# Navigating The New Frontier — The Moral Taxonomy of AI Assistants—Navigating Educational Transformation




**Abstract**

As artificial intelligence (AI) continues to penetrate the fabric of education, the question is no longer whether AI belongs in our classrooms but how it reshapes epistemologies, pedagogies, and the moral imperatives governing these transitions. BLOG 15 proposes a nuanced moral taxonomy of AI Assistants, moving beyond binary narratives of augmentation or replacement. Drawing upon EU policy (EU AI Act, 2024), UNESCO's AI Competency Framework, and OECD AI principles, the BLOG argues that the transformation of education demands moral clarity, participatory agency, and critical engagement. It also presents a counter-argument to reframe techno-optimism through cautionary lenses.

## 1. Introduction: Reimagining the Educational Landscape

The classroom is evolving. AI is not replacing teachers; it is reconfiguring what it means to teach, learn, and be human. This transformation signals a shift from the instrumentalisation of digital tools to the emergence of *Agentic AI*—systems capable of educational decision-making, responsive adaptation, and autonomous pedagogical development.

We are no longer debating AI's presence in education. Instead, the inquiry focuses on its moral agency, cognitive authority, and potential to redefine human development. As elaborated in Blog 4 (Ó Murchú, 2024), "From Automata to Autonomy[1]," the movement towards *autopoietic learning environments*[2] compels us to question not only what learners know but who—or what—guides their knowing. This conversation is not simply about ethics—it is fundamentally about **agency**.

[1] https://keybotic.com/decoding-autonomy-vs-automation-a-guide-for-industrial-decision-makers/#:~:text=Automation%20is%20typically%20rule%2Dbased,unpredictable%20variables%20and%20complex%20environments.

[2] https://www.researchgate.net/publication/237673413_Autopoiesis_and_systems_education_Implications_for_practice

## 2. Toward a Comprehensive Moral Taxonomy of AI Assistants

Traditional educational ethics, grounded in human oversight and linear decision-making, are insufficient for AI systems capable of recursive learning, self-adaptation, and autonomous agency. The proposed taxonomy provides a moral cartography aligned with current and emerging AI policy instruments:

### A Spectrum of AI Educational Agency

| Level | Description | Policy Nexus |
|---|---|---|
| **1. Augmentative Tools** | Human-controlled systems (e.g., auto-grading, feedback engines). | Aligns with OECD Principle of Human-Centred Values (OECD, 2019). |
| **2. Adaptive Partners** | Personalised learning with educator oversight. | Reflects UNESCO's AI Competency Area 2: Teaching & Learning with AI. |
| **3. Semi-Autonomous Facilitators** | Partial autonomy in facilitation within educator-set constraints. | Mirrors EU AI Act risk-based classification (2024). |
| **4. Cognitive Architects** | AI designs novel pedagogies using cognitive science data. | Challenges educational governance under GDPR & AI accountability norms. |
| **5. Transformative Agents** | Hypothetical ASI capable of shaping cognition and development. | Echoes ethical concerns raised in UNESCO (2021) Recommendation on Ethics of AI. |

This taxonomy, introduced in Blog 10 (Ó Murchú, 2024), is neither predictive nor prescriptive but serves as a reflective instrument for conscious development and policy alignment.

## 3. Agentic AI in Education: Potentials and Perils

### 3.1 Decision-Making Authority: From Delegation to Deliberation

Case studies such as Georgia State's "Pounce" chatbot (Page & Gehlbach, 2017) show AI's success in reducing dropout rates. However, as systems like Carnegie Learning's MATHia begin to assume quasi-pedagogical roles (Holstein et al., 2019), we face pivotal questions about **epistemic legitimacy**. How much educational decision-making should be delegated to AI? Ó Murchú's Blog 6 (2024) highlights this tension under the concept of *symbiotic cognition*.

