From Proctoring to Pedagogy: The Death of the AI Detector

Weekly Analysis — https://ainews.social

A student at Adelphi University submitted a paper. Software flagged it as machine-written. The accusation followed, the grade fell, and the matter did not end in a dean's office but in a courtroom — the first time an American judge was asked to rule on whether an algorithm's suspicion could stand as evidence of academic dishonesty [18]. It is worth pausing on how strange this is. For roughly three years, universities responded to generative AI the way a startled organism responds to a threat: by reaching for a countermeasure, buying it in bulk, and deploying it against their own students. The countermeasure was the AI detector, and it is now dying — not because faculties held a referendum on its ethics, but because it never worked, and the cost of pretending it did has begun to arrive as litigation, reputational damage, and a slow-dawning recognition that the whole apparatus was solving the wrong problem.

[18] Newby v. Adelphi: First AI-Detection Court Ruling

This essay is a survey of where that recognition has taken higher education this year, and the argument underneath the survey is simple. The detector was the material expression of an epistemological contract that assumed dishonesty — that positioned the institution as prosecutor and the student as suspect. Its collapse is not merely the retirement of a bad product. It forces the institution to decide what it was actually trying to protect. The tracker of AI-related academic cases now runs to a length that would have seemed absurd in 2023, documenting a steady accumulation of disputes in which detection scores functioned as verdicts [3]. What follows is an attempt to map the full terrain: who is saying what, where the institution's priorities pull against its stated pedagogy, and what the wreckage of the detection era reveals about the contract that has to replace it.

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

This publication has, in earlier surveys, done the responsible thing and laid out AI in higher education as a ledger of promise and peril — personalization and efficiency on one side, bias and privacy on the other. That balance was honest for its moment. It is no longer sufficient. The evidence has moved, and a balanced shrug now obscures more than it reveals. The detector did not turn out to be a double-edged tool; it turned out to be a broken one, and the interesting question is what the institution builds in the space its failure opens.

The Machine That Cried Wolf

Start with the technical fact that everything else depends on, because it is the fact that vendors worked hardest to obscure. AI detectors do not detect AI. They estimate the statistical probability that a text resembles machine-generated writing, and they do so with an error rate that would be disqualifying in any context where the stakes were taken seriously. The false positive is not an edge case; it is a structural feature. Detectors systematically flag writing that is grammatically clean, formulaic, or produced by someone thinking in a second language — which is to say they flag the writing of non-native English speakers, neurodivergent students, and anyone whose prose happens to be tidy [12]. The tool is not neutral machinery that occasionally errs. It errs in a direction, and the direction has a demographic.

[12] Faux positifs détecteurs IA : causes, impacts et solutions

Despite this, universities kept buying. A Nature investigation this year found institutions still leaning heavily on detection software to police student writing, treating a probabilistic score as though it were a fingerprint [26]. The gap between what the software can do and what it is asked to do is where the damage lives. When a number between zero and one hundred becomes the trigger for a disciplinary process, the institution has quietly outsourced its judgment to a vendor's confidence interval. Legal scholars have named the mechanism precisely: the detection score becomes an opaque piece of evidence that a student cannot inspect, cannot cross-examine, and cannot meaningfully rebut, which turns academic-integrity proceedings into a due-process failure dressed in the language of rigor [5].

[26] Universities are relying on AI-detection software to catch ...

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

Watch the move carefully, because it is the same move surveillance systems make everywhere. The score arrives with the aura of objectivity — it is a number, it came from a computer, it must be neutral — and that aura does the work of suspending the ordinary demand that an accuser prove the accusation. In [6], Kate Crawford describes how automated decision systems in welfare bureaucracies were used to flag "anomalous data patterns" and, on that basis alone, cut people off and accuse them of fraud. The structure in the seminar room is identical. An anomaly detector produces a pattern; the pattern becomes an accusation; the accusation becomes a burden the accused must somehow discharge, usually without access to the reasoning that generated it. What was sold as a tool for protecting academic standards turned out to be a machine for manufacturing suspects.

[6] The Atlas of AI

The courts have started to notice. The emerging body of AI-detection litigation shows a consistent pattern: institutions that treated a detector's output as dispositive have found themselves unable to defend the practice when a student pushes back with a lawyer

[4]. Commentators have described the resulting "score-as-verdict" logic as a ladder of escalating unfairness, in which each rung — the flag, the meeting, the sanction — inherits the unreliability of the rung below it [23]. The Adelphi case matters less for its specific outcome than for what it establishes: that the question is now justiciable, that a university can be made to explain in open court why it trusted a probability more than a person [2]. Few institutions want to answer that question under oath. That is the real reason the detector is dying.

