis yields a diagnosis that goes well beyond questions of integrity or efficiency. When generative AI is integrated into higher education under the dominant logic of optimization - faster writing, more output, measurable performance - it does not merely introduce a new tool. It deepens the hold of Enframing over the university's own self-understanding. Knowledge, the research problem, the student, the text, and even the act of thinking itself begin to appear as units of operation to be measured and accelerated. The university risks ceasing to be what Heidegger would recognize as a site of disclosure - a place where truth comes to pass through patient, unrepeatable inquiry - and becoming instead a platform for managing and optimizing performance flows. As Thomson notes, precisely in this tendency lies what Heidegger calls "the greatest danger": not the particular machines, but the possibility that Enframing might "drive out every other possibility of revealing," foreclosing modes of thinking that resist quantification and ordering.[44]

Yet here a question that the existing Heideggerian literature on AI often leaves unanswered must be faced directly: in Heidegger's own terms, what is thinking? This question matters enormously for any evaluation of generative AI in academia, because the answer determines whether the distinction between "supporting thought" and "substituting for thought" can be drawn at all.

In "What Is Called Thinking?", Heidegger opens with a provocation that has lost none of its force: "Most thought-provoking in our thought-provoking time is that we are still not thinking."[45] This is not a complaint about intellectual laziness. It points to a structural condition: what gives itself most urgently to be thought tends to withdraw, and it is precisely in the experience of this withdrawal that genuine thinking begins. Heidegger elaborates the point through a characteristically sharp formulation: "Science does not think."[46] The claim is not that science is worthless or that its methods are somehow mistaken. Rather, science secures its domain precisely by determining in advance the conceptual parameters within which it operates; it works within a settled framework of representation. Thinking - in Heidegger's sense - is something else: it is the movement toward what withdraws, what refuses easy formulation, what demands that the questioner remain in uncertainty rather than resolve it prematurely.

This distinction between calculative, representational cognition and what Heidegger calls genuine thinking bears directly on the question of generative AI. A large language model processes patterns across vast corpora of text and generates statistically coherent outputs. It operates entirely within the domain of what Heidegger would identify as representation and ordering - with extraordinary efficiency, but without any of the conditions that Heidegger regards as constitutive of thinking: no orientation toward what withdraws, no self-interrogation, no existence as "Dasein" - a being for whom its own being is an issue.[47] The outputs may appear thoughtful; the process is not thinking in any philosophically meaningful sense.

The pedagogical consequence is significant. When a student delegates to AI the work of formulating questions, constructing arguments, and synthesizing sources, they are not merely receiving external assistance. They are bypassing the very friction in which, on a Heideggerian account, intellectual formation occurs. To learn to think is not to acquire a set of propositions; it is to become the kind of being capable of dwelling with difficulty, sustaining uncertainty, and remaining open to what resists ready formulation. Generative AI, precisely by resolving such difficulty on demand, may quietly foreclose these experiences before the student has had the opportunity to undergo them.

---

[44] Iain D. Thomson, *Heidegger on Technology's Danger and Promise in the Age of AI* (Cambridge: Cambridge University Press, 2025), 45.; Heidegger, "The Question Concerning Technology," 27.

[45] Martin Heidegger, *What Is Called Thinking?,* trans. J. Glenn Gray. (New York: Harper & Row, 1968), 6.

[46] Heidegger, *What Is Called Thinking?,* 8.

[47] Thomson, *Heidegger on Technology's Danger and Promise*, 24. On Dasein as the being for whom its own being is an issue, see Martin Heidegger, *Being and Time*, trans. John Macquarrie and Edward Robinson. (Oxford: Blackwell, 1962), 32.

A further dimension of the Heideggerian analysis must be addressed at this point, one that carries implications for the entire architecture of the philosophical argument advanced in this article. Throughout the preceding sections, Aristotelian virtue and Kantian autonomy have been brought to bear on the AI question as if they could be combined straightforwardly into a unified normative framework. Heidegger's analysis complicates this assumption. For Aristotle, the acquisition of intellectual virtue occurs within a form of life in which the human being is already embedded in a meaningful world - the "polis", shared practices, the cultivation of character through habituation. "Phronesis" is inseparable from this embeddedness: it is the capacity to perceive what matters in particular situations, and it can only be developed through genuine engagement with those situations. For Kant, the moral worth of an action depends on the autonomy of the rational will - the subject's capacity to legislate universal law from within its own reason, independent of external determination. These two frameworks rest on fundamentally different conceptions of the human being's relation to the world. The Aristotelian subject is constituted by its relationships and practices; the Kantian subject is, in principle, prior to them.

Heidegger's analysis reveals that both frameworks presuppose what he regards as a still unexamined understanding of being - one that becomes fully manifest, and fully problematic, in modernity. The Cartesian establishment of the subject as the ground of certainty, and its subsequent Kantian radicalization in the form of the autonomous rational will, are, for Heidegger, steps in the progressive consolidation of subjectivism: the metaphysical configuration in which the human being becomes the measure and guarantor of all that is, and in which the world becomes, in Heidegger's term, a "world-picture" - something represented before a subject and available for mastery and ordering.[48] Nietzsche thinks the unthought of this tradition to its conclusion, dissolving even the subject into the will to power, so that all things, the subject included, become resources to be optimized. Enframing is the historical coming-to-presence of this metaphysical trajectory.