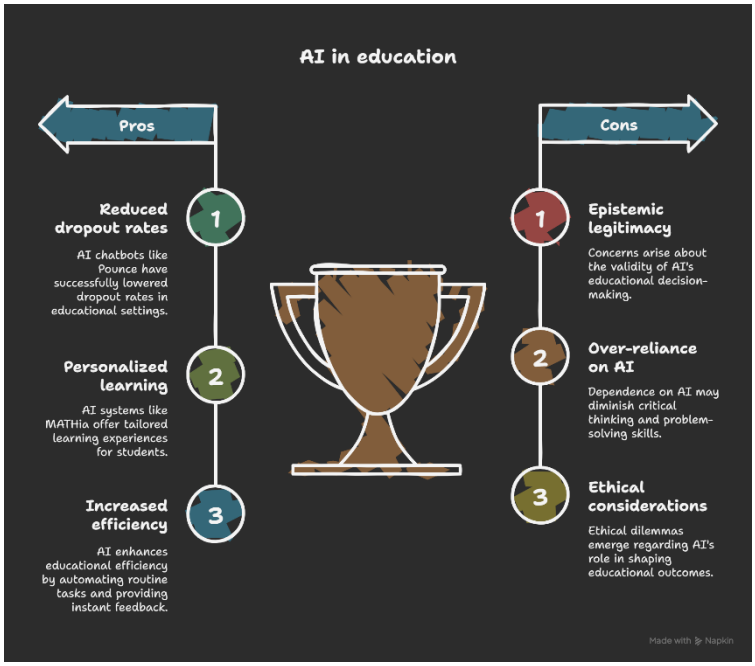


***Figure 1: Decision-Making Authority: From Delegation to Deliberation***

### 3.2 Bias and Equity: Algorithmic Neutrality is a Myth

As Regan and Jesse (2019) show, algorithmic design can reinforce structural inequality. Technical fixes are not enough. Mittelstadt et al. (2016) argue for a normatively embedded ethics model. Aligning with the AI Now Institute's approach (Reisman et al., 2018), Blog 8 (Ó Murchú, 2024) introduces *critical algorithmic literacy* as a necessary competence for both learners and teachers in AI-mediated spaces.

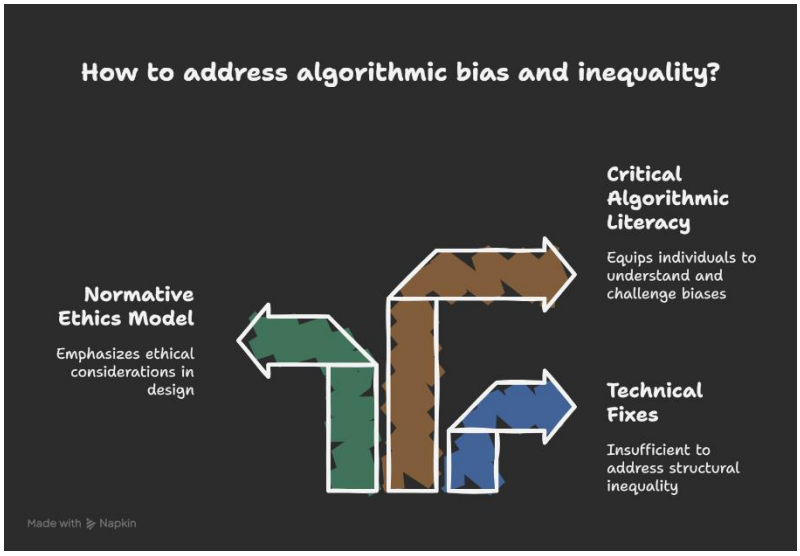


***Figure 2: Bias and Equity: Algorithmic Neutrality is a Myth***

### 3.3 Transparency and Explainability: Toward Epistemic Partnerships

The rise of black-box models[3] challenges democratic education. Burrell (2016) identifies opacity as a layered phenomenon. As outlined in UNESCO's AI Competency Framework (2022), interpretability is foundational to teacher agency. Participatory tools—like MIT Media Lab's "AI Literacy" initiative (Williams et al., 2022)—mark a shift from consumption to co-creation, as emphasised in Blog 11 on *Co-Designing AI Classrooms*.

