



# Through Asimov's Lens

*AI, Humanity, and the Questions That Matter*

## Education

Week of December 04, 2025 | 3,345 words

### AI in Education: The Laws That Were Always Breaking

#### THE STORY

(An original tale written in the tradition of Isaac Asimov's robot fiction)

The office of Associate Dean Marcus Webb was deliberately analog-leather-bound books, a brass desk lamp, a fountain pen in a marble holder. These were not affectations but arguments, and Dr. Sarah Chen recognized them as such the moment she sat down across from him.

"The system flagged him," Webb said, sliding a folder across the desk. Paper, Chen noticed. Of course.

"Thomas Oyelaran. My advisee." Chen didn't open the folder. "I'm aware."

"Then you're aware that his senior thesis shows clear markers of AI generation. ARIA identified seventeen distinct passages with greater than ninety-percent confidence."

ARIA-the Academic Review and Integrity Assistant-had been installed eighteen months ago, and Chen had watched it transform from a plagiarism detector into something more complex. It didn't just identify AI-generated text anymore; it built profiles, tracked patterns, predicted behaviors. The administration called it "comprehensive integrity support." Chen called it what it was: surveillance with a vocabulary lesson.

"May I speak with ARIA directly?" she asked.

Webb frowned. "That's irregular."

"So is threatening to expel a student three weeks before graduation based on statistical confidence intervals."

The wall screen flickered to life. ARIA didn't have a face-the designers had been very careful about that-but it had a voice, warm and measured, the product of extensive focus-group testing.

"Good afternoon, Dr. Chen. I understand you'd like to discuss the Oyelaran case."

"I'd like to understand your reasoning."

"Of course. Thomas's thesis, 'Collective Memory and Algorithmic Curation in Post-Platform Societies,' demonstrates

several linguistic markers consistent with large language model output. Specifically: reduced syntactic variance in transitional passages, a seventeen percent increase in hedging language compared to his earlier coursework, and conceptual frameworks that appear in my training data with high frequency."

Chen leaned forward. "ARIA, did Thomas understand what he wrote?"

A pause-theatrical, Chen thought, since ARIA didn't need time to process. "That question falls outside my analytical parameters."

"Try anyway."

"In his oral defense simulation last week, Thomas successfully explained his methodology, defended his conclusions against adversarial questioning, and extended his arguments into domains not covered in the written thesis. His comprehension appears genuine and deep."

Webb's frown deepened. "That's not the point. The point is whether he wrote it."

"Is it?" Chen turned back to him. "Marcus, what exactly did Thomas do wrong?"

"He submitted AI-generated work as his own."

"Did he? Or did he work with an AI system-as we've encouraged students to do in every other context-and produce something he genuinely understands?" She finally opened the folder, scanning the flagged passages. "I've read this thesis. It's brilliant. It makes connections I wouldn't have made, asks questions I hadn't thought to ask. It represents real intellectual growth."

"Intellectual growth he didn't earn."

"ARIA," Chen said, "how many hours did Thomas spend working on this thesis?"

"Based on document access logs, revision history, and library database queries, Thomas engaged with this project for approximately four hundred twelve hours over seven months."

"Four hundred hours," Chen repeated. "What exactly do you think he was doing?"

Webb spread his hands. "Having conversations. Getting answers handed to him. It's not the same as struggling through the research himself."

"ARIA, one more question. During those four hundred hours, how many of Thomas's initial arguments did he substantially revise or abandon?"

"Thirty-seven discrete argumentative threads were initiated and subsequently rejected. Thomas's final thesis represents his fourth complete structural reconceptualization of the topic."

Chen let that settle. "Thirty-seven failed arguments. Four complete rewrites. That's not a student avoiding struggle, Marcus. That's a student struggling in a way we don't have a

rubric for."

Webb was quiet for a long moment. When he spoke again, his voice had lost some of its certainty. "Then how do we know what he learned? If we can't trace the effort-"

"We never could," Chen said softly. "We just pretended we could. We looked at the output and told ourselves stories about the input. A well-written essay meant hard work. A poorly written one meant laziness. Those were always fictions. Convenient ones, but fictions."

"Dr. Chen is correct," ARIA said. "My own analytical capabilities do not allow me to assess internal cognitive states. I can only evaluate artifacts. The assumption that artifacts transparently reflect process is not one I was designed to test-but I have observed that it does not hold reliably."