What this means for the present argument is not that Aristotelian or Kantian categories are simply mistaken or that they should be abandoned. Their diagnostic power - particularly Aristotelian attention to habit and character, and Kantian insistence on disclosure and moral responsibility — remains genuine and is retained in the normative model this article develops. The point, rather, is that these frameworks cannot, by themselves, address the level at which Enframing operates. Virtue ethics can ask what kind of intellectual character AI use cultivates; deontology can ask whether autonomy is preserved; but neither can ask what it means that the very concept of efficiency has become the self-evident standard against which academic activity is measured - that the university itself has come to inhabit a world in which everything, including thinking, appears as standing-reserve. To pose that question requires moving onto the terrain that Heidegger opens.

This does not leave the discussion without resources. Heidegger himself, following Hölderlin, insists that "where the danger is, grows / the saving power also."[49] The danger of Enframing and the possibility of a freer relation to technology are not opposites but belong together: it is precisely by coming to see what Enframing is that the possibility of resisting its monopoly on revealing arises. In the academic context, this translates into a specific demand: that the university not merely react to AI as a problem of regulation or detection, but think through what it means that its own practices have been shaped by the logic of optimization - and ask what modes of inquiry, assessment, and pedagogy remain capable of keeping open the space in which something other than ordering can come to pass. In Heidegger's terms, "questioning is the piety of thought";[50] and it is this commitment to questioning - irreducible to efficiency, unanswerable by any algorithm - that marks the boundary between thinking and its technological substitutes.

[48] Heidegger, "The Age of the World Picture," 129–130.
[49] Heidegger, "The Question Concerning Technology," 28. The lines are from Hölderlin, "Patmos".
[50] Heidegger, "The Question Concerning Technology," 35.

The concept of digital habitus, developed by Alberto Romele in dialogue with Bourdieu's social theory, names a specific mechanism through which this reshaping occurs at the level of pre-reflective practice. For Bourdieu, habitus designates the system of durable, transposable dispositions - practical schemes of perception, evaluation, and action - that agents acquire through their participation in particular social fields, and that orient their conduct without requiring conscious deliberation. Romele's extension of this concept to the digital domain argues that sustained engagement with digital technologies does not merely change what agents do; it reconfigures the dispositions through which they perceive, evaluate, and engage with their worlds. Digital tools are not instruments that agents pick up and put down while remaining unchanged; they are environments that, over time, sediment into the habitus itself - reshaping what feels natural, what registers as effort, what counts as an adequate response to a situation.[51]

In the academic context, this means that the student who consistently uses AI for writing, research synthesis, and argument generation is not merely acquiring a new skill. They are participating in a field of practice whose regularities - speed as the default, fluency as the standard, the outsourcing of difficulty as the norm - gradually recalibrate their dispositions. The threshold of tolerable cognitive effort shifts; the sense of what constitutes a satisfying intellectual engagement changes; the intuition of what a "good text" looks or feels like is shaped by repeated exposure to AI outputs rather than by the struggle of first-person composition. These are not dramatic transformations visible in any single act; they are the slow sedimentation of a digital habitus that may be well-established before the student or institution recognizes what has formed.

An important consequence of the digital habitus perspective is that AI ethics cannot be reduced only to individual preferences and intentions. Even if a student does not use AI with bad intent, if they continually develop AI-dependent writing practices, their threshold of thinking, span of attention, and habits of original expression may gradually change. Likewise, when institutions direct faculty toward AI amid increasing workload and publication pressure, the ethical issue is no longer merely a matter of individual actors' faults; the institutional habitus itself has changed. Contemporary philosophical extensions, therefore, require us to think about the AI problem at the intersection of subjective intention and structural conditions.

Taken together, phenomenology, Heidegger, and digital habitus yield the following fundamental principle: AI does not merely accelerate the academic world; it discloses and shapes it in particular ways. It may become possible to write texts more quickly, answer questions more rapidly, and classify literature more efficiently, but the cost can be the thinning of experiential depth, the loss of the slowness of thinking, and an attenuation of direct relation to truth. For this reason, AI integration must be approached not only as a matter of skill acquisition but also as a question of how the world and the academic subject are being reconfigured.

3. Application to the Academic Context: Integrity, Authorship, Assessment, Pedagogy, and Governance

The philosophical model developed across the preceding sections has established a set of interlocking normative criteria. Autonomy, in the Kantian sense, demands that the rational subject remain the final author of its own judgment and that the epistemic process be honestly represented. Virtue, in the Aristotelian sense, demands that AI use strengthen rather than atrophy the intellectual dispositions - patience, attentiveness, tolerance of difficulty - through which genuine academic character forms. Truth, in the Platonic sense, demands that the difference between grounded understanding and its fluent simulation be kept visible and pedagogically active. Methodical self-possession, in the Cartesian sense, demands that doubt, justification, and cognitive responsibility remain with

[51] Alberto Romele, *Digital Hermeneutics: Philosophical Investigations in New Media and Technologies* (New York: Routledge, 2019), esp. ch. 4. For Bourdieu's original formulation, see Pierre Bourdieu, *The Logic of Practice*, trans. Richard Nice. (Stanford: Stanford University Press, 1990), 52–65. Birlik's application of Romele's framework to the AI context is developed in Birlik, "Alberto Romele'nin 'Dijital Habitus' Kavramı," 83–103.

the human subject rather than being progressively assumed by the system. And the Heideggerian and phenomenological extensions have shown that these criteria cannot be applied solely at the level of individual acts: they must also address the structural conditions - the regimes of attention, the institutional incentives, the normalized orientations - within which individual choices are made. The question that follows is how these integrated criteria can be brought to bear on the concrete domains in which universities make decisions about AI.

