

The Detector as Doctrine: Higher Ed Chooses Surveillance Over Purpose

Weekly Analysis — <https://ainews.social>

There is a particular sound a university makes when it is frightened, and this year it has been unmistakable: the click of a "check for AI" button, the whir of a proctoring camera, the dry rustle of a freshly drafted code of conduct. Confronted with the arrival of generative models that can produce a passable undergraduate essay in the time it takes to refill a coffee, higher education has reached, almost as a reflex, for the instruments of detection and enforcement. And it has done so even as the evidence that those instruments do not work — that they misfire, that they punish the innocent, that they corrode the very trust they claim to protect — has piled up in plain view. National Public Radio put the paradox in its headline with admirable bluntness: [3]. That sentence is not a description of a technical problem. It is a description of a doctrine.

[3] AI detection tools are unreliable.
Teachers are using them anyway :
NPR

This essay is an attempt to map what that doctrine has done to the institution over the course of a single week's discourse — who is saying what, where the tensions run, and what the whole arrangement reveals about how universities are choosing to understand themselves in the presence of a technology they did not ask for and do not control. The short version is that the sector has staked its legitimacy on the wrong competence. It has redefined academic integrity as detection capacity, governance as press release, and learning as a problem to be policed rather than a purpose to be reimaged. The longer version requires that we watch the moves closely, because each of them is being made in the language of care.

The Arms Race Nobody Admits Losing

Start with the detectors, because everything else in the discourse orbits them. The premise of an AI detector is seductive in its simplicity: if a machine can write an essay, another machine can catch it. The trouble is that the second machine cannot, at least not reliably, and this has been documented so thoroughly that continuing to deploy the tools requires a kind of willed institutional forgetting. When students in California found themselves accused of cheating on the strength

of a detector's verdict, they did not quietly accept the charge; they organized, gathered evidence of their drafting process, and pushed back against faculty and administrators who had treated a probabilistic score as proof, a story the Los Angeles Times followed as it unfolded across campuses [9]. The human cost here is not abstract. It is a nineteen-year-old assembling a forensic defense of her own honesty because a vendor's confidence interval said she was a liar.

The failure is not marginal, and it is not going away with a software update. Detectors produce false positives at rates that would be intolerable in any process with real consequences, and they do so with a documented bias against certain writers — non-native English speakers, students with distinctive or formulaic styles, anyone whose prose happens to resemble the statistical center of gravity the model was trained to flag. Reporting on how the sector has actually responded to this reality reveals an institution caught between knowing and acting, aware that the tools are broken and unwilling to relinquish them [2]. The more honest administrators have begun to say so out loud. Inside Higher Ed reported this year on a quiet reversal in which some institutions are abandoning detection software altogether, having concluded that the accusations it generates cost more in trust and litigation than the cheating it fails to catch [4]. That is genuine progress. It is also, notably, framed as a matter of tools rather than of purpose — detectors are "out," new approaches are "in," as if the sector were updating its inventory rather than confronting a mistake.

Meanwhile the arms-race logic continues to escalate on the faculty side, where suspicion has curdled into something closer to dread. A professor at Brown became briefly notorious for suspecting that most of his class had used AI to cheat — a suspicion that, whether or not it was correct, reveals a teacher who has come to regard his own students as a population of probable offenders [6]. This is the psychological texture of the detection doctrine: it does not merely fail to catch cheaters, it reconstitutes the entire pedagogical relationship as adversarial. Coverage of the "cheating wars" on campus has described exactly this atmosphere — extreme surveillance, false accusations, and a jarring confusion in which neither students nor faculty can any longer tell what counts as legitimate work [11].

And when the accusations fail, they do not simply evaporate. They migrate into the legal system, where a growing docket of disputes now tracks students suing institutions over integrity findings and the occasional institution trying to hold the line. The existence of a dedicated tracker cataloguing every case and its outcome tells you the phenomenon has reached a scale worth indexing [1]. This is what an arms race looks like when it fails: not a decisive defeat but a slow

[9] Falsely accused of using AI, California college students push back as ...

