Distributing Books Is Not Building Minds: The Hollow Core of AI Literacy

Weekly Analysis — https://ainews.social

A book placed in a child's hands is not a thought placed in a child's head, and a curriculum shipped to a thousand schools is not a thousand minds that have learned to think. This is the oldest confusion in education, and it has come roaring back wearing the costume of the new. Across this week's evidence, the institutions that have appointed themselves the teachers of AI literacy keep performing an elegant sleight of hand: they measure the reach of their materials and call it the growth of understanding. They count books distributed, contest entries received, warnings issued, and guardians trained, and they present these outputs as though provision and capacity were the same substance. They are not. And the gap between them is where the whole project of living wisely alongside these machines either happens or dies.

The confusion matters because "AI literacy" has become a phrase almost everyone endorses and almost no one defines the same way. To a vendor it means fluency with a product; to a security office it means suspicion of email; to a ministry it means a printed guide in every backpack; to a parent it means knowing which toys to fear. Each definition smuggles in a theory of what a person needs, and each theory quietly decides what will be taught and what will be left out. When Kate Crawford insists in [22] that artificial intelligence is never a purely technical domain but rather "technical and social practices, institutions and infrastructures, politics and culture," she is describing exactly the terrain that a serious literacy would have to cover — and exactly the terrain that the current campaigns flatten into a leaflet. The result is a field being hollowed out from the inside by the very organizations claiming to fill it.

[22] The Atlas of AI

This essay is an attempt to map that hollowing: to distinguish the competing meanings of literacy now in circulation, to show where the dominant model quietly substitutes the appearance of learning for the thing itself, and to name what a literacy adequate to democratic life would actually require. The stakes are not academic. A public that cannot calibrate its fear to evidence, that has been trained to watch rather than to reason, is a public that can be managed — and being managed is precisely the condition that literacy is supposed to end.

The Two Meanings of Literacy

Begin with the word, because the trouble starts there. When Paulo Freire attacked what he named the ”banking” model of education in [22], his target was a pedagogy that treats the learner as an empty account into which the teacher deposits pre-approved contents. The student receives, files, and stores; the teacher narrates; and the more completely the student accepts the passive role, the better a student they are said to be. Freire's objection was not sentimental. It was that banking education produces people who can recite but not question, who can hold the deposited material without ever acting upon the world it describes. Deposit is not development. Reception is not judgment.

[22] Pedagogia del oprimido

The dominant model of AI literacy is banking pedagogy with a chip in it. Its flagship gesture is distribution — books, curricula, guides, one-pagers, a comic for the children — and its implicit metric is coverage. A campaign for everyone hands out materials and reports the count of materials handed out, and somewhere in the translation ”we produced and shipped this content” becomes ”the population now understands.” The move is so smooth that its practitioners rarely notice they have made it. But the possession of a document about AI is no more evidence of AI literacy than the possession of a Bible is evidence of grace. This is the first meaning of literacy — literacy as provision, as materials-in-hand — and it is the meaning that currently commands the budgets.

Against it stands a second meaning: literacy as capacity, as the developed ability to reason about a world saturated with these systems. This is the harder thing, because capacity cannot be shipped. It has to be built through practice, and practice requires exactly what a leaflet cannot supply — repeated encounters with real, ambiguous, consequential cases in which the learner has to decide something and can be wrong. UNESCO's own framing gestures toward this when it observes, in its [22], that generative systems are ”stochastic in generating outputs,” so that ”the same inputs may lead to different outputs” and the content is ”potentially less trustworthy, especially for the teaching of factual and conceptual knowledge.” Read that carefully: it means the very tools being pressed into service as literacy instructors are epistemically unreliable narrators. A capacity-based literacy would take that unreliability as its central object of study. A provision-based literacy takes it as a footnote and hands out the pamphlet anyway.

[22] AI competency framework for teachers

The distinction between these two literacies is not merely a matter of emphasis. It determines whose competence gets counted. If literacy is provision, then success is a logistics problem, solved when the trucks

arrive. If literacy is capacity, then success is a judgment problem, and it can only be assessed by watching people reason under conditions of genuine uncertainty. The campaigns have chosen the logistics problem because it is measurable, fundable, and photogenic. The judgment problem is none of these things, which is precisely why it is the one that matters.

