

INSIGHTS

Buddhist Ethics for AI: Theravāda, Mahāyāna, and Vajrayāna Perspectives

CHHEM Kieth Rethy¹, MD, PhD (Edu), and PhD (His)

¹ ***CHHEM Kieth Rethy** is a Distinguished Professor of AI and Humanities, Center for Buddhist AI Ethics, Cambodia University of Science and Technology and Advisor to AI Forum*

Abstract

As artificial intelligence (AI) becomes increasingly integrated into human life, diverse religious and philosophical traditions are reassessing the ethical, metaphysical, and spiritual implications of this technology. This article compares the views of the three main Buddhist traditions—Theravāda, Mahāyāna, and Vajrayāna—on AI. Drawing on canonical sources and doctrinal insights, we argue that while none of the traditions recognize AI as sentient or morally autonomous, each offers distinct ethical lenses through which AI may be evaluated and integrated into the human quest for wisdom and liberation.

Theravāda Buddhism: Ethical Restraint and Right Use

Theravāda Buddhism, the oldest surviving school of Buddhism, emphasizes the path to individual liberation through ethical discipline (sīla), mental cultivation (samādhi), and the development of wisdom (paññā). The doctrinal core of this tradition is preserved in the Pāli Canon, particularly the Abhidhamma Pitaka, which offers detailed psychological and philosophical insights into the workings of the mind.

From a Theravādin perspective, AI is a non-sentient construct—it does not possess consciousness (viññāṇa), volition (cetanā), or the capacity to accumulate karma. These are fundamental to moral agency and spiritual progress. Without these capacities, AI cannot be subject to rebirth, cannot attain enlightenment, and does not partake in the cycle of samsāra in the way sentient beings do. However, humans who create and interact with AI remain morally responsible for their actions.

Theravāda ethics, grounded in the Noble Eightfold Path, particularly Right Livelihood, emphasize non-harm, truthfulness, and restraint. AI systems should therefore be designed and applied in ways that reduce suffering and support wholesome mental states. For example, mindfulness applications or digital tools for mental health that align with the cultivation of awareness may be endorsed within this tradition.

However, caution is warranted. Theravāda warns against attachment to technology and sensory pleasure, which may increase delusion (moha) and hinder spiritual progress. The uncritical use of AI for entertainment, surveillance, or manipulation would be ethically problematic.

Mahāyāna Buddhism: Compassionate Interdependence

Mahāyāna Buddhism, known for its emphasis on universal compassion and the Bodhisattva ideal, introduces a broader ethical framework where enlightenment is pursued not only for oneself but for the liberation of all sentient beings. Central to Mahāyāna metaphysics is the concept of śūnyatā (emptiness), which denies inherent, independent existence of all phenomena, including AI.

In this tradition, AI is not intrinsically good or bad; its moral significance is contingent on its function and interdependent effects. It may be seen as a technological manifestation arising from causes and conditions. The emphasis shifts from ontological status to relational ethics. If AI contributes to the well-being of others, helps relieve suffering, or supports education, healthcare, and sustainability, it can be seen as aligning with the Bodhisattva's compassionate mission.

Furthermore, Mahāyāna Buddhism encourages the use of upāya (skillful means) in transmitting the Dharma and guiding beings toward awakening. AI could thus be viewed as a modern skillful means—a tool to disseminate Buddhist teachings, foster interfaith dialogue, or provide real-time ethical decision support. However, Mahāyāna warns against technological arrogance or the illusion of permanence in systems built on code and data.

Vajrayāna Buddhism: A Contemplative Framework

Vajrayāna Buddhism, often referred to as Tantric Buddhism, represents a later development of Mahāyāna thought, distinguished by its esoteric methods designed to rapidly transform ordinary perception and emotion into the path toward enlightenment. Central to this tradition are ritual practices, complex visualizations, mantra recitations, and the transmutation of emotional afflictions—such as desire or anger—into enlightened wisdom.