[2] AI Detection in Education: How Schools Are Responding

[4] AI Detectors Are Out, New Approaches Are In - Inside Higher Ed

[6] Brown Professor Suspects Most of His Class Used AI to Cheat

[11] Inside college AI cheating wars: extreme surveillance, false ... - MSN

[1] AI Cheating Lawsuits Tracker — Every Case, Who Won (2026)

bleed of legitimacy, litigated one accusation at a time. The detector was supposed to protect the credential. Instead it has turned the credential into something you might have to defend in court.

Surveillance Dressed as Guardianship

The detector is only the most visible edge of a broader migration, one in which the university reimagines itself as a monitoring apparatus. The same institutions that would recoil at the phrase "surveillance state" have, in the name of integrity and safety, been steadily assembling one on their own grounds. The Pulitzer Center framed the question with appropriate directness — is this security, or is it privacy invasion? — in a report on the AI monitoring systems that campuses have begun to install to watch students, track behavior, and flag anomalies [26]. The framing matters, because it names the trade being made: the university is purchasing a sense of control at the price of the trust that used to define the relationship between a teacher and a learner.

The Spanish-language coverage has been sharper still on the technical recklessness of these deployments. Reporting on programs used to monitor students has emphasized not only the ethical objection but the security one — that the surveillance software itself introduces vulnerabilities, hoovering up sensitive data into systems whose safety cannot be guaranteed [21]. The controversy this provoked in the United States was covered across the Hispanophone press as a cautionary tale, a polemic over the wisdom of pointing an unaccountable eye at a captive population of young people [20]. What the surveillance vendors sell as a guardian is, on inspection, a new attack surface — a fresh liability wrapped in the vocabulary of protection.

The most instructive spectacle came from France, where an AI-driven exam-proctoring system failed so comprehensively that fifty-eight thousand candidates were forced to retake their examination, a fiasco that exposed the fragility of the whole automated-invigilation dream [23]. It is worth dwelling on the arithmetic of that failure. Fifty-eight thousand people had their time, their preparation, and their nerves sacrificed to a system that was supposed to secure the very legitimacy it destroyed. This is the surveillance guardian revealed as an unreliable narrator of its own competence — and the institution, having outsourced its judgment to the machine, left holding the wreckage.

To understand why universities keep reaching for these tools despite such outcomes, it helps to borrow a diagnosis from outside the sector.

[26] Using AI on Campuses: Security Surveillance or Privacy Invasion?

[21] Programas de IA para monitorear a estudiantes tienen riesgos de ...

[20] Polémica en EE.UU. por uso de IA para vigilar a estudiantes | Estados ...

[23] Surveillance d'examen par IA : le fiasco qui oblige 58 000 candidats à ...

Shoshana Zuboff’s account of how [24] describes forms of power that make “conflict itself impossible” by “refusing to admit the right of combat” reads almost as a prophecy of the proctored exam: a regime that treats the monitored subject not as a party with standing but as a source of data to be extracted and adjudicated. When a student cannot contest the verdict of a detector because the verdict issues from a black box, the university has quietly imported precisely the pathology Zuboff describes — a structure of authority that has removed the possibility of legitimate challenge. And the enthusiasm for extending that structure has no obvious ceiling. One trade outlet argued this year that colleges and schools must block and ban agentic AI browsers immediately, a call to erect ever-higher walls around the campus network as though the institution’s core mission were perimeter defense [7]. Each new tool is justified by the failure of the last, and each justification takes the underlying doctrine — that the answer to a technology is a wall — further from any question about what the university is for.