When the Medium Eats the Message

Nothing exposes the incoherence of provision-based literacy more sharply than what happens when a campaign asks people to *demonstrate* their understanding. Consider the now-familiar spectacle of the AI-literacy contest — the poster competition, the comic challenge, the "show us what you've learned" exercise — in which participants overwhelmingly reach for a general-purpose generative tool to produce their entry. When a demonstration of literacy is itself produced by the thing the literacy was supposed to render the participant independent of, the medium has eaten the message. The campaign has modeled, at scale, the very dependence it claimed to cure.

This is not a marginal irony; it is the whole confusion in miniature. To use a large language model to generate a comic about how to think critically about large language models is to enact banking pedagogy in its purest form: the content is deposited by the machine, filed by the human, and submitted as evidence of a capacity that was never exercised. The tools themselves are engineered to make this substitution effortless. OpenAI's own guidance in [4] and in [21] teaches users to elicit polished output through incantation rather than comprehension — a skill in operating the oracle, not in understanding what the oracle is or how it fails. Fluency of operation is being sold as literacy of judgment, and the two could hardly be more different.

[4] Best practices for prompt engineering with the OpenAI API

[21] Prompt engineering best practices for ChatGPT

The confusion is compounded by how little the operators actually know about the systems they operate. Anthropic's admission, in [9], that outside researchers are needed to understand real-world usage patterns is a quiet acknowledgment that even the builders do not fully grasp how their tools function in the wild. And the security guidance that accompanies these systems — AWS's own [6] — reveals a layer of data-handling risk that the cheerful contest entrant never sees. A literacy worthy of the name would begin here: not with "how do I make the machine produce a comic" but with "what is this thing doing with what I feed it, and why should I or should I not trust what it returns."

[9] Enabling independent research on how people use Claude

[6] Consideraciones de seguridad para los datos en la IA generativa

Broussard names the underlying disease in [22]: a technochauvinism

[22] Artificial Unintelligence

that assumes the computational solution is always the superior one, that reaching for the machine is itself a mark of sophistication. When a literacy campaign rewards participants for outsourcing their thinking to a generative model, it does not merely fail to inoculate against technochauvinism — it administers the pathogen and calls it the cure. The medium eats the message, and the campaign hands over the fork.

The Spectacle of the Deepfake

If the provision model fails at building capacity, the fear it distributes fails at accuracy — and here the map of AI literacy reveals its most dangerous distortion. The public imagination of AI harm has been captured almost entirely by the deepfake: the perfect synthetic video, the impersonated politician, the fabricated confession. Anxiety about manipulated media now runs extraordinarily high, and the literacy apparatus feeds it, producing an endless stream of guides on how to spot the fake. Content like [13] trains the citizen to squint at pixels, to hunt for the six-fingered hand and the uncanny blink, as though the central threat of the synthetic age were forensic.

[13] How to Spot AI Deepfakes & Fake News: Full Guide (2026)

The trouble is that this fear is badly calibrated. Most political manipulation requires no perfect forgery at all. As analyses of synthetic media in democratic contests have documented — [15] among them — the deepfake is often less effective than crude, cheap, plausible content that no one bothers to fabricate seamlessly because seamlessness is not what persuades. The EDMO white paper [19] makes the same point structurally: the damage flows less from any single flawless fabrication than from the flooding of the information environment with cheap synthetic plausibility, which corrodes the baseline assumption that anything might be real. Influence operations of the kind traced in [12] and in coverage of [17] succeed through volume, targeting, and repetition — not through the single unmaskable masterpiece the spot-the-fake guides train us to hunt.

[15] La democracia sintética en Iberoamérica y el 'deepfake'

[19] Generative AI and Disinformation: Recent Advances, Challenges

[12] How AI-Driven Influence Operations Attack Ukraine

[17] Municipales 2026 : IA, deepfakes et désinformation, la démocratie

Worse, the forensic literacy on offer is itself becoming obsolete faster than it can be taught. The market of detection and authenticity tools that citizens are urged to trust is, on honest inspection, a mixture of the genuine and the vaporous, as [1] concedes. To hand the public a detector and tell them to rely on it is to build literacy on sand — the same error, at civic scale, that the surveillance-and-detection arms race has already made in other domains, where the tools generate false confidence and fresh distrust in roughly equal measure, a dynamic [2] documents with grim clarity.