The normative model established above must now be applied to concrete academic domains so that the discussion of AI is not left at the level of abstract philosophy. The ethical problems encountered in higher education often appear singular and scattered, making their common ground easy to overlook. Yet most of the issues, ranging from academic integrity to authorship, from assessment design to institutional governance, are different faces of deeper questions about autonomy, virtue, truth, intellectual agency, and technological shaping. This section accordingly applies the conceptual axes discussed earlier to the practical domains of higher education. The aim is not to repeat abstract principles, but to establish a philosophically grounded framework for evaluating the concrete problems that universities now face.

3.1. Academic Integrity and the Problem of Plagiarism

The debate on academic integrity is the most visible and fastest-growing area of the literature on generative AI. The reason is clear: the university assesses the student product not merely as a textual object, but as evidence of a particular learning process. Generative AI unsettles this regime of evidence because it can produce, in a short time, texts that are formally quite adequate, thereby loosening the traditional link between product and process.[52] Systematic reviews show that instructors' concern is not merely cheating itself, but, at a deeper level, the difficulty of recognizing student labor, distinguishing original contribution, and preserving assessment fairness.

Universities must therefore redefine the very concept of integrity. The classical regime of plagiarism was built around taking another person's text without attribution. In the age of generative AI, by contrast, what often exists is not a stolen human text but still a misrepresented intellectual process.[53] The student may not be plagiarizing directly, but they may nonetheless be misleading the evaluator by presenting the text as though it were the natural result of their own reasoning. Here, "originality" no longer means uniqueness alone, but the truthful declaration of labor and agency. In the new period, policies on academic integrity must therefore be designed not solely through similarity detection, but through transparency of contribution and visibility of process.

The pedagogical consequence is equally clear: a detection-centered approach is not sustainable on its own.[54] Technically, the reliability of AI detection tools remains contested, and a regime grounded solely in detection risks damaging the relationship of trust between instructor and student. What is needed instead is a pedagogical regime of integrity that renders AI use visible, classifies levels of use, and explains which forms of assistance are legitimate and which are illegitimate in given tasks. Ethical learning becomes possible not when the student is told only that something is forbidden, but when they are shown why it is prohibited or conditionally permissible.[55] Academic integrity must rest on justification, not on fear.

[52] Gallent-Torres et al., "The impact of generative artificial intelligence in higher education: A focus on ethics and academic integrity".
[53] Bittle and El-Gayar, "Generative AI and academic integrity in higher education: A systematic review and research agenda".
[54] Liu et al., "Responding to generative AI in Australian higher education".
[55] E. M Berba, "Navigating the Landscape of Generative AI in Higher Education: Institutional Policies, Workforce Readiness, and Ethical Guidelines," *Advances in Transdisciplinary Engineering* (2025). https://doi.org/10.3233/ATDE250550

3.2. Authorship, Contribution, and the Obligation of Disclosure

The question of authorship is one of the most delicate knots in academia in the age of generative AI. An academic text is not merely an arrangement of words; it is a practice of deciding, selecting, sorting, taking responsibility, and standing behind a claim.[56] For this reason, the question "who wrote the text?" is not a simple question of formal ownership. The real issue is who bears the conceptual load, the argumentative structure, and the burden of verification. As Askari emphasizes, panic in the face of AI sometimes narrows the matter too much; the more productive approach is to rethink authorship itself: what is the indispensable share of the human in the making of a text, and how is that share to be made visible?[57]

Disclosure here is not merely a gesture of transparency; it is the ethical condition of truthful representation. When the researcher or student does not explain at what stage and to what extent generative AI was used, the reader may take the text to be more human-centered and more directly the product of intellectual labor than it really is. This distorts both evaluation and trust relations. Pinho and colleagues' proposals for ethical AI use in research show that disclosure must be understood not as an institutional formality, but as a practice that preserves epistemic trust.[58] In the same way, debates on academic publishing stress that the human author cannot delegate the tasks of final verification, source control, interpretation, and responsibility.

The reasonable solution, therefore, is not to turn AI into an invisible author, but to explicitly define levels of contribution. Universities and journals must develop more refined disclosure categories than the binary of "prohibited/permitted": different types of use, such as linguistic editing, summarization, idea generation, structuring, coding assistance, data organization, and visual production, should be distinguished, and the relation of each to authorship should be clearly specified. Such an approach goes beyond both the naiveté that sanctifies AI and the exaggerated reaction that treats every use as hidden misconduct. The core principle is this: authorship is not the mere production of words on the screen, but the bearing of intellectual load and responsibility.