The Charm of “Responsible and Humane Use”

If surveillance is the operational face of the doctrine, its rhetorical face is the governance document — the framework, the guide, the institutional statement that recites the phrase “responsible and humane use” as though the words themselves discharged the obligation. This is the genre that has proliferated most luxuriantly over the past year, and it deserves scrutiny precisely because it presents itself as the antidote to the surveillance panic rather than its accomplice.

Consider the sheer volume of framework production. Québec issued a substantial guide to the responsible integration of AI in higher-education institutions, a document heavy with the language of ethics, equity, and deliberation [18]. WCET produced an entire policy and practice ecosystem framework mapping the terrain of institutional AI decisions [16]. Across Latin America, the Inter-American Dialogue advanced an institutional framework for how universities might appropriate the technology [17], while the Organization of Ibero-American States published a full map of AI’s arrival in the region’s higher education [19]. These are not trivial efforts. Many contain genuinely thoughtful passages. But when you read them alongside one another, a pattern emerges that should trouble anyone who cares about the difference between governing a thing and being seen to govern it.

The pattern is this: the documents are rich in procedure and poor in power. They describe values without specifying who holds the bud-

[24] The Age of Surveillance Capitalism

[7] Colleges And Schools Must Block And Ban Agentic AI Browsers Now. Here’s ...

[18] PDF Intégration responsable de l’IA dans les établissements d’enseignement ...

[16] PDF 2025 AI Education Policy & Practice Ecosystem Framework

[17] PDF Hacia Un Marco Institucional Para Apropiar La Inteligencia Artificial ...

[19] PDF La llegada de la IA a la educación superior en Iberoamérica: Un mapa ...

get, who signs the vendor contract, who decides when a tool is abandoned, or how faculty and students acquire a binding rather than a consultative voice. The word "humane" appears; the word "shared governance," in any enforceable sense, tends not to. This is the tell of policy theater — the ceremony of deliberation performed to legitimate decisions that have, in the operative sense, already been made in a procurement office. The AI-ethics literature itself has an aspiration that cuts against this reduction; as the argument in [24] puts it, to "do AI" well should mean folding ethics into a "more radically interdisciplinary and pluralistic" formation of judgment — a *Bildung*, not a checklist. The frameworks gesture at that horizon and then retreat to the checklist, because the checklist is cheaper and does not require ceding any actual authority.

There are honest attempts to do better, and it is worth naming them so the critique does not collapse into cynicism. The University of Lausanne built a lightweight decision aid — two minutes to determine which AI tool you may use with which data — that at least tries to translate abstract principle into a usable classroom judgment [5]. And institutions have invested real effort in preparing their people, as at one campus where faculty and staff undertook structured preparation for AI's growing role rather than simply waiting for a policy to descend [8]. These are the exceptions that clarify the rule. What distinguishes them is that they treat the faculty member as an agent with judgment to exercise rather than a compliance risk to be managed. The dominant governance genre does the opposite: it addresses the institution's members as potential liabilities and the technology's vendors as partners, an inversion of the relationship a university is supposed to have with its own scholars.

The missing piece across almost all of it is disclosure. A governance framework that recites "responsible use" while declining to say what the institution is paying, to whom, and on what terms has not governed the relationship with the AI industry — it has laundered it. The absence of budget transparency is not an oversight; it is the point at which the rhetoric of humane deliberation meets the reality of institutional complicity, and it is precisely the point the documents are structured to avoid.

Compliance as the Ceiling of Ambition

Beneath the high frameworks sits their operational residue: the codes of conduct, the ethics statements, the lists of buenas prácticas that land on a faculty member's desk and purport to tell them what critical

[24] AI Ethics - The MIT Press
Essential Knowledge series

[5] AI_GO : deux minutes pour savoir
quel outil d'IA vous ...

[8] Faculty and Staff Prepare for AI's
Growing Role in Higher ...

AI use looks like. And here the doctrine reveals its intellectual poverty most clearly, because these documents consistently reduce the whole rich problem of thinking-with-machines to three thin obligations — validate the output, protect the privacy, check for errors. That is the floor of competence. The literature presents it as the ceiling.

