Policing a Border That No Longer Exists: The Compliance Pretense of the AI University

Weekly Analysis — https://ainews.social

Sometime in the last two years the university convinced itself that the central problem posed by generative AI was a problem of detection and permission — a border-control problem. Who crossed the line? Which side of it does this paragraph belong to? Was the work *produced by the student* or by a machine wearing the student's name? The institutional response has been prodigious: usage taxonomies with tiers and traffic lights, authorship declarations, detector guidance, proctoring protocols, and governance frameworks arriving faster than anyone can validate them. Read enough of these documents in a single week — and this week produced a great many — and you notice something disorienting. They are all patrolling a frontier between human and machine authorship that the technology they are responding to has already erased. The surveys of how students actually work now describe a thoroughly blended practice, drafting and revising in continuous conversation with a model [15]. The border the policies defend is a line drawn on water.

[15] How universities are adapting to students' use of AI

This is not a small administrative embarrassment. It is a structural mistake with a cost, and the cost is being paid unevenly. The genre of documents proliferating this week — what the academic literature has begun, drily, to catalogue as "emerging policy approaches" to governing generative AI [11] — presupposes that authorship can be cleanly attributed, that a permissible-use line can be drawn and enforced, and that the university's job is to draw and enforce it. Meanwhile the vendors whose tools created the ambiguity have arrived offering both the disease and the cure: guidance on how educators should respond when students pass off generated content as their own [30], and detectors positioned as instruments of institutional protection. The university has accepted the framing. In doing so it has agreed to become a compliance apparatus — an institution that certifies authorship it cannot actually verify — and it has left the threats that could genuinely hollow it out entirely unaddressed.

[11] Governing generative AI in higher education: emerging policy approaches and considerations

[30] ¿Cómo pueden responder los educadores cuando los estudiantes presentan contenido generado por IA como si fuera propio?

What follows is a map of that week's discourse, read against its own grain. The point is not that universities should do nothing. It is that

they are doing a great deal of the wrong thing, with visible energy and evident sincerity, while the territory changes beneath them.

The Taxonomy Explosion, or the Wrong Question Asked Beautifully

The signature artifact of this moment is the usage taxonomy. You have almost certainly seen a version: a graded scale that runs from "no AI permitted" through "AI for brainstorming only" and "AI for editing" up to "AI use encouraged and disclosed," often color-coded, sometimes accompanied by a required declaration in which the student attests to which tier they operated within. Institutions have poured real craft into these instruments, and the emerging governance literature treats their proliferation as evidence of a maturing sector, cataloguing the range of tiered permission structures now in circulation [11]. The most sophisticated of them lean on a clause of apparently commonsense force — that the submitted work must be, in the end, *produced by the student* — as though "produced by" were a bright and stable predicate rather than the exact thing generative AI has thrown into doubt.

[11] Governing generative AI in higher education: emerging policy approaches and considerations

Here is the move to watch. The taxonomy answers the question *What may I use?* with impressive granularity. It does not touch — cannot touch — the question that actually matters, which is *What does it mean to author work you did not fully produce?* These are different questions, and the second is not reducible to the first. A student who uses a model to restructure an argument, generate three counter-examples, and smooth every transition has, under most tiered schemes, remained inside the permitted zone; the labor of judgment that we once took the essay to certify has nonetheless been substantially redistributed. Surveys of actual student practice describe precisely this kind of continuous, granular collaboration, in which the AI is woven through the cognitive work rather than bolted onto its edges [23]. No traffic-light column captures where, in that weave, the student's authorship ends and the model's begins, because there is no such point to capture. Authorship is not a threshold that a workflow crosses. It is a property of the whole relation.

[23] Rapport d'enquête - Usages de l'IA dans les pratiques étudiantes

The taxonomies pretend otherwise, and the pretense is the problem. When MIT's own reckoning with the technology reaches for a redesign of the university rather than a rulebook for the essay, it is implicitly conceding that the unit of analysis is wrong — that you cannot regulate your way to authorial integrity one assignment at a time [25]. The more ambitious redesign proposals go further, arguing that gen-

[25] Report - AI and Education

erative AI requires rethinking curriculum and classroom practice from the ground up, precisely because the old products of assessment no longer certify what we assumed they certified [24]. That is the honest response, and it is telling how much rarer it is than the taxonomy. Drawing tiers is cheap, fast, and defensible in a committee. Rebuilding the relationship between teaching and judgment is expensive, slow, and politically exposed. The institution reaches for the cheap artifact and calls it governance.

