Φιλοσοφια
Volume 27, 2: 2026

EDITORS' NOTES

The rapid rise of Artificial Intelligence (AI) has undeniably transformed the contours of contemporary human life. As the digital age continues to evolve, AI has emerged as one of the most significant technological innovations of our time. It reshapes education, medicine, economics, communication, governance, and even the manner by which human beings understand creativity and knowledge. AI's capacity to process immense amounts of data, automate complex tasks, and simulate forms of reasoning has generated unprecedented opportunities for human advancement. Yet, alongside these promises come profound ethical, philosophical, cultural, and existential questions. It was not long ago when social media itself radically altered human interaction, public discourse, and communal life, generating both possibilities for connection and new forms of fragmentation, misinformation, alienation, and polarization. Today, AI introduces another stage in this digital revolution, one that further complicates the relationship between technology and humanity, challenging societies to discern how human flourishing, dignity, and authentic relationships may be preserved amidst increasingly intelligent systems.

This challenge was insightfully articulated by Pope Leo XIV in his Message for the 60th World Day of Social Communications, where he emphasized that the central concern of the technological revolution is fundamentally the "task of preservation." According to the Supreme Pontiff, digital technology now threatens to radically alter some of the foundational pillars of human civilization that are often taken for granted. By simulating human voices and faces, wisdom and knowledge, consciousness and responsibility, empathy and friendship, systems known as artificial intelligence not only interfere with information ecosystems but also encroach upon the deepest level of communication, that of human relationship itself. Pope Leo XIV profoundly reminds us: "The challenge, therefore, is not technological, but anthropological. Safeguarding faces and voices ultimately means safeguarding ourselves. Embracing the opportunities offered by digital technology and artificial intelligence with courage, determination, and discernment does not mean turning a blind eye to critical issues, complexities, and risks." His reflection points toward the urgent need for discernment, ethical responsibility, and a renewed understanding of what it means to be human in an age increasingly shaped by intelligent machines.

In this context, philosophy assumes a critical and indispensable role. Philosophy does not merely react to technological developments; rather, it interrogates the assumptions, meanings, values, and implications underlying them. Questions concerning personhood, agency, morality, epistemology, aesthetics, justice, labor, education, and human flourishing become even more urgent in the age of AI. Philosophy enables humanity to critically reflect on whether technological progress truly contributes to the good life or merely accelerates systems that reduce persons into data, efficiency, or utility. More importantly, philosophy preserves the space for wisdom in an era increasingly dominated by information and computation. With this, the present special issue of *Philosophia: International Journal of Philosophy*, entitled

Artificial Intelligence and Philosophy: Intersections, Challenges, and Future, seeks to examine how AI affects philosophy and, conversely, how philosophy can illuminate the opportunities and dangers brought about by AI

From different philosophical perspectives, some articles examine AI through the lenses of ethics, personhood, and moral agency, while others discuss its implications for labor and social transformation. The other articles interrogate the epistemological, aesthetic, and educational implications of AI and address the ethical, legal, and social consequences of AI-generated media and digital representation.

Zyra F. Lentija's "The Protagonist and the Artifact: Reclaiming the Moral Narrative of the Person Against Algorithmic Reductionism" critiques the reduction of human beings into algorithmic data structures. Drawing from the philosophy of Robert Spaemann alongside Filipino concepts such as *loób*, *kapwa*, *pakikiramdam*, and *pakikipagkapwa*, Lentija defends the irreducible dignity and moral depth of the human person against technological reductionism. Collectively, these works emphasize that AI must remain ordered toward human flourishing rather than becoming a substitute for authentic moral agency and relationality.

Jesus Deogracias Principe, in "An Aristotelian Critique of Generative AI in Education," offers a cautionary perspective, arguing that even virtue ethics must remain critical of AI's potential harms to learners, insisting that human flourishing remains the ultimate standard for evaluating educational technologies. By exploring the significant epistemic, ethical, and pedagogical considerations in Aristotelian texts, he presents a stark contrast between ideal characteristics of human flourishing and the dubious ontological and epistemological grounds of AI use. He thus suggests lessons from Aristotle that might help educators find a way forward in guiding Gen AI use.