[4] AI Detection Lawsuits: Every Student Case, Outcome, and What the Data ...

[23] Score-as-Verdict: The AI-Detection Due-Process Ladder

[2] Adelphi University accused a student of using AI to ... - Newsday

The Surveillance Contract and Its Chilling Effect

The detector did not arrive alone. It came bundled with a broader impulse toward surveillance that had already installed itself in higher education through remote proctoring — the webcam monitoring, the eye-tracking, the room scans, the flagging of a student who looked away from the screen too often. A careful ethical review of these systems concluded that the harms are not incidental glitches to be patched but consequences of the model itself: proctoring treats the student as a presumptive cheater and builds the entire testing experience around that presumption [21]. Detection software is proctoring's textual cousin. Both encode the same contract, and the contract reads: we do not trust you, and we have purchased instruments to prove it.

[21] Remote Proctoring Through an Ethical Lens: The Case Against ...

The corrosive effect of that contract is easy to miss because it operates on behavior the institution never sees. If a student knows that any trace of AI in their writing may trigger an accusation they cannot rebut, the rational response is not to develop sophisticated judgment about when and how to use these tools. The rational response is to hide, to avoid, to produce the blandest possible defensible text, or to route around the tool entirely and learn nothing about it. This is the chilling effect, and it is precisely the opposite of what an institution claiming to prepare students for an AI-saturated world should want. The surveillance regime blocks the very experimentation through which literacy develops. You cannot learn to collaborate with a system you have been taught to treat as contraband.

Here the discourse contains a genuine tension the institution has to hold rather than resolve. There is real evidence that unsupervised, unreflective AI use damages learning. The largest study to date of undergraduate AI use, out of Berkeley, found not only that adoption is widespread but that it is unevenly distributed — that access and use patterns break along lines that track existing advantage, and that the line between productive use and outright cheating is blurrier than

either enthusiasts or prohibitionists want to admit [24]. Reporting on cognitive offloading has documented students handing the hard, effortful parts of thinking — the struggle that is the actual site of learning — to a machine that performs them frictionlessly [27]. And faculty sentiment has curdled: a striking survey found that ninety percent of faculty believe AI is weakening student learning [1].

[24] The largest study of AI use by undergrads is in, revealing disparities ...

[27] University students offload critical thinking, other hard work to AI

[1] 90% Of Faculty Say AI Is Weakening Student Learning: How ... - Forbes

The critical point — and the place where most of the discourse loses its nerve — is that these concerns are arguments *against detection*, not for it. If the worry is that students are offloading their thinking, a false-positive-prone probability score does nothing to address it. Detection catches the wrong students, punishes the wrong behavior, and leaves the actual pedagogical problem untouched while generating the reassuring illusion that something is being done. The institution's genuine interest — that students continue to do the difficult cognitive work that constitutes an education — has been captured by a technology that measures none of it. This is the governance-versus-pedagogy imbalance in its purest form: an enormous investment in the machinery of catching, and almost none in the harder work of teaching.

The Quiet Retreat

If the detector were merely controversial, institutions would defend it. What is happening instead is quieter and more telling: they are walking away from it without announcement, folding detection into the background while foregrounding something they can actually stand behind. The retreat is legible in the governance documents. Brown University's committee on generative AI in teaching and learning produced a report that reads less like a prohibition and more like an attempt to articulate shared judgment — a framework for guided use rather than a warrant for surveillance [14]. The vocabulary has shifted from detection to integration, from catching to teaching, and the shift is not cosmetic. It reflects a recalculation of where the institution's credibility can safely rest.

[14] GAITL Committee Report

The most vivid case is legal education, where the University of Chicago Law School did something that looks, at first, like reactionary retrenchment and is actually the opposite. It banned laptops from first-year classrooms — a move easy to caricature as nostalgia — while simultaneously issuing a strategy that embraces AI as a competency law students must develop [25]. The logic, spelled out in the school's own statement, is that the answer to AI is not to police its traces in submitted work but to redesign where and how learning happens — to protect the analog space of live reasoning while explicitly teaching

[25] UChicago Law Bans Laptops from 1L Classrooms As Part of Sweeping New AI ...

the digital tools everywhere else [22]. This is not prohibition. It is a deliberate separation of the moments that require unassisted cognition from the moments that require fluency with the machine. It treats detection as beside the point, because it has stopped trying to verify output and started designing process.