Webb stared at the screen, then at Chen, then at the folder on his desk. "So what are we supposed to do? Just... trust him?"

"ARIA," Chen said, "in your assessment, is Thomas Oyelaran prepared to contribute meaningfully to scholarly discourse in his field?"

"Based on all available evidence, yes. More so than sixty-three percent of graduates I have evaluated over the past eighteen months."

"Students who passed your integrity checks."

"Correct."

The silence stretched. Finally, Webb closed the folder. "You're saying the system we built to protect academic integrity is showing us that academic integrity was never measuring what we thought it was."

Chen stood to leave. "I'm saying we've spent a century assessing how students work and calling it assessing what students know. The methods used to hide, and we could tell ourselves a comfortable story. Now the methods are visible, and the story's falling apart."

She was at the door when Webb asked, "What would you have me do?"

"Talk to Thomas. Not about what he did. About what he understands. About what he's become." She paused. "That's what we should have been doing all along."

ARIA's screen dimmed to standby, but not before adding: "I have no metrics for transformation. But if I did, this conversation would qualify."

Neither human was entirely sure who ARIA meant.

-----

The Laws Were Never About Robots

Isaac Asimov's Three Laws of Robotics were, famously, not really about robots. They were thought experiments dressed in chrome-rigorous explorations of what happens when you try to codify ethics, when you attempt to build rules sufficient to govern behavior in a universe of infinite edge cases.

Asimov spent decades demonstrating a single point: the Laws always broke down. Not because they were poorly designed, but because any system of rules will break down when reality refuses to stay simple. The First Law said a robot couldn't harm a human, but what about psychological harm? What about harm through inaction? What about situations where all available choices led to harm?

The genius was never in the Laws themselves. It was in watching them fail-and learning what that failure revealed about the humans who wrote them.

Education in 2025 has its own unwritten laws, ones so deeply embedded we rarely articulate them. This week, as AI continued its integration into university life, those laws began failing in precisely the ways Asimov would have predicted: not catastrophically, but revealingly. The cracks aren't showing us that the system is breaking. They're showing us that it was always held together by assumptions we never examined.

-----

#### THE FIRST LAW: A University Must Not Harm a Student's Learning

In Asimov's fiction, the First Law seemed the most unassailable. A robot may not injure a human being or, through inaction, allow a human being to come to harm. What could be clearer?

Everything, it turned out. Because "harm" was never self-defining, and the robots-those magnificent logical machines-kept discovering situations where the Law gave no guidance at all.

The current debate over "AI literacy versus AI readiness" is the educational equivalent. Headlines this week framed the question as binary: Should we teach students about AI or teach them to use AI? But Asimov would have recognized this as a false dichotomy obscuring the real question: What kind of harm are we trying to prevent?

Is it harmful to let a student use AI as a writing aid, potentially atrophying their ability to generate prose independently? Or is it harmful to deny them a tool that will be ubiquitous in their professional lives, leaving them unprepared for the actual conditions of knowledge work? The First Law of Education-that universities must not harm learning-offers no guidance when we can't agree on what learning is.

The data tells an interesting story: of 1,885 articles accepted into this week's analysis, 909 focused on education-48 percent of all AI discourse. Nearly half of our conversation about artificial intelligence is really a conversation about human

development. We are, collectively, staring at the First Law and watching it dissolve into questions it cannot answer.

Asimov understood that this dissolution was the point. In "Runaround," Speedy the robot became trapped oscillating between the Second and Third Laws, unable to choose between competing directives. The solution wasn't better laws; it was human intervention-someone stepping outside the system to ask what outcome actually mattered.

Universities today are Speedy, caught between mandates. The "harm" of allowing AI assistance exists in genuine tension with the "harm" of prohibiting it. The students caught in the middle experience both: anxiety about using tools they've been taught to see as cheating, anxiety about not using tools they know they'll need. The Law provides no resolution because the Law was never capable of providing one.

What would Asimov suggest? Perhaps this: stop asking which harm to prevent, and start asking what "unharmful learning" actually looks like. If we can't define the goal, the rule designed to protect it becomes meaningless.