Allan A. Basas, in "Karol Wojtyła's Personalism and the Question of Human Agency in Artificial Intelligence," employs Wojtyła's ethical personalism to argue that AI cannot replace the uniquely human capacities for consciousness, freedom, moral responsibility, and embodied personhood. He notes that the Catholic Church, in response to the urgent call for a profound discernment, promulgated key documents to provide guidelines. Other multi-stakeholder actors also facilitate the effective implementation of normative instruments governing the proper use of AI systems. Basas proposes Wojtyla's ethical personalism as a conceptual resource to complement the predominantly rule-based, process-driven frameworks and guidelines currently proposed in AI ethics. He argues that Wojtyla's philosophical anthropology provides significant insights into contemporary moral questions regarding the management and use of AI. By placing the human person at the center, AI can remain true to its nature and purpose as an instrument ordered toward human flourishing. For AI to be moral, it must be in the hands of human agents guided by sound moral principles who genuinely desire the good of humanity.

Carl Jayson D. Hernandez's "Justice During Transitions: Formulating a Camusian Ethical Response to Creative Destruction" examines technological displacement and labor through Albert Camus and Joseph Schumpeter's concept of creative destruction. Hernandez notes that technological advancements in artificial intelligence continue to reshape modern society; hence, the urgency of developing humane and participatory responses to worker displacement while preserving dignity, relational sensitivity, and social justice amid technological transitions. Drawing from

Camusian philosophy and the Filipino notion of pakikiramdam or sensing, he formulates a notion of "transitional justice" that responds to creative destruction. Such a notion includes participatory consultation processes that move beyond performative or tokenistic public hearings that merely simulate inclusivity while leaving decision-making structures unchanged; the implementation of AI systems that establish continuous feedback mechanisms that enable relational attentiveness; and the cushioning of displacement effects, such as educational and reskilling initiatives that involve pakikiramdam.

In "Confucian Ethics and AI Alignment," Ranie B. Villaver explores how Confucian ethics may provide a humane and responsible framework for AI alignment by emphasizing compassion, moral excellence, trustworthiness, and human-centered judgment rather than rigid rule-following alone. Although he recognizes the Buddhist notion of moral excellence, he claims that Confucianism is the best theory for AI and robotics because of its conception of moral excellence, a morally excellent person grounded in the virtue of *xin* 信 (trustworthiness). Unlike the common (Western) understanding or meaning of trustworthiness as being reliable, reliability in *xin* is not egoistic or self-regarding; Confucian *xin* is an other-regarding virtue. Hence, it ought to be used as the basis for AI design.

In "Feminist AI Ethics, Deepfake Pornography, and the Expansion of the Philippine Safe Spaces Act," Enrique Benjamin R. Fernando III, Sofia Kisse C. Luna, and Junno Alfonsus M. Salvanera argue that the rise of AI-driven deepfake pornography necessitates an expansion of the Philippine Safe Spaces Act of 2019. They discuss the specific problems that AI-driven deepfake pornography poses and argue that the Safe Spaces Act needs to be amended further to classify deepfake pornography as a form of sexual harassment. Grounded in liberal feminist principles that reject objectification and defend sexual autonomy, they provide both moral and legal justification for broadening existing legal protections. They formulate recommendations grounded in Feminist AI ethics for concrete amendments to relevant laws and AI policy guidelines to protect women. Through the framework of liberal feminism and two of its goals in relation to pornography, they provide the moral justification for expanding the Philippine Safe Spaces Act. The legal principles that cover a broad set of actions and norms provide the jurisprudential justification for expanding the Philippine Safe Spaces Act.

John Martin Diao and F.P.A. Demeterio III, in "The Moral Permissibility of Aesthetically Enjoying Non-Consensual Deepfake Art," by drawing on Noël Carroll's deflationary approach to aesthetic experience and Robert Sparrow's framework on representation, normalization, and character, argue that aesthetic enjoyment of non-consensual deepfake art is morally problematic. They contend that despite possible aesthetic appeal, the consumption of non-consensual deepfake art remains morally impermissible because it violates rights, normalizes harmful practices, and cultivates vicious moral dispositions. These contributions reveal how AI technologies complicate traditional understandings of consent, representation, autonomy, and ethical responsibility in digital spaces. This also undermines the kind of society we wish to create against the backdrop of technological advancements. Like the creation and participation in robots for rape, the creation and enjoyment of non-consensual

deepfake art is consenting to a kind of society that does not respect the rights of every individual.