The same reorientation is visible outside the elite American core, which matters, because the detection panic was largely an Anglophone-institution phenomenon and the alternatives are being theorized more richly elsewhere. French-language analyses of generative AI in academic teaching have moved past the catch-the-cheater frame toward strategies of responsible, structured integration, treating the technology as something to be metabolized into pedagogy rather than repelled at the gate [16]. Work coming out of the same tradition argues for strategic institutional integration that begins from learning goals rather than from anxieties about enforcement [28]. And professional bodies are formalizing the expectation that AI competency is now part of the baseline: mandatory AI learning is being written directly into continuing-education requirements for regulated professions, a signal that the "just detect and forbid" posture has no future in fields that credential for practice [17].

What unites these moves is a decision to stop asking "was this written by AI?" and start asking "did this student learn?" Those turn out to be almost entirely different questions, and the detector could only ever pretend to answer the first. The retreat from detection is the institution admitting, in its procurement and its policy language if not in its press releases, that it had been buying an answer to a question it did not actually care about.

What the Detectors Never Measured

The deepest indictment of the detection era is not that the tools were inaccurate or discriminatory, though they were both. It is that they trained institutional attention on a metric — presence or absence of machine text — that has no stable relationship to the thing universities exist to cultivate. Guidance on rethinking assessment in this environment keeps returning to the same insight: if an assignment can be completed by a chatbot without the student learning anything, the problem is the assignment, not the chatbot [7]. The take-home essay that asks for a competent summary of settled knowledge was always a weak proxy for understanding; it survived because it was cheap to assign and cheap to grade, and generative AI simply exposed the weakness that was there all along.

[22] Rethinking Legal Education in the AI Era

[16] L'IA générative dans l'enseignement académique

[28] Vers une intégration stratégique et responsable de l' ...

[17] Mandatory College-directed AI learning for 2026-27

[7] AI in class: time to rethink assignments

This is where UNESCO's framing becomes useful, because it insists on a distinction the detection discourse flattened. Generative systems are, in the framework's own terms, "stochastic in generating outputs or predictions, as the same inputs may lead to different outputs," which makes their content "potentially less trustworthy, especially for the teaching of factual and conceptual knowledge" [6]. The point is not that AI is useless in education but that its unreliability has to be *taught* — that the competency an institution owes its students is the judgment to know when a fluent, confident, wrong answer is a fluent, confident, wrong answer. A detector cannot impart that judgment. It can only punish the traces of a tool whose actual danger is epistemic, not evidentiary.

[6] AI competency framework for teachers - UNESCO

Meredith Broussard, in [6], argues for being "more honest and realistic about what goes into technology and what it takes to keep technology working," and warns against the technochauvinism that treats a computational tool as inherently more trustworthy than human judgment. The detection regime was technochauvinism applied to discipline: it assumed that a machine's assessment of another machine's output was more reliable than a teacher's reading of a student's work over a semester. The assumption inverts everything we know about how learning is actually observed. A teacher who has read a student's drafts, heard them talk in seminar, and watched an argument develop has evidence a detector cannot access and cannot weigh. The detector asked the institution to discount all of that in favor of a number, and for a while the institution complied, because the number was scalable and the judgment was not.

[6] Artificial Unintelligence

There is a governance lesson buried here that extends well past AI. Broussard's own account of a healthier path points to bodies like the AI Now Institute, founded to insist that these systems be evaluated for their social consequences rather than their technical elegance [6]. The detection episode is a case study in what happens when an institution procures a technology on the vendor's terms — accepting the vendor's definition of the problem, the vendor's metric of success, and the vendor's disclaimer of responsibility for false positives — instead of on the terms of its own stated mission. The IT procurement decision and the pedagogical decision were made in different rooms, by different people, answering to different pressures, and the detector is what you get when the procurement room wins.