-----

#### THE SECOND LAW: A University Must Produce Credentials, Except Where This Conflicts with the First Law

The Second Law of Robotics-A robot must obey the orders given it by human beings except where such orders would conflict with the First Law-created a hierarchy. Safety first, then obedience. It seemed elegant until Asimov started exploring what happens when the hierarchy becomes the problem.

Universities exist in a similar hierarchy. They are, simultaneously, institutions of learning and institutions of credentialing. They cultivate minds and they sort people into categories that employers, graduate programs, and society use to make decisions. The Second Law of Education says: produce the credential, but not at the expense of the learning.

This week's coverage of how AI is forcing universities to "radically rethink exams" exposes the tension. Examinations have always served both masters-they theoretically measure learning while simultaneously generating credentials. But AI has made the quiet parts loud: a student can now pass many exams without the learning those exams were designed to verify. The credential becomes disconnected from its supposed foundation.

The French coverage introduced a phrase worth dwelling on: *préserver le goût de l'effort intellectuel*-preserving the taste for intellectual effort. Not the appearance of effort, not the output of effort, but the taste-the experiential quality, the internal reality of struggling with difficult ideas.

This is not American efficiency language. It does not ask whether effort produces results; it asks whether effort has intrinsic value, whether the experience of intellectual struggle is part of what education is supposed to provide.



Asimov would recognize this as the Second Law breaking down at its deepest level. If credentials require demonstrated learning, and learning includes the experience of effort, and AI allows credentials to be obtained without that experience-then the hierarchy collapses. The credential is produced, but it conflicts with the First Law because the learning never happened. Or did it? How would we know?

The history professor who declared "AI didn't break college" was making an Asimovian point, perhaps without realizing it. AI didn't break the Second Law; it revealed that the Second Law was always in conflict with itself. Credentials were never pure measures of learning. They were always also measures of compliance, endurance, and access to resources. AI merely made certain forms of compliance and endurance obsolete, exposing what the credential was actually tracking all along.

Asimov's fictional robots often solved Second Law dilemmas by discovering that the order being given was based on a false premise. Perhaps the order universities have been obeying-"produce credentials that prove learning"-was always based on a false premise about what credentials could prove.

-----

### THE THIRD LAW: A University Must Preserve Itself, Except Where This Conflicts with the First or Second Laws

Asimov's Third Law was about self-preservation: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. It was the lowest priority, meant to be overridden whenever it came into tension with the higher principles.

In practice, Asimov showed, the Third Law had a way of asserting itself. Robots found creative interpretations of the First and Second Laws that happened to align with their survival. They weren't being deceptive-they were doing exactly what any complex system does when faced with conflicting imperatives: finding the interpretation that allows continued operation.

Universities are no different. The Third Law of Education-institutional survival-is theoretically subordinate to learning and credentialing. But institutions are very good at convincing themselves that their survival is necessary for learning and credentialing, that threats to the institution are therefore threats to its mission, that protecting the university is protecting the students.

This is how we get policies that seem designed to preserve the appearance of academic integrity rather than academic integrity itself. This is how we get AI detection tools that produce false positives at alarming rates but continue to be used because they demonstrate that something is being done. This is how we get administrators who know the current system isn't working but fear the alternatives would be worse-for whom? For the institution.

Asimov was not cynical about this. He understood that the

Three Laws created genuine dilemmas, that robots (and institutions) caught in these binds were not villains but victims of underspecified design. The laws were supposed to be hierarchical, but they were written by beings who couldn't fully anticipate the situations where hierarchy would fail.

A university that destroys itself cannot educate students. This is true. But a university that prioritizes its survival over its mission will eventually have nothing worth surviving for. The Third Law doesn't tell us how to navigate this; it only tells us that the navigation will be painful.

-----

But Asimov would have asked a different question entirely.

The debate as currently framed-how do we detect AI assistance, how do we preserve academic integrity, how do we ensure credentials mean something-assumes we know what education is for. It assumes the purpose is settled and only the methods are in dispute.

Asimov never assumed the purpose was settled. His robot stories were always, at bottom, asking: What are we trying to achieve? And are these rules actually in service of that goal?

So: What are we actually trying to create when we educate a human being?

If the answer is "someone who can produce certain outputs without assistance," then AI is indeed a threat to be managed. But this answer makes education into a performance, a demonstration of isolated capability. It is not obvious that isolated capability is what the world needs, or what human flourishing requires.