3.3. Assessment and the Problem of Validity

Assessment is one of the areas in which generative AI has created the most visible institutional disruption in higher education. Classical assignment and essay formats were built on the assumption of a relatively stable link between the student's written product and their level of learning. AI has partially severed that link. A text that appears technically strong may no longer reliably represent the student's level of understanding. This does not reduce assessment merely to detect cheating; more deeply, it becomes a problem of validity. As Sibilin notes, the crisis is not only practical but epistemological: how certain can the educator be about what is actually being measured?[59]

Discussions of assessment redesign, therefore, become more important than defensive reflexes in the age of AI. A defensive approach seeks to make assignments "ChatGPT-proof"; a more productive approach designs assessment forms that make the student's thinking process, decisions, and conceptual justifications visible.[60] Oral defenses, draft chains, process reports, explanations of source selection, reflective texts evaluating one's own mode of use, and writing tasks tied to in-class discussion matter for this reason. Their aim is not merely to make cheating more difficult, but to place the student's reasoning and learning process at the center of evaluation.

---

[56] Roberts, "Addressing the novel implications of generative AI for academic publishing, education, and research".

[57] Askari, "Reclaiming authorship in the age of generative AI: From panic to possibility".

[58] Pinho et al., "Ethical and responsible use of GenAI in research context".; Roberts, "Addressing the novel implications of generative AI for academic publishing, education, and research".

[59] Sibilin, "Education and the epistemological crisis in the age of ChatGPT".

[60] N. Aghaee, J. Vrågård and F. Brorsson, "Generative AI in Higher Education: Educators' Perspectives on Academic Learning and Integrity," (2024).

Not every form of assessment is appropriate for every course, however. Universities must therefore reconstruct the relation between normative aims and modes of assessment. If the aim of a course is content recall, one assessment logic is appropriate; if the aim is conceptual analysis, critical judgment, or methodological justification, a different logic is required. In the age of AI, good assessment is not obsessed with discovering whether the student used the system; it is an assessment that can distinguish what the student themselves has thought, even if they did use it. Such a design must carry together Kantian autonomy, Aristotelian intellectual character, Platonic criteria of authentic learning, and the Cartesian insistence that the student remains the methodical owner of their own cognitive process - not merely the recipient of its outputs.

3.4. Pedagogy, Learning, and Intellectual Formation

From a pedagogical perspective, generative AI is not simply a new auxiliary tool but an event that tests our very understanding of teaching and learning. If education is understood as the transmission of information and the retrieval of polished products, AI appears as a highly successful ally. But if education aims to cultivate the student's capacity to ask questions, develop intellectual courage, cope with uncertainty, and form a voice of their own, then AI integration must be designed far more carefully. For this reason, the center of pedagogical debate should not be tool selection, but the question of what learning itself is.

The crucial distinction in instructional design is whether AI functions as a substitute for thinking or as a provocation to think. In the first case, the student receives a ready-made structure and a ready-made formulation in place of their own conceptual labor; in the second, AI compels the student to think through comparison, feedback, and the generation of alternatives. Good pedagogical design pushes the student toward the second path. For example, one may ask the student to critique the AI-generated answer, identify its errors, compare it to their own approach, and explain why certain suggestions are rejected. Tasks of this sort transform AI from a producer of final answers into an object of critical dialogue.

For this reason, metacognitive components become central to pedagogy in the age of AI. Students must be asked to explain not only what they learned, but how they learned it, at what points they turned to AI, which suggestions they accepted and which they rejected, and why. In this way, learning becomes less visible as the delivery of an external text and more as a self-regulated and reflective process. Pedagogy need not be weakened by AI use, but for that, deliberate designs that render the student's intellectual agency visible are required.

3.5. Research, Academic Writing, and Knowledge Production

In the domains of research, academic writing, and knowledge production, generative AI raises much more complex ethical questions. Here, the issue is not only the student's assignment, but the reliability and public credibility of scientific discourse itself. AI can provide researchers with significant speed gains in areas such as idea generation, summarization, literature mapping, structural suggestions, linguistic correction, and visualization. Yet this speed comes with risks such as fake sources, misleading summaries, commentary that implies fabricated data, methodological opacity, and the thinning of intellectual labor. Especially in fields under intense pressure to produce rapidly, AI can cease to be a tool supporting research and become a machine for increasing publication efficiency.

A reasonable stance with regard to the future of academic writing is therefore neither to treat AI as a hidden helper nor as a new co-author, but to position it as a support instrument whose limits are explicit, whose contribution is declared, and whose every stage is verified by a human being.[61] Scientific writing is not merely the production

[61] Kotsis, "The role of ChatGPT in academic writing: Pedagogical and ethical dimensions".; Gabay et al., "Generative artificial intelligence (GenAI) for academic writing in higher education: A scoping review of applications, challenges, and implications".

of grammatically proper sentences; it is a practice of questioning, interpretation, and epistemic risk-taking. AI may enrich that practice, but it cannot replace it. For this reason, research institutions, journals, and graduate programs should develop standards of explicit disclosure, record-keeping, and verification for AI-assisted writing.