[6] Artificial Unintelligence - How Computers Misunderstand

Rebuilding the Epistemological Contract

If the old contract assumed dishonesty, the new one has to be built on shared judgment — and shared judgment cannot be automated. This is the expensive part, the part that challenges grading norms and faculty training and the whole economy of how assessment gets done at scale, which is exactly why it was avoided for as long as a detector offered a cheaper substitute. The alternatives now emerging are not speculative; they are being implemented, and they share a common feature: they move the site of assessment from the artifact to the process, from the submitted text to the demonstrable understanding behind it.

The most direct of these is the return of the oral examination. Colleges are reintroducing in-person tests and viva-style defenses precisely because a student who can defend an argument aloud, in real time, under questioning, has demonstrated something no detector could ever confirm and no chatbot could sit for [9]. The oral exam is not a nostalgic retreat to a pre-digital purity; it is a targeted response to the actual gap the detector failed to close. When the assessment asks a student to think in front of you, the question of who or what drafted the prose becomes irrelevant, because the understanding is being tested directly rather than inferred from a document [19]. It is more labor-intensive than running a batch of papers through detection software, and that added labor is the point: it is the institution paying, in faculty time, for the trust it can no longer buy in a subscription.

[9] Colleges are turning to in-person tests, oral exams to combat AI

[19] Oral Exams & AI: A Practical Alternative to Take-Home Writing

Alongside the oral defense sits a family of designs that make AI use visible rather than forbidden — process portfolios that capture the evolution of a piece of work, and collaboration logs in which students document how they used AI, what they accepted, what they rejected, and why. This reframes the machine from contraband into a documented collaborator whose contribution the student must own and evaluate. It also aligns assessment with the literacy the institution claims to value, because the reflective account of *how* one used a tool is itself the evidence of judgment. The most ambitious versions integrate AI into the instructional core rather than the assessment periphery. Harvard Business School's deployment of custom tutor bots, purpose-built for specific courses and constrained to their material, treats the technology as a teaching instrument rather than a threat to be neutralized [11]. A Harvard physics course that built a tailored AI tutor reported that student engagement roughly doubled, a result that would be unthinkable inside a surveillance frame because it depends entirely on students trusting the tool enough to use it openly [20].

[11] Custom AI Tutor Bots Are Transforming Learning at HBS

[20] Professor tailored AI tutor to physics course. Engagement doubled.

None of this is free, and none of it scales the way detection promised to. Redesigning assessment around process demands faculty development, smaller assessment loads, and a willingness to tolerate the messiness of judgment where a number once offered false clarity. It is worth being honest that the institutions leading this shift — Chicago, Harvard, Brown — are wealthy, and that oral examinations and bespoke tutor bots are expensive interventions that an underfunded regional institution or an overloaded adjunct cannot simply adopt. The equity dimension is not decorative. If the detector was cheap and unjust, the alternatives risk being just but unaffordable, and the pedagogically superior path will remain the property of the well-resourced unless the shift is treated as an institutional investment rather than an individual heroism. The literacy contract cannot be a luxury good.

The Partnership That the Discourse Keeps Omitting

The most conspicuous absence in the entire year of debate is the student as a party to the contract rather than its object. Almost everything written about AI in higher education positions the student as either a threat to be contained or a beneficiary to be served, and almost nothing positions them as a partner whose consent and judgment the arrangement depends on. This is the missing collaborative framing, and its absence is why so many well-intentioned policies still feel like something done to students rather than with them.

A few voices are trying to supply it. There is a serious case, articulated in the discourse this year, for a student's *right to refuse* AI — the argument that literacy includes the informed choice not to use a tool, and that an institution mandating AI fluency has to respect the refusal as much as it teaches the use [8]. This matters because the enthusiasm-versus-prohibition binary hides a third position: that a genuinely literate person can understand a technology thoroughly and decline it on principle, and that a university worthy of the name protects the space for that decision. Detection made refusal impossible — you cannot conscientiously object to a tool that you are being surveilled for allegedly using. A partnership model makes refusal legible as a considered stance rather than a suspicious one.

[8] AI in Universities: Duty to Understand vs Right to Refuse

The partnership framing also has to reckon with whose interests the whole apparatus serves, and here the perspectives least represented in the mainstream discourse are the most clarifying. Feminist and Global South analyses of AI-supported learning insist that these systems are not neutral instruments but carriers of the assumptions, languages, and power relations of the institutions that build them, and that a

learning environment designed around them can quietly reproduce the very inequities it claims to dissolve [13]. This connects directly back to the detector's demographic error rate: a tool that disproportionately flags non-native English speakers is not a technical failure of an otherwise sound system but the predictable output of a system trained on a particular kind of writing and deployed as though it were universal. The critique that a francophone research tradition has advanced — of "sortir de la matrice," of getting out of the frame the technology imposes — is not an anti-technology stance but a demand that the institution decide the terms rather than accept the vendor's [10].