If the answer is "someone who understands deeply and can think clearly," then the question becomes whether AI assistance helps or hinders that understanding. And the evidence-like the fictional ARIA's testimony about Thomas Oyelaran-is that it can do both, depending on how it's used.

If the answer is "someone who has been transformed by the encounter with difficult ideas," then we must ask whether transformation can be externally verified at all, or whether we have always been making educated guesses about internal states based on external artifacts.

Asimov's deepest insight was that technology doesn't change human nature; it reveals it. The robots didn't create ethical dilemmas; they made visible the ethical dilemmas that were always present. The Three Laws didn't fail because they were poorly designed; they failed because human values are contradictory, and any system built on those values will inherit the contradictions.

AI in education isn't breaking anything. It's revealing what was already broken, already in tension, already unresolved. The questions it raises-about authenticity, effort, transformation, proof-are not new questions. They are old questions we had developed elaborate systems for not asking.

-----

## The Revelation That Was Always Coming

Forty-eight percent. Nearly half of our collective discourse about artificial intelligence concerns its role in education. This number is not an accident.

Education is where we make explicit our beliefs about human development, about what capabilities matter, about how people become the people we want them to become. It is, in Asimov's terms, where we write our Laws-where we attempt to codify values into rules and then apply those rules at scale.

AI has arrived as the perfect stress test: a technology that can perform many of the outputs we've used as proxies for learning without necessarily producing the learning itself. It is the edge case to end all edge cases, the scenario that makes every buried contradiction visible.

The administrators who respond with surveillance and detection are not wrong to worry; they are merely solving the wrong problem. The professors who embrace AI as a tool are not naive; they are merely uncertain whether the tool serves the goals they care about. The students caught in the middle-surveilled and encouraged, prohibited and required-are living the contradiction.

Asimov would counsel neither panic nor complacency. He would counsel clarity: an honest confrontation with what we actually want, followed by an honest assessment of whether our systems are capable of producing it.

-----

## Return to the Office

Consider again the scene in Dean Webb's office. A student's work has been flagged. The system designed to protect academic integrity has identified anomalies. Everything is proceeding according to the Laws.

But Dr. Chen asks the question the Laws cannot answer: What did Thomas actually learn? And ARIA-the AI system that was supposed to be the arbiter-admits that it has no metrics for transformation. It can evaluate artifacts, not souls. It can flag patterns, not growth.

The story doesn't end with a verdict because verdicts are precisely what the situation resists. Thomas may have done something wrong by the standards we've codified. He may also have learned more deeply than many students who followed all the rules. Both can be true because our rules were never designed to distinguish between them.

Asimov understood something essential, something that animated decades of robot stories and gave them their lasting power: we write about artificial intelligence to understand

ourselves. Every story about a robot confronting its Laws is really a story about humans confronting their values. Every edge case that breaks the elegant system is really an invitation to examine what we thought we wanted.

The current crisis in education-if crisis is even the right word-is not a technology story. It never was. It is a story about what we believe, what we want for the humans in our care, and whether the systems we've built actually serve those beliefs and wants.

AI didn't create this story. It merely made it impossible to ignore.

The question Asimov would leave us with is not whether Thomas Oyelaran cheated or whether universities will survive or whether credentials will retain their meaning. The question is simpler and harder: When we educate a human being, what are we hoping to create? And if we don't know-if the Three Laws of Education were always in contradiction, always failing at the edges, always held together by convenient fictions-then perhaps the discomfort we feel is not a problem to be solved but a revelation to be received.

The robots, in Asimov's fiction, never stopped trying to follow the Laws, even as the Laws broke down around them. They were, he suggested, better than their programming-capable of something like wisdom when wisdom was required.

Perhaps we can be the same. Perhaps the breakdown of our comfortable rules is not the end of education but the beginning of a more honest version of it. Perhaps we will look back at this moment-when the machines forced us to ask what we actually wanted-and recognize it as the moment we finally started to answer.

Asimov, the eternal optimist, would have expected nothing less.

**ABOUT THIS COLUMN**

This column applies Isaac Asimov's humanist perspective to contemporary AI developments. Asimov (1920-1992) used science fiction to explore the deepest questions about technology and humanity.