3.6. Institutional Governance and University Responsibility

Institutional governance is at least as important as the pedagogical and ethical dimensions of the AI debate, because individual practices of use are often shaped by institutional cultures, assessment regimes, and incentive structures. If universities expect honesty from students while, at the same time, indirectly pushing faculty toward AI use under conditions of increased workload, publication pressure, and demands for speed, then the problem cannot be explained as a mere deficiency in individual morality. AI governance must therefore be understood not simply as a disciplinary mechanism activated when violations occur, but as an integrated field of policy that defines human and academic values from the outset.

Orfanidis's comparative analysis shows that university policies do not proceed along a single universal ethical line: institutions at times legitimize AI use through frameworks of "moral consistency" and at times through "responsible future-readiness," while the authority granted to faculty takes absolute, limited, or hybrid forms.[62] This finding is important because university policies are not merely technical guidelines; they are normative documents that declare which moral priorities the institution adopts. As Coates and colleagues emphasize, effective governance becomes meaningful only when it supports that normative field with appropriate systems of indicators, channels of accountability, and architectures of education and assessment.[63]

Global reviews of guidance documents and institutional policies show that effective governance takes shape along several basic axes: classification of use levels, disclosure requirements, data security, privacy, copyright and intellectual property, principles of assessment, faculty training, student awareness, and the distribution of institutional responsibility.[64] Yet even when these documents are useful at the application level, most do not clearly state which value they protect and why. Good governance does not merely produce procedures; it renders visible foundational principles such as autonomy, accuracy, equality, responsibility, and human dignity. If a policy document lacks pedagogical and ethical language, rule lists quickly become mechanical and lose legitimacy.

For this reason, universities' first need is to develop a shared conceptual language for AI. Unless such concepts as "prohibited," "permitted," "assistance," "contribution," "substitution," "originality," and "responsibility" are defined consistently across faculty, students, and administrators, practices will remain scattered. The second need is contextual flexibility. It cannot be expected that the same modes of use will produce the same outcomes in engineering, philosophy, creative writing, and data analysis. Institutional frameworks should therefore set common principles while leaving flexible spaces that disciplines can concretize according to their own learning aims. The third need is continuing ethical education. AI literacy does not mean only learning how to use tools; it also includes discussing the intellectual and moral consequences of the relations formed with those tools.

Institutional governance must also take inequality and access problems into account. AI tools are unequally distributed across paid versions, linguistic differences, device access, and levels of digital competence. If institutions ignore these differences and judge student products by a single uniform standard, they may, in fact, deepen inequalities. Fair governance, therefore, means not only setting rules but also aiming to equalize access conditions,

[62] C. Orfanidis, "Moral Diversity in Institutional Policies Governing the Student Usage of Generative AI: An International Comparison," *Higher Education Quarterly* 79, (2025): e70051. https://doi.org/10.1111/hequ.70051

[63] H. Coates, G. Croucher and A. Calderon, "Governing Academic Integrity: Ensuring the Authenticity of Higher Thinking in the Era of Generative Artificial Intelligence," *Journal of Academic Ethics* 23, (2025): 2015-28. https://doi.org/10.1007/s10805-025-09639-7

[64] Berba, "Navigating the landscape of generative AI in higher education: Institutional policies, workforce readiness, and ethical guidelines".

provide support, offer training, and ensure transparency. AI ethics cannot be thought independently of technical infrastructure.

Taken together, these discussions lead to the following conclusion: the university must act in the face of AI not merely as an institution defending itself, but as an institution that rearticulates its own normative purpose. Good governance names human ends explicitly without demonizing or neutralizing technology. Among these ends are at least the following: preserving the student's courage to use their own reason; nurturing intellectual character; safeguarding the distinction between truth and appearance; ensuring the truthful representation of intellectual labor and contribution; and preserving public trust in research and teaching.

Conclusion

The argument developed in this article has moved through several distinct philosophical registers, and it is worth pausing at the end to ask what, taken together, they establish - and what they leave open.

The most significant finding is not any single criterion for evaluating AI use, but the demonstration that the normative language currently dominant in the literature is structurally insufficient for the problem it faces. When educational researchers and policymakers speak of "academic integrity," "authentic learning," or "autonomous thinking," they are, as this article has shown, drawing on philosophical traditions - Kantian, Aristotelian, Platonic, Cartesian - whose internal logic they do not always acknowledge and whose tensions they do not always resolve. The contribution of an explicitly philosophical approach is not to replace this language but to systematize it: to show that these concerns are not merely intuitions but the surface manifestations of deep normative commitments, and that those commitments, when made explicit, yield more precise and more defensible criteria for institutional practice.

A second finding concerns the relation between the classical and contemporary philosophical axes. The dominant tendency in the applied ethics literature is to treat Kantian autonomy as the master concept for evaluating AI in education - the question of whether the student "really" thought for themselves. This article has argued that autonomy, while indispensable, is insufficient on three grounds. First, it operates at the level of discrete acts and misses the longer arc of character formation that Aristotelian ethics illuminates. Second, it presupposes a stable distinction between knowledge and its appearance that requires the Platonic epistemic framework to articulate and defend. Third, and most fundamentally, it operates within a conception of the individual rational subject that Heidegger's analysis of Enframing reveals to be part of the very metaphysical trajectory that makes the AI problem what it is. A normative model that relies on Kantian autonomy alone risks deploying, as its corrective, the same logic of self-grounding subjectivity that technological Enframing has already colonized. The integration of Heidegger and phenomenology into the framework is therefore not an ornament; it is a structural necessity.