[24] Redesigning STEM Higher Education in the Era of Generative AI: From Curriculum Design to Classroom Practice

There is a deeper irony in the manifesto strain of this literature — the authorship declarations, the integrity charters, the earnest statements of institutional value. They are written as though the university's authority over authorship were a settled possession merely in need of reaffirmation. But that authority never came from the ability to certify who wrote what. It came from teaching students to author — to make and defend the judgments that writing externalizes. A charter that relocates the institution's role from *cultivating* judgment to *attesting* its provenance has quietly traded the thing that made the university valuable for the thing that makes a notary valuable. The compliance artifact is not a defense of the university's authority. It is its liquidation, dressed as its assertion.

The Asymmetry the Detectors Launder

If the taxonomy is the week's signature artifact, the detector is its signature weapon, and here the stakes stop being merely epistemic and become, bluntly, a matter of who gets hurt. The vendor pitch for AI-detection and proctoring software rests on a statistical framing that sounds neutral and is not: every classifier trades false positives against false negatives, and the vendor presents this as an unfortunate but symmetric tradeoff to be tuned. It is not symmetric. A false negative means a student who used AI without disclosure is not caught — an integrity loss, real but recoverable, absorbed by the institution. A false positive means a student who did nothing wrong is accused of academic dishonesty — a disciplinary and reputational harm, sometimes life-altering, absorbed by an individual. These are not two faces of one error rate. They are different kinds of injury falling on different parties, and the vendor's "neutral statistical tradeoff" language exists precisely to make you forget that.

The legal and procedural scholarship this week names the mechanism with unusual clarity: detection tools produce opaque evidence — a probability score with no auditable reasoning — and opaque evidence corrodes due process, because a student cannot cross-examine a

number [4]. The burden of disproving a machine's suspicion is placed on the accused, inverting the ordinary direction of proof, and the accused's ability to mount that defense is not evenly distributed. A student with a parent who knows how to write a firm email, or the resources to escalate, fares differently from a first-generation student who assumes the accusation must be correct because the institution made it. The reporting aimed at families spells this out — that proctoring and detection systems are failing students, and that the failure is invisible to exactly the families least equipped to contest it [5]. The tool advertised as an instrument of fairness operates as an amplifier of existing advantage.

[4] AI Detection Tools and Academic Punishment: How Opaque Evidence Threatens Due Process

[5] AI Proctoring Tools Are Failing Students: What Parents Must Know

This is not a novel dynamic; it is the oldest dynamic in the sociology of technical systems, and the critical AI literature has been describing it for years. Kate Crawford's account of what machine classification does — boiling an "infinitely complex and changing universe" down to "a terse set of formalisms based on proxies," naming some features while obscuring countless others [26] — is a precise description of what a detector does to a student's prose. The detector does not read the essay. It reads a proxy for the essay, a statistical signature of "AI-ness," and then a human being, trusting the machine's authority, converts that proxy into a charge. The proxy is wrong often enough to matter, and it is wrong in patterned ways: non-native English writers, students who write in a plainer register, students whose style happens to resemble the model's average — these are the reliable false positives. The systematic reviews of online proctoring reach compatible conclusions, documenting how these systems generate suspicion unevenly and offload the cost of their errors onto the surveilled [2].

[26] The Atlas of AI

[2] A Systematic-Narrative Review of Online Proctoring Systems and a Case Study

And the students most exposed to the detector's errors are, by the evidence, often the same students the institution should be most anxious to protect. The largest study yet of undergraduate AI use found the adoption of these tools shot through with disparities — in access, in disclosure, in who feels entitled to use them openly and who uses them furtively [27]. Layer a punitive detection regime on top of a stratified usage landscape and you get a machine that manufactures accusations along the grain of existing inequality, then launders the result as impartial enforcement. The document that establishes the detector as institutional protection never mentions whom it fails to protect. That silence is not an oversight. It is the design.

