Beyond AI Literacy: Teaching Intelligence Management

Sai Gattupalli, Ph.D.
Society and AI Research Group
sai@societyandai.org

April 2026

Abstract

What does it mean to teach students to govern AI rather than just use it? That is the question driving this paper. I argue that existing AI literacy frameworks, however well-intentioned, were designed for a world of AI tools. We now live in a world of AI judgment, where frontier models like Claude Mythos Preview (Anthropic, 2026) perform at expert human level across medicine, engineering, and scientific reasoning. The educational response required is not incremental. This paper advances *intelligence management* as a broader pedagogical and conceptual framework for directing, interrogating, and remaining accountable for AI outputs in consequential settings. The paper situates this framework in relation to my earlier concept of *promptgramming*, which described the emerging craft of structured interaction with generative systems (Gattupalli, 2024). Building on that earlier work, the present argument proposes intelligence management as a more comprehensive umbrella term for the forms of epistemic, operational, and civic authority that education must now cultivate. The paper proposes three structural shifts, a tiered competency framework grounded in learning science, and closes with generative questions for the field rather than settled conclusions.

Keywords: *Intelligence Management, Promptgramming, AI Literacy, Future of Education, Curriculum Redesign, Teacher as Orchestrator, On-Device AI, Epistemic Cognition, Educational Equity.*

The Reckoning We Have Been Avoiding

Recent developments in frontier AI systems illustrate where we actually are.

In April 2026, Anthropic announced Claude Mythos Preview, a model it described as representing a step change in AI performance. The announcement was not primarily about what the model could explain or generate. It was about what the model could *find*. According to Anthropic's technical documentation, Mythos Preview autonomously identified thousands of previously unknown vulnerabilities across every major operating system and web browser tested, and in several cases developed working exploits for them before the vulnerabilities had been patched (Anthropic, 2026). The same capabilities that make the model effective at patching security flaws also make it effective at exploiting them. Anthropic restricted access to roughly forty organizations and launched a coordinated defensive effort called Project Glasswing. The American spy agency, National Security Agency, is reportedly among those with access (Curi & Sabin, 2026).

This example matters because it illustrates something educators need to take seriously: the gap between what AI systems can now do and what our educational frameworks are preparing students for is no longer a gap. It is a chasm.

For years, we wondered whether AI would ever match human expert performance in complex domains. Recent evidence suggests that frontier models now score at the ninety-fourth percentile on graduate-level scientific reasoning benchmarks and resolve complex software engineering tasks at rates exceeding most junior practitioners (Anthropic, 2026). And yet the curriculum frameworks, teacher preparation programs, and assessment systems in which most students spend their educational lives were designed for a world where expertise was scarce, hard-won, and the legitimate object of years of formal training. That world has changed. The curriculum has not.

This paper argues that the right response is not alarm or minor adjustment. It is the deliberate construction of a new pedagogical aim: teaching students not to compete with AI at tasks AI now performs better, but to become the people qualified to direct it, interrogate its outputs, and take responsibility for what is done with its results. In this paper, that broader pedagogical aim is developed under the term *intelligence management*, which is proposed as the defining educational challenge of this decade.

The paper moves through several connected ideas. I first clarify what intelligence management is and how it differs from existing AI literacy frameworks. I then describe three structural shifts the imperative requires: a reconceived role for teachers, a curriculum built around simulation and consequence, and an equity infrastructure that reaches every student regardless of connectivity. I propose a tiered competency framework, address the strongest objections honestly, and close with generative questions rather than confident conclusions.

From Users to Governors: What This Transition Actually Requires

The Challenge. When intelligence management is described to educators, a common response is: "Isn't that just AI literacy with a new name?" The conflation is understandable, but it points to exactly the problem this paper is trying to name.

Current AI literacy frameworks, as codified by UNESCO and others, typically organize around four competencies: understanding how AI systems work conceptually, using AI tools effectively, evaluating AI outputs critically, and reflecting on social and ethical implications (UNESCO, 2021). These are genuine capacities worth developing. The trouble is the role they implicitly prepare students for. AI literacy prepares students to be competent, reflective users of AI tools. It does not prepare them to be the accountable human authorities operating inside systems where AI does the analytical, generative, and evaluative heavy lifting. Those are different roles. They require different preparation.

My earlier work on *promptgramming* addressed one important part of this shift. In that work, promptgramming named the art and science of deliberately structuring prompts so that human users could work more effectively with generative AI systems (Gattupalli, 2024). That concept remains useful, but it is narrower than the challenge now in front of education. Promptgramming concerns the craft of interacting with models. Intelligence management concerns the broader authority to frame problems, interrogate outputs, calibrate trust, and retain responsibility in environments where AI systems increasingly shape analysis, recommendation, and action.