The Iberoamerican discourse has been unusually candid about how this plays out, describing what happens when the university becomes, in one memorable framing, a circus of machines — the institution deploying AI to grade what AI produced, an ouroboros of automation in which no human judgment survives at either end [14]. That image is worth holding onto, because it exposes what the compliance checklist cannot see: if critical use means only verifying and error-checking, there is nothing in the doctrine that would prevent the entire loop of writing and assessment from being handed to machines with a human reduced to notary. Validation, privacy, and error-checking are hygiene. They are not education. A code that mistakes hygiene for the whole of critical practice has confused not being negligent with being thoughtful.

The transparency demanded of students, meanwhile, is riddled with a contradiction the literature has begun to name. There is a genuine paradox in requiring scholars to disclose their AI use in research: the more honestly a researcher declares the assistance they received, the more suspect their work becomes in the eyes of a system primed to treat AI contact as contamination [13]. Honesty is punished; concealment is rewarded. This is what happens when an institution builds its integrity regime on detection and disclosure rather than on a shared understanding of legitimate practice — it manufactures exactly the dishonesty it set out to police. The compliance ceiling, in other words, does not merely fail to lift anyone toward critical practice; it actively incentivizes the concealment it exists to prevent.

And the students are not fooled by any of it. When professors themselves began using ChatGPT to prepare materials and provide feedback, some students objected — not to the technology as such but to the double standard, the spectacle of an institution that surveils undergraduates for AI use while its own faculty quietly adopt the same tools [15]. This is the compliance ceiling generating its own delegitimation. A rule that binds the watched and exempts the watchers does not read as ethics; it reads as power, and students of even average perceptiveness can tell the difference. The floor-as-ceiling problem is thus not only intellectually thin — it is politically unstable, because it asks a generation to accept a standard the institution will not apply to itself.

[14] La universidad en el circo de las máquinas: cuando la IA ...

[13] La paradoja de la transparencia en el uso de la IA generativa en la ...

[15] Los profesores utilizan ChatGPT y algunos estudiantes no están ...

The Purpose Question They Won't Ask

Here is where the whole survey converges, because underneath the detectors, the surveillance systems, the frameworks, and the checklists lies a single question the institution is spending enormous energy to avoid. Every detector purchase, every proctoring contract, every "responsible use" statement is a way of treating AI as a problem of speed — students can now produce work too fast, too easily, so we must catch them, slow them down, wall them off. But speed was never the real challenge. The real challenge is purpose: if a machine can generate a competent essay, then the point of assigning the essay — indeed the point of much of what the university does — is suddenly, urgently, up for renegotiation. The detection arms race is, at bottom, a strategy for not having that conversation.

The Latin American commentary has been most willing to name the true stakes. There is a growing warning about the cognitive debt that universities are unknowingly accumulating — the erosion of the very thinking capacities the institution exists to cultivate, incurred quietly each time the hard cognitive work of a course is offloaded without anyone asking what the work was for [12]. Cognitive debt is a better frame than cheating, because it locates the danger where it actually lives — not in a student's dishonesty but in the institution's failure to say why the intellectual labor mattered in the first place. You cannot defend the value of struggle with an essay if you have never articulated what the struggle was supposed to produce. The detector cannot answer that question. It was never designed to.

[12] La deuda cognitiva que las universidades no saben ...

The most rigorous voices in the literature have reframed the entire crisis as one of institutional truthfulness. The scholarship on universities at the crossroads of generative AI argues that what is really being tested is not the student's honesty but the institution's — its capacity to be truthful about what it values, what it assesses, and why [25]. A university that responds to AI by buying detectors has answered the truthfulness question in the negative: it has revealed that it would rather police the symptom than examine the practice. And the reimagining that the moment demands runs deeper than assessment design. Scholars working on global citizenship education have begun to reconceptualize the entire aim of higher learning for an age in which machines share the cognitive terrain, treating AI not as a threat to smuggle out of the classroom but as an occasion to ask what a formed human being now needs to become [22]. That is the purpose conversation the detectors are built to postpone.