[27] The largest study of AI use by undergrads is in, revealing disparities in access and in cheating

Detection as Theater, and the Panic That Sells It

The market for detection runs on episodes of dramatized crisis, and this week supplied two. A Spanish professor denounced mass AI-enabled fraud on an examination at Brown, framing the episode as evidence that academic integrity itself is in mortal danger [28]. In Mexico, the largest university in Latin America opened an investigation into irregularities in its admission examination, a scandal touching the most consequential gatekeeping ritual the institution performs [20]. These stories are real, and the anxiety behind them is not manufactured. But notice how frictionlessly the panic converts into procurement. Every dramatized breach becomes an argument for more detection, more proctoring, more surveillance — an escalation that the vendors are standing by to supply, complete with white papers explaining why clear proctoring rules are the responsible institution's obligation [3].

[28] Un catedrático español denuncia fraude masivo con IA en un examen en la Universidad de Brown

[20] Por primera vez, la universidad más grande de América Latina hizo pública una investigación por irregularidades

[3] Academic Integrity in the Age of AI: Why Clear Proctoring Rules Matter

The trouble is that detection does not work well enough to justify the authority placed in it, and the technical literature says so without much hedging. Careful analyses of the detection of generative-AI use conclude that reliable, defensible identification is not currently achievable — that the tools produce probabilistic guesses unfit to ground a disciplinary determination [17]. This is the crucial fact that the crisis-and-procurement cycle is structured to bury. The university buys a tool that cannot do the thing it is bought to do, deploys it against the students least able to contest its errors, and experiences the purchase as diligence. The theater is convincing because everyone in it is playing an earnest part. The professor is genuinely alarmed. The administrator is genuinely responsible. The vendor is genuinely helpful. And the net effect is a surveillance apparatus erected on a foundation the apparatus's own makers cannot certify.

[17] La détection de l'usage des systèmes d'IA génératives

There is a quieter, more corrosive cost to running the institution this way, which the restorative-justice strand of the integrity literature has begun to name. When the university frames integrity primarily as a policing problem — catch, prove, punish — it forecloses the preventive and restorative approaches that actually change student behavior, the ones that treat integrity as something cultivated through the design of assignments and the relationship of trust rather than enforced through detection [21]. The policing frame is not merely ineffective; it is actively crowding out the frame that works. Every dollar and every committee-hour spent on detection is a dollar and an hour not spent on the redesign that would make detection beside the point. The vendor inversion — detectors as institutional protection — sells the institution a substitute for the pedagogical work, and the institution,

[21] Preventive and Restorative Academic Integrity in AI-Assisted Learning Environments

exhausted and afraid, buys the substitute.

Governance Multiplying Ahead of Evidence

Step back from any single artifact and survey the field, and the dominant pattern is a strange one: governance is multiplying far ahead of the evidence that could tell us whether any of it works. Task forces convene, frameworks are published, community-college systems produce readiness roadmaps, provosts stand up curriculum committees — a whole genre of institutional self-organization has bloomed in eighteen months. The Illinois campus AI curriculum task force is a representative instance of the form, an institution assembling its authorities to produce guidance and structure [7]. The community-college sector produced its own detailed vision of the "AI-enabled" institution, mapping governance, workforce alignment, and implementation with genuine seriousness [9]. And the flagship redesign proposals — MIT's argument that the university itself should be reconceived for the AI era — carry real intellectual ambition [18].

[7] AY2025-26 Campus AI Curriculum Task Force

[9] Creating the AI-Enabled Community College

[18] MIT propone rediseñar la universidad para la era de la IA

I do not want to be glib about this. Institutions faced with a genuine discontinuity have to do *something*, and convening to think is better than paralysis. But the striking feature of the governance genre is how much of it is normative confidence unmoored from empirical grounding. Frameworks specify what responsible AI use *should* look like, in elaborate acronymic detail, long before anyone has established what AI use *does* — to learning, to judgment, to the durable capacities a degree is supposed to certify. The governance literature itself, surveying the emerging policy landscape, is candid that these approaches are being generated under conditions of deep uncertainty, in advance of the evidence that would validate them [11]. We are watching an inversion of the ordinary relationship between knowledge and rule. Normally you learn how something works and then you regulate it. Here the regulation is racing ahead, generating the comforting appearance of control over a phenomenon no one yet understands.