To clarify the distinction between the more familiar discourse of AI literacy and the broader framework proposed here, Table 1 contrasts the two across several core dimensions.

The Response. Intelligence management, as the term is used here, refers to the capacity to exercise principled authority over AI systems in consequential contexts. It includes framing the questions a system should address, contesting its outputs on reasoned grounds even when one cannot replicate the underlying computation, recognizing characteristic failure modes of highly capable systems such as confident confabulation, optimization for measurable proxies instead of genuine targets, and overconfidence at the edges of the training distribution, and holding firm to the commitment that certain decisions require human accountability regardless of what the machine recommends.

Dimension	AI Literacy	Intelligence Management
Primary role	Competent and reflective user of AI tools	Accountable human authority operating in relation to AI systems
Focus	Understanding, using, and evaluating tools	Governing, interrogating, directing, and remaining responsible for system outputs
Core activity	Tool interaction	System oversight in consequential settings
Epistemic stance	Critical evaluation of outputs	Contestation, calibration, and accountable judgment under delegation
Context	General AI use across learning and work	High-stakes domains where AI performs substantial analytical or generative labor
Educational aim	Competent use	Responsible authority
Relation to promptgramming	May include basic prompting as a use skill	Incorporates promptgramming within a larger framework of governance and accountability

Table 1: Distinguishing AI literacy from intelligence management

This is not an exotic skill. It is what a senior physician exercises when she interrogates a diagnostic AI before acting on it, asking not just whether the output is plausible but whether it accounts for clinical particulars the training data did not represent. It is what a public defender exercises when he challenges an algorithmic risk score used to determine bail (Eubanks, 2018). What these professionals need is not mastery of mathematical internals. They need the epistemological authority to push back: the training, the conceptual framework, and the institutional support to act as a competent principal rather than a passive delegate.

The Reflection. The learning sciences offer real grounds for optimism here. Research on epistemic cognition, specifically how students develop beliefs about the nature of knowledge and the conditions under which claims warrant confidence, consistently shows that students can be taught to reason more carefully in complex, contested information environments (Barzilai & Chinn, 2018). Intelligence management extends this tradition into an urgent new domain. It is buildable through deliberate instructional design, structured practice, and teachers who model epistemic authority in real time. That is the foundational claim of this paper.

The Teacher Has Always Been the Point

The Challenge. A central question keeps returning: if a teacher's authority rests on knowing more than her students about history, biology, and mathematics, what happens to that authority when AI systems can outperform most human experts in those same domains, on demand, at any moment?

Selwyn (2019) has argued that most discussions of AI in education dodge this question by leaning on a comfortable division of labor: AI handles the routine, the teacher handles the relational. It sounds tidy. But it does not actually answer what the teacher's authority is grounded in, or what it becomes when the informational substrate she draws on is now machine-generated.

The Response. The teacher's authority in the intelligence management classroom cannot rest on superior content knowledge alone. It must rest on something AI cannot replicate: the capacity to model the intellectual virtues of intelligence management in real time, with real stakes, in front of students who are watching closely.

Consider a concrete example. A group of students asks an AI system to analyze a historical dataset on school discipline and identify patterns. The system produces a fluent, confident summary: certain demographic groups are disciplined at higher rates, the disparity is statistically significant, and the model suggests differential behavioral patterns in the student population. An AI-literate student might flag this as potentially biased. An intelligence manager asks a harder question: what assumptions does this model make about what behavior looks like, who gets to record it, and what counts as a disciplinary event? That answer requires knowledge of how administrative data is produced, who holds the power to label and categorize, and what historical context shapes the numbers (Eubanks, 2018). The teacher who walks her students through that interrogation is doing something no AI can do. She is showing what it looks like to hold authority over the machine.

As an orchestrator, this teacher designs learning environments where students and AI systems work together on complex problems, deciding which tools to deploy for which subtasks and intervening when the human-in-the-loop is at risk of being squeezed out by the convenience of automated answers. As an ethical mentor, she makes the principled grounds for contestation visible through demonstration and dialogue. These capacities are shown, not just told.

The Reflection. This reconception is not a diminishment of teaching. It is an elevation. The teacher becomes the highest-order intelligence manager in the room. The question for teacher education programs is whether they are prepared to produce that kind of educator at scale. Currently, most are not.