[25] Truth and truthfulness: Universities at the crossroads of generative AI.

[22] Reconceptualising global citizenship education for the age of ...

It is worth naming what this avoidance obscures, because the mechanism is well understood outside education. Kate Crawford's account

of AI’s political economy in [24] warns of an “enchanted determinism” in which we are “told to focus on the innovative nature of the method rather than on what is primary: the purpose of the thing itself” — an enchantment that “obscures power and closes off informed public discussion, critical scrutiny, or outright rejection.” The detection industry runs on precisely this enchantment. It invites the university to marvel at the sophistication of the tool, to argue about accuracy rates and false-positive thresholds, to treat the whole matter as a technical problem of method — and in doing so it forecloses the prior question of whether the institution should be running a surveillance operation against its own students at all.

There is, to be fair, a version of the discourse that claims to be looking forward. Consulting firms have begun to advise universities on how to lead AI research into the coming years, casting the institution as a driver of innovation rather than a victim of disruption [10]. But leadership in research is not the same as clarity of purpose in teaching, and it is telling that the ambitious, future-facing framing tends to appear in the register of competitiveness — how to win, how to lead, how to capture the opportunity — rather than in the register of formation, of what the university is for. The strategic vision and the pedagogical vacuum coexist comfortably, which is itself a sign that the sector has not yet connected its ambition to its actual reason for existing.

[24] The Atlas of AI

[10] How Universities Can Lead AI Research in 2026

What the Missing Conversation Would Sound Like

The absence that haunts this entire landscape is the one the discourse is least able to admit: the missing framing is partnership. Nearly every position surveyed here — the detector’s, the proctor’s, the code of conduct’s, even many of the earnest frameworks’ — treats the relationship between the institution and its members as adversarial or custodial. The student is a suspect; the faculty member is a liability; the technology is a threat to be contained. What almost no one is building is a collaborative settlement in which students, faculty, and the institution together decide what AI is for in a given discipline, what legitimate use looks like, and how the work of thinking will be protected and valued rather than merely surveilled.

This is not a utopian demand. It is the ordinary practice of shared governance applied to a new circumstance, and its near-total absence from the operational documents is the most damning finding of the survey. When the University of Lausanne’s decision tool or one campus’s faculty-preparation effort stands out, it is because they alone

treat the university's members as collaborators in judgment rather than objects of policy [5]. The rarity of that stance is the measure of how far the doctrine has spread. And the doctrine spreads because it is easy: a detector can be purchased and a code can be circulated in a week, while a genuine partnership over purpose requires the institution to expose its own uncertainty about what it is doing and why.

The governance-versus-pedagogy imbalance runs through the entire body of discourse like a fault line. The frameworks are dense with governance — committees, principles, compliance regimes, ecosystem maps — and thin on pedagogy, on the actual question of how a person learns to think in a world where a machine can produce the appearance of thought [16]. This is not because pedagogy is harder to write about, though it is. It is because governance can be performed at the level of the press release while pedagogy has to be lived in the classroom, and the institution under pressure will always prefer the performance it can complete to the practice it must sustain. A code of conduct can be adopted by a committee before lunch. A reimagined relationship to intellectual work is the labor of years, and it cannot be delegated to a vendor.

The Cost of the Wrong Competence

What, then, is happening to higher education as AI arrives? The comfortable answer, the one the sector tells itself, is that it is responding — vigilantly, responsibly, humanely — to a technological threat. The truer answer, the one the week's evidence forces, is that the institution is answering a crisis of purpose with a display of enforcement, and staking its legitimacy on a competence that will not save it. Detection does not work and, as reporting on the sector's own reversals has shown, some institutions already know it [4]. Surveillance introduces liabilities it cannot contain, as fifty-eight thousand French candidates learned when the proctoring system collapsed under them [23]. Frameworks recite values while declining to disclose the budgets and contracts that would make those values enforceable. And the checklists mistake the hygiene of validation and error-checking for the substance of critical thought, incentivizing the very concealment they were meant to prevent [13].