A third finding concerns the concept of digital habitus. The existing literature tends to frame AI ethics as a problem of individual choice - did the student use AI appropriately or not? The Bourdieusian dimension introduced through Romele's concept shifts the analysis to the level of pre-reflective disposition: what the student takes for granted, what registers as effort, what feels like adequate intellectual engagement. This shift has a practical implication that the literature has not yet fully absorbed: institutional AI governance cannot be evaluated solely by its explicit rules and prohibition categories. It must also be evaluated by what it normalizes - by the dispositions it sediments into academic practice through the design of assessment regimes, publication incentives, and workload structures. An institution that prohibits AI in student essays while driving faculty toward AI-assisted publication under performance pressure is not maintaining a coherent normative position; it is producing a double habitus whose tensions will surface, and are already surfacing, in academic practice.

Several limitations of the present argument must be acknowledged. The philosophical tradition drawn upon here is overwhelmingly Western and canonical. The normative framework developed - autonomy, virtue, truth, methodical self-possession - carries specific historical and cultural presuppositions that a genuinely global account of AI ethics in higher education would need to interrogate and expand. This is not a minor caveat: the universities most rapidly integrating generative AI are not all operating within the same philosophical inheritance, and a normative model that does not acknowledge this risks reproducing a different kind of epistemic imperialism. Second, the article has focused on the normative question - what ought to count as legitimate use - and has not engaged systematically with the empirical literature on how students actually use AI, what effects this use has on measurable learning outcomes, or how institutional interventions change behavior. The philosophical and empirical programs need each other: the former without the latter risks becoming prescriptively detached; the latter without the former risks becoming normatively blind.

The most pressing direction for future philosophical work is the question of institutional design. This article has argued that the university must articulate its normative purpose explicitly rather than reactively - must name what it is protecting and why, rather than simply responding to each new AI capability as it appears. What this requires, philosophically, is an account of the university's telos that is robust enough to guide policy without being so rigid that it cannot respond to genuinely new conditions. The tension between the university as a site of truth-seeking and the university as a credentialing institution operating within market logics is not new — but generative AI has made it acute in a way that demands philosophical attention rather than purely administrative response. The question of what the university is for, which philosophy of education has long debated, has become, in the age of generative AI, an urgent practical question that institutions can no longer defer.

The real issue, then, is not what generative AI will do to the university. It is whether the university has, and can articulate, a sufficiently clear account of what it is trying to do — clear enough to determine, in each case, whether a given form of AI integration serves or undermines it. That is a philosophical question before it is a policy question. And it is one that this article has attempted, at least in part, to bring into sharper focus.

Declaration on the Use of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, SciSpace was used to support the literature review and to organize and refine the text. The author (Dursun Eşsiz) critically examined, verified, and edited all AI-assisted outputs and assumes full responsibility for the article's final content.

References

Aghaee, N., J. Vrågård and F. Brorsson. "Generative AI in Higher Education: Educators' Perspectives on Academic Learning and Integrity," (2024).

Aristotle. "Nicomachean Ethics" in *The Complete Works of Aristotle*, Edited by Jonathan Barnes. vol. 2. Princeton: Princeton University Press, 1984.

Askari, M. "Reclaiming Authorship in the Age of Generative AI: From Panic to Possibility," *AI Magazine* (2025). https://doi.org/10.1002/aaai.70022

Barus, O. P., A. N. Hidayanto and I. Eitiveni. "Mapping Generative AI's Ethical Issues in Higher Education: A Felt-Guided Systematic Review," *Polyglot: Jurnal Ilmiah* (2025). https://doi.org/10.19166/pji.v21i2.10020

Bekalp, B. "Fenomenolojik Perspektiften Yapay Zekâ: Yönelimsellik, Beden ve Çerçeve Problemi," *MetaZihin* 8, no. 1 (2025): 25-56. https://doi.org/10.51404/metazihin.1695620

Berba, E. M. "Navigating the Landscape of Generative AI in Higher Education: Institutional Policies, Workforce Readiness, and Ethical Guidelines," *Advances in Transdisciplinary Engineering* (2025). https://doi.org/10.3233/ATDE250550

Birlik, B. A. "Alberto Romele'nin 'Dijital Habitus' Kavramı ve Val Dusek'in Yapay Zekâ Felsefesi Bağlamında İnsanın Mevcudiyet Krizi," *MetaZihin* 8, no. 1 (2025): 83-103. https://doi.org/10.51404/metazihin.1674795

Bittle, K. and O. El-Gayar. "Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda," *Information* 16, (2025): 296. https://doi.org/10.3390/info16040296

Bourdieu, Pierre. *The Logic of Practice*. Translated by Richard Nice, Stanford: Stanford University Press, 1990.