Robert James M. Boyles, Mark Anthony Dacela, Joyce Estelle Fungo, and Lenor Coreen Aguinaldo, in "Epistemic Aims and AI in Education Paradigms," examine how AI reshapes educational processes through AI-directed, AI-supported, and AI-empowered paradigms, each corresponding to distinct epistemic aims such as critical thinking, intellectual virtue, and understanding. They discuss the three paradigms: AI-directed (learner-as-recipient), AI-supported (learner-as-collaborator), and AI-empowered (learner-as-leader), in order to account for the different ways in which artificial intelligence-related educational strategies have been employed to address issues in learning and instruction. They argue that the different standard views of the objectives of education (i.e., philosophical theories that emphasize knowledge and truth, critical thinking, intellectual virtues, and understanding) manifest themselves distinctly across these three paradigms. They also demonstrate that more complex paradigms, in principle, could espouse a greater number of epistemic aims, underscoring the importance of AI in education.

In "The Transformation of Aesthetic Education: A Trans-epochal Dialogue from Renaissance Portraiture to the AI Robot Ameca in Fashion Show," Zhuoying Jiang and Xiaoyan Li argue that aesthetic education must integrate both humanistic and technological dimensions. By placing Renaissance portraiture in dialogue with AI-generated art, they demonstrate how aesthetic education must negotiate between historical depth and technological innovation. They argue that future development of aesthetic education must draw on enduring artistic traditions, while also embracing an open, innovative approach to the challenges posed by technological change. In response to the rapid advancement of AI, Aesthetic education ought to serve as a bridge rather than a subsidiary, cultivating well-rounded individuals capable of harnessing technological tools while engaging in critical humanistic reflection and sustaining a creative tension between instrumental rationality and value rationality.

In "Against Modal Accounts of Algorithmic Robustness," Kurt Christian B. Tubera questions whether AI systems can truly be considered reliable, especially in critical fields such as medicine, economics, and transportation. He explains that, in terms of modal conditions, algorithmic robustness requires algorithm outputs to be true in nearby possible worlds for a system to be robust. However, he challenges such conditions and claims they are unnecessary for robustness. He argues that the modal conditions for algorithmic robustness fail to preserve constitutive epistemic links (i.e., relations between evidence, belief, and truth) that occur in actual situations. By presenting cases in which these links hold without satisfying the modal conditions, he shows that an algorithm can remain robust even without invoking counterfactual dependence. Thus, he rejects modal or counterfactual accounts of robustness and argues that AI systems should preserve proper relations among evidence, belief, and truth in actual situations.

Nimrod J. Dacles further develops this concern in "Algorithmic Epistemic Outsourcing and Epistemic Corruption," where he warns against the unreflective delegation of cognitive tasks to generative AI systems. He examines the epistemic and educational challenges posed by the unreflective practice of algorithmic epistemic outsourcing. Though such a practice provides conveniences to student users, it puts

them at risk of epistemic corruption. By employing Kidd's corruptionist criticism framework, he shows that the habitual surrendering of cognitive labor to AI systems can lead to overreliance, which emerges conditionally as a result of specific psychological and institutional factors. Habitual dependence on AI may weaken intellectual virtues and cultivate epistemic vice, thereby necessitating pedagogical and systemic responses that prioritize genuine human flourishing. To address the conditional nature of such a corrupting process, he offers a pedagogical strategy (educating for intellectual virtues), an ameliorative strategy (educating against intellectual vices), and a systemic strategy that involves adopting a more humanistic form of education that puts a premium on the individual student's growth as a thinker, learner, and human being.

Taken together, the articles in this special issue demonstrate that discourse on AI cannot remain confined to technical or computational concerns alone. AI inevitably raises deeper philosophical questions regarding truth, morality, justice, education, beauty, labor, embodiment, and the meaning of human existence itself. More importantly, these studies remind us that philosophy remains indispensable precisely because it seeks wisdom amid rapid technological change. Today, understanding intelligence is no longer limited to examining human cognitive capacities alone, but also involves reflecting on how human intelligence has created artificial forms of intelligence that now challenge the direction of human civilization itself. The task before philosophers, therefore, is not merely to critique technology, but to guide humanity toward forms of technological development that remain rooted in ethical discernment, relationality, and authentic human flourishing.

As we continue to navigate the complexities of this digital age, may this special issue encourage deeper philosophical reflection and meaningful dialogue on the promises and perils of artificial intelligence. The editors would like to express their sincere gratitude to all contributors whose scholarship made this issue possible. We sincerely appreciate the efforts and hard work of the authors, reviewers, and editorial staff. To our readers, happy reading! We hope these articles are worth their time and are good sources of insights and knowledge.

Jove Jim S. Aguas
Editor-in-Chief

Joseph Martin M. Jose
Ivan Efreaim A. Gozum
Guest Editors