The tragedy is that each of these moves is made in the name of trust, and each of them corrodes it. A student accused by a detector trusts the institution less [9]. A professor who suspects his whole class trusts his students less [6]. A campus that surveils its members trades the relationship it exists to cultivate for a control it cannot actually

[5] AI_GO : deux minutes pour savoir quel outil d'IA vous ...

[16] PDF 2025 AI Education Policy & Practice Ecosystem Framework

[4] AI Detectors Are Out, New Approaches Are In - Inside Higher Ed

[23] Surveillance d'examen par IA : le fiasco qui oblige 58 000 candidats à ...

[13] La paradoja de la transparencia en el uso de la IA generativa en la ...

[9] Falsely accused of using AI, California college students push back as ...

[6] Brown Professor Suspects Most of His Class Used AI to Cheat

achieve [26]. The detection arms race claims to prevent the erosion of trust; it is the erosion of trust, conducted at institutional scale and billed as prudence.

The way out is not another tool, and it is not another framework. It is the conversation the whole apparatus was built to avoid — the one about purpose, about why the intellectual labor of a course matters when a machine can counterfeit its product, about what a university is for once the production of competent text is no longer scarce. The scholars warning of cognitive debt and the ones reframing the crisis as a test of institutional truthfulness have already begun that conversation [12]. What they are waiting for is an institution brave enough to admit that the detector was never a doctrine of integrity at all. It was a way of not deciding what the university believes. And a university that will not decide what it believes has already conceded the only competence that could have justified its existence — the courage to say, out loud and to its own members, what all of this was ever for.

[26] Using AI on Campuses: Security Surveillance or Privacy Invasion?

[12] La deuda cognitiva que las universidades no saben ...

References

1. AI Cheating Lawsuits Tracker — Every Case, Who Won (2026)
2. AI Detection in Education: How Schools Are Responding
3. AI detection tools are unreliable. Teachers are using them anyway : NPR
4. AI Detectors Are Out, New Approaches Are In - Inside Higher Ed
5. AI_GO : deux minutes pour savoir quel outil d'IA vous ...
6. Brown Professor Suspects Most of His Class Used AI to Cheat
7. Colleges And Schools Must Block And Ban Agentic AI Browsers Now. Here's ...
8. Faculty and Staff Prepare for AI's Growing Role in Higher ...
9. Falsely accused of using AI, California college students push back as ...
10. How Universities Can Lead AI Research in 2026
11. Inside college AI cheating wars: extreme surveillance, false ... - MSN
12. La deuda cognitiva que las universidades no saben ...

13. La paradoja de la transparencia en el uso de la IA generativa en la ...
14. La universidad en el circo de las máquinas: cuando la IA ...
15. Los profesores utilizan ChatGPT y algunos estudiantes no están ...
16. PDF 2025 AI Education Policy & Practice Ecosystem Framework
17. PDF Hacia Un Marco Institucional Para Apropiar La Inteligencia Artificial ...
18. PDF Intégration responsable de l'IA dans les établissements d enseignement ...
19. PDF La llegada de la IA a la educación superior en Iberoamérica: Un mapa ...
20. Polémica en EE.UU. por uso de IA para vigilar a estudiantes | Estados ...
21. Programas de IA para monitorear a estudiantes tienen riesgos de ...
22. Reconceptualising global citizenship education for the age of ...
23. Surveillance d'examen par IA : le fiasco qui oblige 58 000 candidats à ...
24. The Atlas of AI
25. Truth and truthfulness: Universities at the crossroads of generative AI.
26. Using AI on Campuses: Security Surveillance or Privacy Invasion?