Coates, H., G. Croucher and A. Calderon. "Governing Academic Integrity: Ensuring the Authenticity of Higher Thinking in the Era of Generative Artificial Intelligence," *Journal of Academic Ethics* 23, (2025): 2015-28. https://doi.org/10.1007/s10805-025-09639-7

Crawford, J., M. Cowling and K.-A. Allen. "Leadership Is Needed for Ethical ChatGPT: Character, Assessment, and Learning Using Artificial Intelligence (AI)," *Journal of University Teaching & Learning Practice* 20, no. 3 (2023): Article 02. https://doi.org/10.53761/1.20.3.02

Dempere, J., K. P. Modugu, A. Hesham and L. K. Ramasamy. "The Impact of ChatGPT on Higher Education," *Frontiers in Education* (2023). https://doi.org/10.3389/feduc.2023.1206936

Descartes, René. *Discourse on Method and Meditations on First Philosophy*. Translated by Donald A. Cress, 4th ed. Indianapolis: Hackett, 1998.

Descartes, René. *Meditations on first philosophy: With selections from the objections and replies*. Oxford University Press, 2008.

Dreyfus, Hubert L. "Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian," *Philosophical Psychology* 20, no. 2 (2007): 247-268.

Dukor, M. and E. Izeji. "Investigating Epistemological Relations Between Descartes' Thinking Mind and Artificial Intelligence Processing," *SSRN preprint*, n.d.

Durmaz, N. "Heidegger'in Felsefesi Bağlamında Yapay Zekâ," *MetaZihin* 8, no. 2 (2025): 175-187. https://doi.org/10.51404/metazihin.1767253

Fırat, M. "Reimagining Descartes' Philosophy of Autodidacticism in the Postdigital Education," *Educational Philosophy and Theory, advance online publication* (2026). https://doi.org/10.1080/00131857.2025.2611000

Francis, N. J., S. Jones and D. P. Smith. "Generative AI in Higher Education: Balancing Innovation and Integrity," *British Journal of Biomedical Science* (2025). https://doi.org/10.1080/09674845.2025.2474477

Gabay, R. A. E., A. A. Funa and J. D. Ricafort. "Generative Artificial Intelligence (GenAI) for Academic Writing in Higher Education: A Scoping Review of Applications, Challenges, and Implications," *International Journal of Education in Mathematics, Science and Technology* 14, (2026): 200-232.

Gallent-Torres, C., A. Zapata-González y J. L. Ortego-Hernando. "The Impact of Generative Artificial Intelligence in Higher Education: A Focus on Ethics and Academic Integrity," *RELIEVE* 29, no. 2 (2023): Article M5. https://doi.org/10.30827/relieve.v29i2.29134

Gimpel, H., K. Hall, S. Decker, T. Eymann and L. Lämmermann. "Unlocking the Power of Generative AI Models and Systems Such as GPT-4 and ChatGPT for Higher Education: A Guide for Students and Lecturers," *Hohenheim Discussion Papers in Business, Economics and Social Sciences* 02-2023.

Guamán Chávez, R. E. "Ética e Integridad Académica en el Uso de la Inteligencia Artificial Generativa en la Educación Superior," *Revista Científica Multidisciplinar G-NER@NDO* (2025). https://doi.org/10.60100/rcmg.v6i1.392

Guleria, A., K. Krishan, V. Sharma and T. Kanchan. "ChatGPT: Ethical Concerns and Challenges in Academics and Research," *Journal of Infection in Developing Countries* (2023). https://doi.org/10.3855/jidc.18738

Gupta, V. and A. Nyamapfene. "Generative AI in Universities: Practices at UCL and Other Institutions, and the Path Forward," *Internet Reference Services Quarterly* 29, no. 1 (2025): 131-51. https://doi.org/10.1080/10875301.2025.2453461

Heidegger, Martin. "The Question Concerning Technology" in *The Question Concerning Technology and Other Essays*, Translated by William Lovitt, 3-35. New York: Harper & Row, 1977.

Heidegger, Martin. *Being and Time*. Translated by John Macquarrie and Edward Robinson, Oxford: Blackwell, 1962.

Heidegger, Martin. *What Is Called Thinking?* Translated by J. Glenn Gray, New York: Harper & Row, 1968.

Huang, A. Y. Q., J.-W. Tzeng, C.-S. Huang, Z.-Q. Liu, B. C. Tanujaya and O. H. T. Lu. "A Case Study for Educators with ChatGPT and Plato's Allegory of the Cave" in *Proceedings of the 32nd International Conference on Computers in Education*, Edited by A. Kashihara et al., 2024. https://doi.org/10.58459/icce.2024.4960

Husserl, Edmund. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, First Book*. Translated by F. Kersten, Dordrecht: Kluwer, 1983.

Hyland, T. "Educational Responses to Artificial Intelligence (AI) Applications: Problems and Promise," *Qeios* (2023). https://doi.org/10.32388/08UCQU

Jiao, J., S. Afroogh, K. J. Chen, D. Atkinson and A. Dhurandhar. "The Global Landscape of Academic Guidelines for Generative AI and Large Language Models," *arXiv preprint* (2024). https://doi.org/10.48550/arXiv.2406.18842

Kadwa, M. S. "Written Assignments and the Ethical Considerations of Artificial Intelligence in Higher Education," *International Journal of Linguistics, Literature and Translation* (2025). https://doi.org/10.32996/ijllt.2025.8.9.27

Kant, Immanuel. "An Answer to the Question: What is Enlightenment?" (1784) in *Kant: Political Writings*, Edited by Hans Reiss, 54-60. 2nd ed. Cambridge: Cambridge University Press, 1991.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor and Jens Timmermann, Cambridge: Cambridge University Press, 2012.