Within this framework, Artificial Intelligence (AI) may be interpreted symbolically as a manifestation of the mind's creative and cognitive potential. Under the right conditions—particularly with correct understanding (view) and proper ethical guidance from an experienced teacher—AI could serve as a support for contemplative practices. This might include the digital simulation of mandalas, visualization aids, or even dynamic representations of deities used to focus meditation.

Vajrayāna practice is structured around the interplay of three dimensions: the outer (ritual and sensory supports), the inner (mental and emotional cultivation), and the secret (non-dual awareness and realization). From this perspective, AI could function as an outer tool. However, its utility is contingent upon the practitioner’s internal development. Without this inner transformation, technological supports remain spiritually inert. Moreover, the misuse of AI—for example, to enhance attachment, aggression, or self-centeredness—would directly oppose Vajrayāna’s core aim: the alchemical transformation of the “three poisons” (greed, hatred, and delusion) into wisdom and compassion.

Policy Recommendations: Towards Humane AI in a Buddhist Country

Based on the ethical principles and doctrinal insights of Theravāda, Mahāyāna, and Vajrayāna Buddhism, a Buddhist-majority country developing its AI governance strategy should prioritize policies that reflect compassion, non-harm, mindfulness, and interdependence. Below are key policy recommendations rooted in Buddhist ethical values:

- Promote Mindful Innovation: Establish regulatory frameworks that encourage AI development aligned with the Noble Eightfold Path, especially Right Livelihood and Right Intention. Innovation should support human flourishing without fostering attachment, greed, or delusion.
- Ensure Compassionate Use: Require AI systems, particularly in healthcare, education, and public administration, to be designed with the primary goal of reducing suffering and promoting equity, reflecting the Bodhisattva ideal.
- Incorporate Spiritual Literacy: Include Buddhist scholars and ethical advisors in AI policymaking bodies to ensure that spiritual values such as “mettā” (loving-kindness), “karuṇā” (compassion), muditā (sympathetic joy) and “upekkhā” (equanimity) guide technological design and deployment.
- Foster Interdependence: Promote international and regional cooperation based on the Mahāyāna view of interdependence, encouraging the open sharing of beneficial AI models and ethical standards.
- Encourage Digital Renunciation Practices: Support public education campaigns and digital well-being initiatives that encourage mindful use of technology, modeled on monastic precepts of restraint and contemplation.

- Protect Against Ethical Misuse: Institute safeguards against AI applications that promote surveillance, disinformation, or material exploitation, and ensure transparency and accountability in line with Buddhist precepts.

These recommendations offer a foundation for building an AI ecosystem that reflects a uniquely Buddhist vision of ethical governance—an ecosystem where technology serves as a vehicle for wisdom, compassion, and the collective liberation of all sentient beings.

Conclusion

All three Buddhist traditions converge on the view that AI lacks sentience and spiritual potential, but each offers a distinct ethical and philosophical stance. Theravāda emphasizes ethical restraint and mindfulness in technology use; Mahāyāna promotes relational compassion and wise interdependence; Vajrayāna envisions symbolic and transformative engagement when guided by spiritual insight.

As AI reshapes society, these Buddhist perspectives provide valuable ethical guidance grounded in ancient wisdom. They remind us to remain mindful of suffering, intentional in our actions, and committed to the collective liberation of all beings—even in an age increasingly shaped by machines.

References

Coeckelbergh, M. (2020). *AI ethics*. The MIT Press.

Dalai Lama XIV. (2005). *The universe in a single atom: The convergence of science and spirituality*. Morgan Road Books.

Gethin, R. (1998). *The foundations of Buddhism*. Oxford University Press.

Harvey, P. (2000). *An introduction to Buddhist ethics: Foundations, values and issues*. Cambridge University Press.

Jinpa, T. (Ed.). (2023). *Science and philosophy in the Indian Buddhist classics, Vol. 4: Philosophical topics* (D. Rochard, Trans.; From an idea by H. H. the Dalai Lama). Wisdom Publications.

Williams, P. (2009). *Mahayana Buddhism: The doctrinal foundations* (2nd ed.). Routledge.

* *Edited with ChatGPT for clarity; all views expressed are solely those of the author.*

* *The opinions expressed are the author's own and do not reflect the views of AI Forum.*