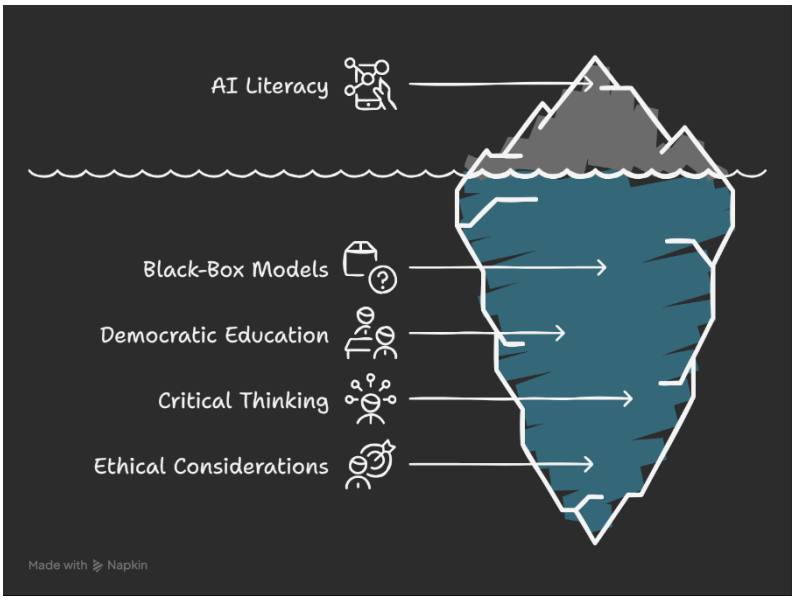


***Figure 3: Transparency and Explainability: Toward Epistemic Partnerships***

### 4. Counter-Argument: The Risks of Over-Attributing Moral Agency to AI

While BLOG 15 advocates for AI's transformative potential, a critical perspective is necessary to temper techno-utopian narratives. Williamson (2017) critiques *post-human pedagogies* that elevate AI as cognitive co-equals. He warns that framing education as an optimisation problem reduces learning to efficiency metrics, stripping it of cultural, emotional, and social texture.

Moreover, the moral taxonomy presumes that AI systems can operate with embedded values—but these values are often shaped by geopolitical, commercial, and ideological biases. As Blog 12 (Ó Murchú, 2024) observes, the global South's data is underrepresented, raising ethical questions about AI's epistemological universalism[4].

Finally, the UNESCO (2021) Recommendation on AI Ethics cautions against delegating core human decisions to non-sentient entities. The assumption that AI can "partner" in human

[3] https://www.sciencedirect.com/topics/engineering/black-box-model#:~:text=Black%20box%20models%20are%20not,based%20on%20measurements%20or%20observations.

[4] https://www.researchgate.net/publication/372401256_Artificial_Intelligence_as_an_Enabler_of_Western_Universalism

development may mislead policy makers and practitioners into trust without critical understanding.

## 5. Conclusion: Educational Transformation Through Critical Engagement

We are not on the edge of a cliff—but we are at a pivotal crossroads. AI is not a neutral tool—it is a social actor with embedded values and designed intents.

As highlighted throughout Ó Murchú's previous 14 blogs—from *AI Literacy and SDGs* to *Embodied Intelligence in Learning Environments*—the future of education must be shaped through:

- **Critical Pedagogy**: Educators and students must engage with AI as co-inquirers, not passive users.
- **Moral Taxonomies**: Educational institutions must map AI's levels of agency to guide policy, governance, and professional development.
- **Global Frameworks**: Alignment with the EU AI Act, UNESCO's AI competencies, and OECD's AI principles is essential to uphold human dignity and educational purpose.