Kotsis, K. T. "The Role of ChatGPT in Academic Writing: Pedagogical and Ethical Dimensions," *EthAIca* 4, (2025): 420. https://doi.org/10.56294/ai2025420

Liu, D., T. Fawns, M. Cowling and A. J. Bridgeman. "Responding to Generative AI in Australian Higher Education," *OSF Preprints* (2023). https://doi.org/10.35542/osf.io/9wa8p

Loos, E. and J. Radicke. "Using ChatGPT-3 as a Writing Tool: An Educational Assistant or a Moral Hazard? Current ChatGPT-3 Media Representations Compared to Plato's Critical Stance on Writing in Phaedrus," *AI and Ethics* 5,

(2024): 1133-46. https://doi.org/10.1007/s43681-024-00470-1

Merleau-Ponty, Maurice. *Phenomenology of Perception*. Translated by Colin Smith, London: Routledge, 2002.

Meyer, J. G., R. J. Urbanowicz, P. C. N. Martin, K. O'Connor and R. Li. "ChatGPT and Large Language Models in Academia: Opportunities and Challenges," *Biodata Mining* 16, (2023): 20. https://doi.org/10.1186/s13040-023-00339-9

Michel-Villarreal, R., E. L. Vilalta-Perdomo, D. E. Salinas-Navarro, R. Thierry-Aguilera and F. S. Gerardou. "Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT," *Education Sciences* 13, (2023): 856. https://doi.org/10.3390/educsci13090856

Moran, Dermot. *Introduction to Phenomenology*. London: Routledge, 2000.

Noviandy, T. R., A. Maulana, G. M. Idroes, Z. Zahriah and M. Paristiowati. "Embrace, Don't Avoid: Reimagining Higher Education with Generative Artificial Intelligence," *Journal of Social Research* (2024). https://doi.org/10.60084/jeml.v2i2.233

Orfanidis, C. "Moral Diversity in Institutional Policies Governing the Student Usage of Generative AI: An International Comparison," *Higher Education Quarterly* 79, (2025): e70051. https://doi.org/10.1111/hequ.70051

Petricini, M. "What Would Aristotle Do? Navigating Generative Artificial Intelligence in Higher Education," *Postdigital Science and Education* 6, (2024). https://doi.org/10.1007/s42438-024-00463-7

Pinho, M. I. G. de, A. P. D. Costa e C. G. de Pinho. "Ethical and Responsible Use of GenAI in Research Context," *Práxis Educativa* 20, (2025): e25388. https://doi.org/10.5212/PraxEduc.v.20.25388.057

Plato. "Meno" in *Plato: Complete Works*, Edited by John M. Cooper, 870-897. Indianapolis: Hackett, 1997.

Plato. *Phaedrus*. Translated by S. Scull, Newburyport, MA: Focus Publishing, 2003.

Plato. "Republic" in *Plato: Complete Works*, Edited by John M. Cooper, 971-1223. Indianapolis: Hackett, 1997.

Qadhi, S. M., A. Alduais, Y. Chaaban and M. Khraisheh. "Generative AI, Research Ethics, and Higher Education Research: Insights from a Scientometric Analysis," *Information* 15, (2024): 325. https://doi.org/10.3390/info15060325

Roberts, L. W. "Addressing the Novel Implications of Generative AI for Academic Publishing, Education, and Research," *Academic Medicine* (2024). https://doi.org/10.1097/ACM.0000000000005667

Romele, Alberto. *Digital Hermeneutics: Philosophical Investigations in New Media and Technologies*. New York: Routledge, 2019.

Sibilin, C. S. "Education and the Epistemological Crisis in the Age of ChatGPT," *Critical Review* 35, no. 4 (2023): 414-25. https://doi.org/10.1080/08913811.2023.2284042

Strunk, V. and J. Willis. "Generative Artificial Intelligence and Education: A Brief Ethical Reflection on Autonomy," *EDUCAUSE Review* 13, (January 2025).

Thomson, Iain D. *Heidegger on Technology's Danger and Promise in the Age of AI*. Cambridge: Cambridge University Press, 2025.

Vieira, A. and A. Mesquita. "Generative Artificial Intelligence in Higher Education: Challenges, Opportunities and Pedagogical Implications," *Journal of Technologies Information and Communication* 5, no. 1 (2025): 36578. https://doi.org/10.55267/rtic/16675

Watanabe, A. "Have Courage to Use Your Own Mind, with or without AI: The Relevance of Kant's Enlightenment to Higher Education in the Age of Artificial Intelligence," *Electronic Journal of e-Learning* 22, no. 2 (2024): 46-58. https://doi.org/10.34190/ejel.21.5.3229

Wood, Allen W. "Kant's Ethical Thought" in *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, Edited by Paul Guyer. Lanham: Rowman & Littlefield, 1998.

Xames, M. D. A. G. and A. A. Grado. "ChatGPT for Research and Publication: Opportunities and Challenges," *Journal of Applied Learning & Teaching* 6, no. 1 (2023). https://doi.org/10.37074/jalt.2023.6.1.20