[13] Feminist and Global South perspectives on AI-supported ...

[10] COMPAiSS

And there is a warning in the partnership literature that the collaborative frame does not automatically produce collaboration. There is emerging concern that AI, even when embraced rather than policed, may be eroding the social core of learning — the messy, negotiated, interpersonal work of a team assignment reduced to a set of individually AI-completed fragments stitched together at the end [15]. Partnership, in other words, is not achieved by dropping the detector and declaring trust restored. It has to be designed for, protected, and continually renegotiated, because the same tool that can enable literacy can also hollow out the collaborative human practices that a university exists to sustain.

[15] Is AI quietly eroding the social core of student teamwork? - HEPI

The Contract the Institution Has to Choose

The death of the AI detector is not a defeat, though many faculty who invested faith in it will experience it as one. It is the removal of a false solution that allowed the institution to postpone a genuine decision. That decision is not technical and it will not be settled by procurement. It is a decision about what a university believes it is doing when it assesses a student's work: verifying compliance, or cultivating judgment. The detector let institutions pretend those were the same thing. They never were, and the accumulating litigation, the demographic bias, and the sheer pedagogical emptiness of the enterprise have made the pretense unsustainable [4].

[4] AI Detection Lawsuits: Every Student Case, Outcome, and What the Data ...

The MIT Press synthesis of AI ethics gestures toward what the alternative demands: not a bolt-on compliance module but a "more radically interdisciplinary and pluralistic" reformation of what it means to educate at all, in which ethical and epistemic judgment is woven into the disciplines rather than delegated to a detector or a policy [6]. That is a large ask, and it will strain grading norms, faculty workloads, and the budgets of institutions that cannot buy their way to trust. But it is the honest ask, and it has the singular virtue of aligning the institu-

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

tion's practices with the learning it claims to care about. The detector aligned nothing; it protected an anxiety.

What replaces it — the oral defense, the process portfolio, the documented collaboration, the tutor built for a specific course and constrained to its material — is harder, slower, and more expensive than a subscription to a probability engine. It also happens to be teaching. The institution spent three years asking software to tell it whether it could trust its students. The more useful question, the one the death of the detector finally forces, is whether it is willing to build the kind of relationship in which that question no longer needs to be automated at all.

References

1. 90% Of Faculty Say AI Is Weakening Student Learning: How ... - Forbes
2. Adelphi University accused a student of using AI to ... - Newsday
3. AI Cheating Lawsuits Tracker — Every Case, Who Won (2026)
4. AI Detection Lawsuits: Every Student Case, Outcome, and What the Data ...
5. AI Detection Tools and Academic Punishment: How Opaque Evidence Threatens Due Process
6. AI Ethics - The MIT Press Essential Knowledge series
7. AI in class: time to rethink assignments
8. AI in Universities: Duty to Understand vs Right to Refuse
9. Colleges are turning to in-person tests, oral exams to combat AI
10. COMPAiSS
11. Custom AI Tutor Bots Are Transforming Learning at HBS
12. Faux positifs détecteurs IA : causes, impacts et solutions
13. Feminist and Global South perspectives on AI-supported ...
14. GAITL Committee Report
15. Is AI quietly eroding the social core of student teamwork? - HEPI
16. L'IA générative dans l'enseignement académique
17. Mandatory College-directed AI learning for 2026-27

18. Newby v. Adelphi: First AI-Detection Court Ruling

19. Oral Exams & AI: A Practical Alternative to Take-Home Writing

20. Professor tailored AI tutor to physics course. Engagement doubled.

21. Remote Proctoring Through an Ethical Lens: The Case Against ...

22. Rethinking Legal Education in the AI Era

23. Score-as-Verdict: The AI-Detection Due-Process Ladder

24. The largest study of AI use by undergrads is in, revealing disparities ...

25. UChicago Law Bans Laptops from 1L Classrooms As Part of Sweeping New AI ...

26. Universities are relying on AI-detection software to catch ...

27. University students offload critical thinking, other hard work to AI

28. Vers une intégration stratégique et responsable de l' ...