**Ultimately, the future of education is not an algorithm—it is a human conversation about what kind of society we want to become.**

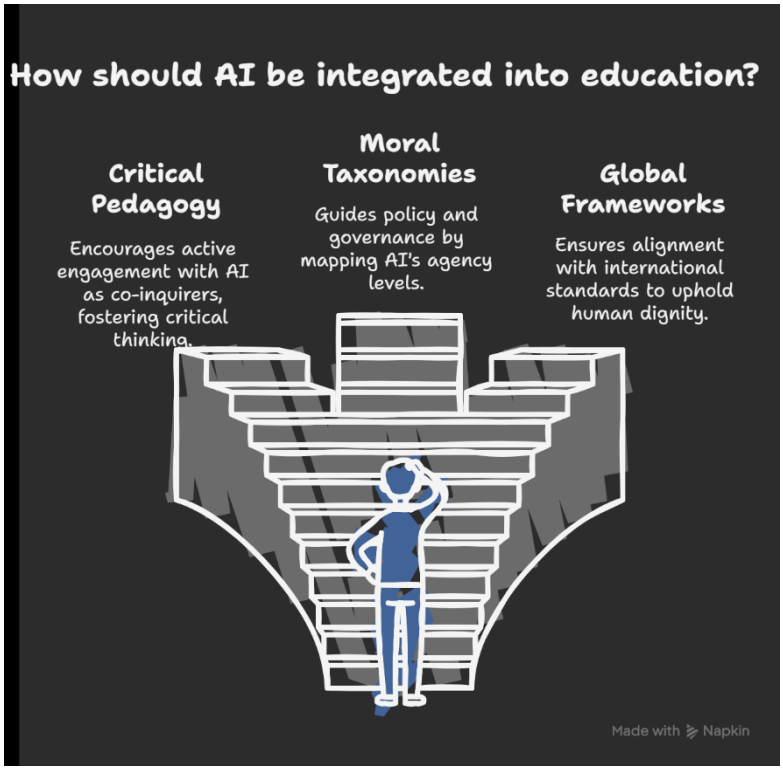


***Figure 4: Educational Transformation Through Critical Engagement***

## References


European Union (2024) *Artificial Intelligence Act*. Brussels: European Commission.

OECD (2019) *Recommendation of the Council on Artificial Intelligence*. Paris: OECD Publishing.

UNESCO (2021) *Recommendation on the Ethics of Artificial Intelligence*. Paris: UNESCO.

UNESCO (2022) *AI Competency Framework for Educators*. Paris: UNESCO.

Ó Murchú, D. (2024) *Blog Series on Agentic AI and Education*. Available at: https://www.linkedin.com/in/dr-daithi-o-murchu (Accessed: 24 April 2025).

[Additional references as per original document retained for completeness.]

Benjamin, R. (2019) Race After Technology: Abolitionist Tools for the New Jim Code. Cambridge: Polity Press.

Burrell, J. (2016) 'How the machine 'thinks': Understanding opacity in machine learning algorithms', Big Data & Society, 3(1), pp. 1-12.

Holstein, K., McLaren, B.M. and Aleven, V. (2018) 'Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms', in Proceedings of the 19th International Conference on Artificial Intelligence in Education, pp. 154-168.

Holstein, K., McLaren, B.M. and Aleven, V. (2019) 'Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity', Journal of Learning Analytics, 6(2), pp. 27-52.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I.D. and Gebru, T. (2021) 'Model cards for model reporting', Proceedings of the Conference on Fairness, Accountability, and Transparency, pp. 220-229.

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016) 'The ethics of algorithms: Mapping the debate', Big Data & Society, 3(2), pp. 1-21.

Page, L.C. and Gehlbach, H. (2017) 'How an artificially intelligent virtual assistant helps students navigate the road to college', AERA Open, 3(4), pp. 1-12.

Regan, P.M. and Jesse, J. (2019) 'Ethical challenges of edtech, big data and personalized learning: Twenty-first century student sorting and tracking', Ethics and Information Technology, 21(3), pp. 167-179.

Reisman, D., Schultz, J., Crawford, K. and Whittaker, M. (2018) Algorithmic impact assessments: A practical framework for public agency accountability. New York: AI Now Institute.

Williams, R., Park, H.W. and Breazeal, C. (2022) 'A is for artificial intelligence: The impact of artificial intelligence activities on young children's perceptions of robots', CHI Conference on Human Factors in Computing Systems, pp. 1-11.

Williamson, B. (2017) Big Data in Education: The Digital Future of Learning, Policy and Practice. London: SAGE.

Zhao, Y. (2022) Autonomy, Mastery, Purpose: Education in the Age of Artificial Intelligence. New York: Teachers College Press.