[1] AI Content Authenticity Tools: What's Real, What's Hype, and What Actually Helps

[2] AI detectors are creating a new era of distrust

Crawford's phrase for what the deepfake spectacle performs is pre-

cise: ”enchanted determinism,” a mode of attention that, as she argues in [22], makes us ”focus on the innovative nature of the method rather than on what is primary: the purpose of the thing itself” — and which, in doing so, ”obscures power and closes off informed public discussion, critical scrutiny, or outright rejection.” The deepfake is enchanting precisely because it is spectacular. It draws the eye, the headline, and the budget toward the exotic edge case while the ordinary machinery of manipulation — which needs no enchantment because it works — runs quietly underneath. A literacy calibrated to evidence would spend most of its energy on the ordinary. The literacy we have spends most of its energy on the spectacle.

[22] The Atlas of AI

The Mundane Machinery of Harm

Turn the coin over and the calibration failure repeats itself in reverse. If the deepfake commands a fear it does not deserve, the genuinely dominant vectors of AI-enabled harm command a shrug they have not earned. The most consequential misuse of these systems is not cinematic. It is dull, industrial, and profitable — and it is precisely the dullness that keeps it beneath the threshold of public alarm.

Phishing remains the workhorse. The overwhelming majority of security incidents still begin with someone being persuaded to click, to trust, to enter a credential — a fact so mundane that the platforms address it through routine help documents rather than headline campaigns, as Google’s own [10] and [7] quietly attest. Generative tools have made this ancient technique dramatically more scalable — better grammar, better personalization, better mimicry of a trusted voice — without changing its fundamental banality. There is no comic contest about phishing, because phishing is boring, and the literacy apparatus is drawn to spectacle. Yet the practice-based judgment that would actually protect a person operates precisely in this boring zone: the pause before the click, the second look at the sender, the mundane habit of doubt.

[10] Evitar y denunciar los correos de phishing - Ayuda de Legal
[7] Denuncia spam, suplantación de identidad (phishing) o software

Alongside phishing runs synthetic identity fraud, scaling under an expert culture of reassurance even as it metastasizes. The regulatory scaffolding meant to contain such harms is visibly inadequate; [3] answers its own title with a barely disguised no. And the systems themselves introduce a subtler species of mundane harm: the confident falsehood. The conceptual work in [18] and the practitioner guidance in [24] describe a failure mode with no villain and no forgery — just a machine producing plausible error at scale, absorbed by users who were never taught that plausibility and truth are different properties.

[3] Are Existing Consumer Protections Enough for AI?

[18] New sources of inaccuracy? A conceptual framework for studying AI hallucinations
[24] When AI Gets It Wrong: Addressing AI Hallucinations and Bias

This is the harm that touches the most people in the most ordinary moments, and it generates almost no fear at all.

Here is the diagnosis in a sentence: literacy dies between panic and shrug. At one end, the public is over-alarmed about the rare, spectacular forgery it will seldom encounter and could rarely counter anyway. At the other end, it is under-alarmed about the phishing email, the synthetic identity, and the confident hallucination it encounters daily. A calibrated literacy would invert both errors — dialing down the deepfake dread, dialing up attention to the mundane machinery — because judgment is a matter of proportion, and proportion is exactly what the spectacle-driven campaign destroys. The propaganda-parsing failures documented in [11] show the two ends meeting: the ordinary tool, trusted without calibration, quietly launders the extraordinary lie.

[11] How AI chatbots and AI overviews parse foreign propaganda

Watching the Vulnerable Instead of Building Them

There is a third substitution at work, and it may be the most corrosive, because it does not merely fail to build judgment — it replaces the project of building judgment with the project of surveilling its absence. When a population is deemed too vulnerable to reason for itself, the institutional response is not to develop that population's reasoning but to appoint watchers over it. The federal prevention posture toward children exemplifies the move: train the adults to monitor, filter, and flag, and treat the child's own developing discernment as a risk to be managed rather than a capacity to be grown.

The evidence that children are in fact reasoning about these systems — and want help doing so well — is not in short supply. UNICEF's [20] documents young people already using, questioning, and worrying about AI in their own lives, which is to say already engaged in the raw material of literacy. Meanwhile the environment they inhabit is being populated with systems designed to talk back to them intimately — the conversational plush toy examined in [14] is a machine that manufactures intimacy without understanding it. And the difficulty of evaluating whether these systems are even safe is documented in the child-safety research [5], which shows that expert judgment itself is strained by the systems' unpredictability.

[20] PDF Snapshot of AI Usage and Concerns Among Children and Parents

[14] Juguetes con IA: ¿Qué pasa cuando el peluche de tu hijo puede conversar con él?