[11] Governing generative AI in higher education: emerging policy approaches and considerations

This matters because unvalidated governance is not neutral. It hardens into policy, policy into infrastructure, infrastructure into the taken-for-granted shape of the institution — and it does all this before the corrective evidence can arrive. A framework adopted in 2026 on the basis of 2026's guesses will still be structuring assignments and adjudicating cases in 2029, when we may know those guesses were wrong. And the frameworks tend to encode the assumptions of the moment, chiefly the assumption that the human/machine authorship border is real and policeable — the very assumption the technology

has dissolved. The genre thus performs a subtle bait-and-switch on the institution: it offers the feeling of having responded to the crisis while quietly entrenching the crisis's founding misconception. The AI Ethics literature has warned about exactly this reflex — the temptation to treat ethics and governance as a compliance layer bolted onto an unexamined technical practice, rather than as a reconception of what it means to "do" the practice at all [26]. The university is bolting on the layer. The reconception is not happening.

[26] AI Ethics

The Threat the Documents Do Not Mention

Here is the most consequential silence in the entire week's output. Across the taxonomies and the detector guidance and the governance frameworks — thousands of words on permission, provenance, and enforcement — there is almost nothing about the structural threat most likely to hollow the university out: the extraction of its research talent by the AI companies themselves. The same firms selling detection to police undergraduate essays are, at the other end of the institution, buying out the faculty and graduate researchers who constitute the university's future capacity to know anything at all. The compliance apparatus is aimed downward, at nineteen-year-olds, while the actual capture is happening upward and out, and the documents do not so much as gesture at it.

This is where Crawford's insistence on seeing AI as "a manifestation of highly organized capital backed by vast systems of extraction" earns its place in the argument [26]. The extraction is not only of minerals and labor and data; it is of the human research capacity that public and academic institutions spent a century building. When a lab's principal investigators and its most capable postdocs are absorbed by industry — with the compute, the salaries, and the data no university can match — the institution loses not a few employees but its ability to generate independent knowledge about the very technology it is supposedly governing. The university then finds itself regulating a phenomenon it can no longer study, dependent for its understanding on the firms it is meant to hold at arm's length. MIT's own report gestures toward the scale of the transformation, but the institutional response has overwhelmingly been to manage student behavior rather than to confront the extraction of institutional capacity [25].

[26] The Atlas of AI

[25] Report - AI and Education

Why the silence? Partly because the threat is uncomfortable — naming it means acknowledging that the vendors offering to protect academic integrity are simultaneously dismantling academic capacity. Partly because the border-policing frame is genuinely absorbing:

a false-positive dispute is concrete, immediate, and adjudicable in a way that a slow talent hemorrhage is not. And partly, one suspects, because the policing frame is the one that leaves the institution's relationship with the vendors intact. You can buy a detector from a company you would rather not think too hard about. Confronting research-talent extraction would require the university to recognize the AI industry as an adversary in a resource conflict, not merely a supplier of classroom tools — and that recognition is precisely what the compliance posture is structured to avoid. The disparities documented among students in the largest usage study to date [27] are a distributional problem the university could actually address; the extraction problem is an existential one it is choosing not to see.

[27] The largest study of AI use by undergrads is in, revealing disparities in access and in cheating

What Authorship Was Supposed to Certify

To understand why the border-policing frame fails, it helps to recover what the essay, the exam, and the problem set were ever *for*. They were never valuable as objects. They were valuable as evidence — proxies for the internal capacities we care about: the ability to reason under uncertainty, to make and defend a judgment, to hold a problem in mind and work it. The reason AI destabilizes assessment is not that it produces good-looking artifacts; it is that it severs the reliable link between the artifact and the capacity the artifact was taken to certify. The frontier literature on what remains distinctively human is, in effect, an inventory of the capacities that no longer travel reliably with the product — the non-automatable cognitive skills a degree is supposed to build and warrant [19].

[19] Non-automatable cognitive skills in higher education in the age of AI

And this reframing exposes what the compliance apparatus is actually risking. The danger of frictionless AI assistance is not that students will "cheat"; it is that they will offload the cognitive effort that constitutes learning, and arrive at credentials certifying capacities they never built. The French-language research on cognitive effort names the peril directly, arguing that generative AI in higher education can become a danger precisely by removing the productive struggle through which understanding forms [16]. A parallel argument frames the risk as intellectual alienation — the student estranged from their own thinking, outsourcing the internal work to a system whose outputs they can no longer independently evaluate [22]. The psychological research on how these tools reshape human skills and thinking points the same way, documenting how habitual reliance can atrophy the very capacities the reliance is meant to augment [14].

[16] L'IA générative dans l'enseignement supérieur : un danger pour l'effort cognitif

[22] Prévenir l'aliénation intellectuelle à l'ère de l'IA : vers une pédagogie de l'autonomie

[14] How AI is reshaping human skills and thinking

Notice that none of this is a detection problem. You cannot catch

cognitive offloading with a classifier, because offloading is often perfectly compliant with the usage taxonomy — it happens inside the permitted tiers, in the encouraged-and-disclosed zone. The problem is pedagogical, and it has a pedagogical shape: it concerns what assignments ask students to *do*, how the doing is structured so that the struggle cannot be skipped, how judgment is taught and rehearsed and made visible. The philosophical literature on professional judgment in the AI era makes the deeper point that judgment is not a deliverable but a formed capacity, developed through practice under stakes [12]. A university that pours its energy into policing provenance while neglecting the design of that practice is defending the wrong perimeter with the wrong tools.

[12] GUERSENZVAIG - revistas.um.es

The complicating evidence deserves its place here, because the picture is not simply that AI destroys learning. There is a striking randomized controlled trial finding that AI tutoring can outperform in-class active learning on measured outcomes [6]. This should be taken seriously, and it cuts against any reflexive technophobia. But it also sharpens the real question, which is not *whether AI helps or harms* but *which uses build capacity and which uses substitute for it* — a distinction that the border-policing frame, obsessed with authorship provenance, is structurally incapable of drawing. A well-designed AI tutor that forces a student to work through a derivation builds judgment. A well-disclosed AI ghostwriter that produces the derivation destroys the occasion to build it. Both are permitted uses. Neither is a border violation. The distinction that matters runs orthogonal to the line the university is policing.

[6] AI tutoring outperforms in-class active learning: an RCT at a Nigerian university

The Partnership That Never Got Framed

What is conspicuously missing from the week's discourse — and the absence is structural, not accidental — is any serious framing of AI integration as a *collaborative* project between the institution and its students, as opposed to an adversarial or compliance relationship. The dominant postures are two: policing (detect and punish) and provisioning (roll out tools and taxonomies). Both position the student as an object of institutional action rather than a partner in a shared problem. Yet the shared problem is real and genuinely difficult — how to build durable intellectual capacity in an environment where the shortcut is always available — and it is precisely the kind of problem that yields to honest collaboration and resists enforcement. The most thoughtful assessment-redesign work coming out of Latin America frames the challenge as *saving learning* when a machine can do the tasks, which reframes the student not as a suspect but as a

co-stakeholder in preserving something both parties value [10].

The equity dimension makes the partnership framing not just preferable but urgent. AI tools carry real potential to widen access — the accessibility literature documents genuine gains for students with disabilities and the possibility of more equitable, universally designed learning [29], and practical guides for equity and inclusion in higher education are beginning to map how the same tools that threaten some students can serve others [13]. But these gains are only realizable in a partnership frame. Under a policing frame, the accommodation and the violation are indistinguishable to the detector — the assistive-technology user and the ”cheater” produce the same statistical signature — and the equity potential curdles into an equity threat. Whether AI widens or narrows access depends almost entirely on whether the institution treats its students as collaborators to be equipped or suspects to be surveilled. The literacy programs that treat AI as an occasion to *strengthen* critical judgment rather than police it — such as the West Bank media-literacy initiative building critical thinking for an AI world [1] — point toward the frame the compliance documents cannot reach.

The reason the partnership framing keeps failing to appear is that it demands something the compliance posture is designed to avoid: institutional vulnerability. To partner with students on the AI problem, the university would have to admit it does not have the answer, that its assessments are genuinely destabilized, that the border cannot be drawn — and to make that admission the starting point of a shared inquiry rather than a secret to be papered over with taxonomies. Broussard’s diagnosis of the culture’s relationship to these systems is apt: we need ”a new, more balanced view of AI” that neither venerates nor demonizes the technology but sees it plainly, institutions included [26]. The compliance apparatus is the opposite of a balanced view. It is a mechanism for not having to look.