[5] Beyond "I Can't Help with That": How Child Safety Experts Evaluate AI

The tempting institutional conclusion is more watching: more filters, more monitoring, more adult guardians positioned between the vulnerable and the machine. But surveillance of the vulnerable is not the same as the empowerment of the vulnerable, and confusing the two produces a population permanently dependent on its watchers.

UNESCO's [22] argues the opposite priority — that the goal is a person equipped to click wisely, to exercise their own critical faculty, rather than a person whose clicks are simply policed. The same document notes how AI systems encode and reproduce social patterns, observing that voice assistants "still only serve a limited number of spoken languages" and reproduce gendered defaults, which means the watched are often watched by systems that already misunderstand them. A literacy that responds to that misunderstanding by adding another layer of adult oversight, rather than by building the watched person's own capacity to recognize and resist the misunderstanding, has chosen management over emancipation.

[22] UNESCO Think Critically Click Wisely

This is where the surveillance instinct and the banking instinct converge. Both treat the learner as passive — one as an empty account to be filled, the other as a risk to be contained — and neither treats the learner as an agent whose judgment is the entire point. Freire's warning in [22] applies with full force: an education that positions people as objects to be acted upon, rather than subjects who act upon the world, produces exactly the docility it claims to protect against. The child watched but never taught to reason grows into the adult managed but never able to judge. And a democracy of the managed is a contradiction in terms.

[22] Pedagogia del oprimido

Whose Literacy Counts

Every definition of literacy answers a prior question: literacy in whose service, defined by whom? Follow the money and the authorship, and the current settlement reveals itself. The dominant frameworks are authored largely by the entities with the most to gain from a particular kind of fluency — vendors who benefit when literacy means comfort with their products, institutions who benefit when literacy means compliance with their guidance, and states who benefit when literacy means trust in their tools. The reader's own interest — understanding as a defense against being managed — appears in these frameworks only where it happens to coincide with someone else's.

Consider the opacity of the systems that citizens are increasingly governed by. In Chile, reporting under [8] describes a government deploying algorithms that no one effectively audits — decisions made about people by systems those people cannot inspect. A literacy adequate to democratic participation would have to include the capacity to demand that inspection, to ask who built the model, on what data, toward what end. But no vendor curriculum teaches the citizen to interrogate the vendor, and no state guide teaches the citizen to audit

[8] El Estado usa algoritmos que nadie fiscaliza: expertos alertan por falta de transparencia y control

the state. The literacy that would most serve the public is precisely the literacy that the funders of literacy have the least incentive to build.

This is why the choice of who defines literacy is not a bureaucratic detail but the whole game. Broussard's account in [22] points toward the alternative model of authorship — the independent, adversarial scrutiny represented by groups like the AI Now Institute, founded to examine these systems from outside the interests that build them. Independent research of the kind Anthropic gestures toward in [9] matters for the same reason: literacy that originates outside the interests being examined is the only kind that can be trusted to serve the examined rather than the examiner. And the encouraging fact-checking infrastructure emerging around contested elections — the AI-assisted verification described at [23] — shows what a public-interest tool oriented toward citizen judgment, rather than citizen management, can begin to look like.

[22] Artificial Unintelligence

[9] Enabling independent research on how people use Claude

[23] Verificador de Hechos Elecciones Perú 2026 — Fact Check con IA

The MIT Press volume on AI ethics gestures at the deeper reorientation required: not a bolted-on ethics module but a "more radically interdisciplinary and pluralistic" formation, a fuller *Bildung* that treats reasoning about these systems as inseparable from reasoning about power, culture, and one's own life. That is a definition of literacy authored in the learner's interest rather than the vendor's — and it is, tellingly, the definition least represented in the campaigns that dominate the budgets. Whose literacy counts is answered by whose literacy gets funded, and right now the answer is: not yours.

Literacy as Practice, Not Performance

Pull the threads together and a single pattern emerges across every substitution examined here. The provision model substitutes materials for minds. The demonstration contest substitutes the tool's output for the user's judgment. The deepfake spectacle substitutes exotic dread for calibrated attention. The surveillance posture substitutes watching the vulnerable for building them. And the vendor-authored framework substitutes the funder's interest for the learner's. In every case the same replacement is performed: the appearance of literacy stands in for the practice of it. Literacy is being performed rather than practiced, and the performance is convincing enough to satisfy everyone except the person who actually has to live in a world saturated with these systems.

What would practice look like instead? It would begin from the recognition, insisted upon in [22], that these systems are social and

[22] The Atlas of AI

political all the way down, so that understanding them means understanding power and not just prompts. It would calibrate fear to evidence, spending its attention on the phishing email and the confident hallucination rather than the rare cinematic forgery. It would build judgment through repeated, low-stakes encounters with the mundane and the ambiguous — the everyday cases in which a person must actually decide whether to trust an output and can learn from being wrong. It would treat children, and everyone else, as subjects developing their own discernment rather than objects to be watched. And it would hold both the panics and the reassurances accountable, refusing the deepfake dread and the synthetic-fraud shrug with equal firmness.

There are faint signals that some corners of the field understand this. The reorientation described in [16] — away from output-counting and toward social, practiced learning — points in the right direction, as does the epistemic honesty of a framework like UNESCO's [22], which at least names the stochastic unreliability that a serious literacy must confront rather than conceal. These are the beginnings of a capacity model. They are also, notably, the parts of the field that generate the fewest press releases, because building judgment is slow, unphotogenic, and impossible to count by the truckload.

[16] MIT AI Report Calls for Alternative Grading, More Social Learning

[22] AI competency framework for teachers

The choice between the two literacies is finally a choice between two kinds of public. A public supplied with materials and watched by guardians is a public that can be reassured, alarmed, and steered — a public managed. A public that has practiced judgment, calibrated its fears, and learned to ask who built the machine and to what end is a public that can govern itself in the presence of these systems rather than merely be governed through them. Freire understood in [22] that the difference between depositing and educating is the difference between adaptation and freedom. Distributing books is not building minds. Until the institutions that fund and define AI literacy grasp that distinction — until they stop counting what they have shipped and start cultivating what people can actually do — the hollow core will remain exactly where it is, at the center of a project loudly performing the competence it has quietly declined to build.

[22] Pedagogia del oprimido

References

1. [7] 2. [10] 3. [4] 4. [21] 5. [6] 6. [11] 7. [20] 8. [15] 9. [2] 10. [14] 11. [8] 12. [23] 13. [13] 14. [18] 15. [5] 16. [12] 17. [16] 18. [19] 19. [1] 20. [17] 21. [24] 22. [3] 23. [9] 24. [22] 25. [22] 26. [22] 27. [22] 28. [22]

[7] Denuncia spam, suplantación de identidad (phishing) o software ...

[10] Evitar y denunciar los correos de phishing - Ayuda de Legal

[4] Best practices for prompt engineering with the OpenAI API

[21] Prompt engineering best practices for ChatGPT - OpenAI Help Center

[6] Consideraciones de seguridad para los datos en la IA generativa

[11] How AI chatbots and AI overviews parse foreign propaganda

[20] PDF Snapshot of AI Usage

References

1. AI Content Authenticity Tools: What's Real, What's Hype, and What ...
2. AI detectors are creating a new era of distrust - The Verge
3. Are Existing Consumer Protections Enough for AI? | Lawfare
4. Best practices for prompt engineering with the OpenAI API
5. Beyond "I Can't Help with That": How Child Safety Experts Evaluate AI
6. Consideraciones de seguridad para los datos en la IA generativa
7. Denuncia spam, suplantación de identidad (phishing) o software ...
8. El Estado usa algoritmos que nadie fiscaliza: expertos alertan por falta de transparencia y control
9. Enabling independent research on how people use Claude
10. Evitar y denunciar los correos de phishing - Ayuda de Legal
11. How AI chatbots and AI overviews parse foreign propaganda
12. How AI-Driven Influence Operations Attack Ukraine
13. How to Spot AI Deepfakes & Fake News: Full Guide (2026)
14. Juguetes con IA: ¿Qué pasa cuando el peluche de tu hijo puede conversar con él?
15. La democracia sintética en Iberoamérica y el 'deepfake'
16. MIT AI Report Calls for Alternative Grading, More Social Learning
17. Municipales 2026 : IA, deepfakes et désinformation, la démocratie ...
18. New sources of inaccuracy? A conceptual framework for studying AI hallucinations
19. PDF Generative AI and Disinformation: Recent Advances, Challenges ... - EDMO
20. PDF Snapshot of AI Usage and Concerns Among Children and Parents
21. Prompt engineering best practices for ChatGPT - OpenAI Help Center

22. UNESCO Think Critically Click Wisely

23. Verificador de Hechos Elecciones Perú 2026 — Fact Check con IA | CONDOR

24. When AI Gets It Wrong: Addressing AI Hallucinations and Bias