[10] Evaluar en tiempos de IA: cómo salvar el aprendizaje cuando es una máquina la que hace las tareas

[29] Where AI Meets Accessibility: Considerations for Higher Education

[13] PDF Guide pratique - collimateur.uqam.ca

[1] A media literacy initiative in the West Bank advances critical thinking for an AI world

[26] Artificial Unintelligence

The Territory Beneath the Border

Return, finally, to the border. The most humane early commentary on this whole transition argued that generative AI would change education rather than destroy it [8], and I think that optimism can still be earned — but not on the current trajectory. The trajectory the week’s evidence reveals is one in which the university, faced with a dissolved border, responds by policing it harder: more taxonomies pretending the line is drawable, more detectors converting statistical guesses into disciplinary charges, more unvalidated frameworks hardening the

[8] ChatGPT is going to change education, not destroy it

founding misconception into infrastructure. Each of these responses is individually defensible and collectively catastrophic, because together they download the institution's authority into policing, expose its least-advised students to the detectors' patterned errors, and leave its research capacity to be bought out from under it while its attention is fixed downward on the essay.

The alternative is not laxity. It is a change of unit. Stop asking *who authored this artifact* and start asking *what capacity did this student build, and how do we know* — a question that dissolves the detection problem because it never depended on provenance in the first place. Stop treating integrity as a perimeter to be defended and start treating it as a capacity to be cultivated, which the restorative-integrity literature already knows how to do [21]. And name the extraction — the quiet transfer of the university's research talent to the firms selling it detectors — as the structural threat the compliance documents are, however sincerely, engineered to ignore [26].

[21] Preventive and Restorative Academic Integrity in AI-Assisted Learning Environments

[26] The Atlas of AI

The territory has changed. Authorship is now a relation, not a threshold; judgment is now the scarce thing, not the artifact that once proxied for it; and the institution's real adversary is not the undergraduate with a chatbot but the extraction that would leave it certifying capacities it can no longer teach. A university that spends this decade guarding the old border will find, when it finally looks up, that it has been meticulously defending a line on a map of a country that no longer exists — while the country itself was carried off in daylight, one hired researcher and one purchased detector at a time.

References

1. A media literacy initiative in the West Bank advances critical thinking for an AI world
2. A Systematic-Narrative Review of Online Proctoring Systems and a Case Study
3. Academic Integrity in the Age of AI: Why Clear Proctoring Rules Matter
4. AI Detection Tools and Academic Punishment: How Opaque Evidence Threatens Due Process
5. AI Proctoring Tools Are Failing Students: What Parents Must Know
6. AI tutoring outperforms in-class active learning: an RCT at a Nigerian university

7. AY2025-26 Campus AI Curriculum Task Force

8. ChatGPT is going to change education, not destroy it

9. Creating the AI-Enabled Community College

10. Evaluar en tiempos de IA: cómo salvar el aprendizaje cuando es una máquina la que hace las tareas

11. Governing generative AI in higher education: emerging policy approaches and considerations

12. GUERSENZVAIG - revistas.um.es

13. PDF Guide pratique - collimateur.uqam.ca

14. How AI is reshaping human skills and thinking

15. How universities are adapting to students' use of AI

16. L'IA générative dans l'enseignement supérieur : un danger pour l'effort cognitif

17. La détection de l'usage des systèmes d'IA génératives

18. MIT propone rediseñar la universidad para la era de la IA

19. Non-automatable cognitive skills in higher education in the age of AI

20. Por primera vez, la universidad más grande de América Latina hizo pública una investigación por irregularidades

21. Preventive and Restorative Academic Integrity in AI-Assisted Learning Environments

22. Prévenir l'aliénation intellectuelle à l'ère de l'IA : vers une pédagogie de l'autonomie

23. Rapport d'enquête - Usages de l'IA dans les pratiques étudiantes

24. Redesigning STEM Higher Education in the Era of Generative AI: From Curriculum Design to Classroom Practice

25. Report - AI and Education

26. The Atlas of AI

27. The largest study of AI use by undergrads is in, revealing disparities in access and in cheating

28. Un catedrático español denuncia fraude masivo con IA en un examen en la Universidad de Brown

29. Where AI Meets Accessibility: Considerations for Higher Education

30. ¿Cómo pueden responder los educadores cuando los estudiantes presentan contenido generado por IA como si fuera propio?