What the Curriculum Actually Needs to Do

If the teacher's role must change, so must the curriculum she orchestrates.

From Receiving Information to Interrogating It

The Challenge. The static textbook was a sensible artifact of an economy where curated content was scarce. That economy has dissolved. Students carry devices that can generate explanations of any concept at any level of sophistication, on demand, tailored to their prior knowledge. The textbook's function is not merely redundant. It is misleading about what learning is actually for.

The Response. The shift being called for moves students from the lower rungs of Bloom's revised taxonomy, remembering and understanding, to the upper ones: analyzing, evaluating, and creating. A student who reads a chapter on climate feedback loops and answers comprehension questions has demonstrated the capacity to receive and retain information. A student who uses a simulation environment to build and test a climate model, adjusts parameters, observes their effects, and defends her assumptions to a skeptical peer panel has practiced being an intelligence manager.

This is not speculative. Web-based simulation frameworks, accessible coding environments, and AI-assisted modeling platforms can be deployed in schools at modest cost today. Problem-based learning research consistently demonstrates that students who engage with complex, ill-structured problems in realistic contexts develop qualitatively different reasoning capacities than those trained through knowledge transmission alone. The student does not need to be a professional data scientist to build a model of how a policy change propagates through a social system, or how algorithmic recommendations shift depending on what gets counted as relevant data. She needs structured scaffolding, a teacher who guides the interrogation of the model's assumptions, and an assignment where the consequences of her choices are real and debatable.

The Reflection. The immersive simulation replaces the textbook not because it is more engaging, though it typically is, but because it places students in a different epistemic relationship to the content: one where the output is falsifiable, the parameters are interrogable, and productive uncertainty is the whole point.

Computational Thinking as the Language of Governance

The Challenge. Students have been asked to use algorithmic systems without being given a working vocabulary for how those systems produce their outputs. That gap matters more now than it ever has. The capabilities of a system like Mythos Preview are not magic. They are the downstream consequence of design choices about objectives, training data, and optimization pressure (Anthropic, 2026). Students who do not understand this have no basis for interrogating those choices.

The Response. Computational thinking, defined as the set of problem-solving methods that draw on concepts fundamental to computing, including abstraction, decomposition, pattern recognition, and algorithmic reasoning, provides the technical substrate for intelligence management. A student who understands how a classifier processes training data, how a model's objective function shapes its outputs, and how optimization pressure creates systematic blind spots is better equipped to contest algorithmic decisions than one who understands only the interface (Lee & See, 2004). This fluency can and should be developed across disciplinary contexts, not only in computer science classrooms. It belongs in biology, social studies, and mathematics. It is a cross-disciplinary governance literacy, not a technical elective.

The Reflection. Whether weaving computational thinking across disciplines is achievable at scale remains an open implementation question, but it is difficult to imagine intelligence management without it. This requires teacher preparation programs to treat computational thinking as a foundational literacy rather than a specialized skill. That is a significant structural ask, and one most programs are not yet making.

The Students Who Cannot Wait

The Challenge. There is a version of this curriculum that should genuinely trouble us. If intelligence management education is available only to students in well-resourced schools with reliable broadband and institutional subscriptions to frontier AI platforms, then we will have built an educationally exciting but socially catastrophic system. Uneven access to the environments where intelligence management is developed will encode a new and durable form of educational inequality, deepening every existing axis of disadvantage along lines of connectivity and income.

The Response. The equity infrastructure of this curriculum is not a supplementary consideration to be addressed after the pedagogical design is settled. It is a design constraint on every other decision.

The students most likely to lack reliable internet access are precisely the students for whom intelligence management education is most consequential. They are most likely to encounter AI systems making automated decisions about their lives, in benefits administration, hiring, and public services, with the least institutional support to contest those decisions (Eubanks, 2018). The intelligence management curriculum must be built to run entirely offline, with AI systems resident on local hardware and capable of supporting tutoring, adaptive feedback, and simulation environments without any network dependency.

Edge computing architectures that distribute AI inference to local devices are not speculative. They are technically mature today. The policy challenge is ensuring that institutions serving the least-resourced students are equipped with this infrastructure before the curriculum diverges irrevocably along connectivity lines. This requires deliberate public investment and procurement standards that treat offline capability as a non-negotiable requirement, not a premium feature.

The Reflection. The classroom being described is a deeply social, collaborative workshop. Students work in teams, argue with each other, and receive guidance from a human educator who is present not as an information source but as a model of intellectual integrity. AI systems are tools in that room, extraordinarily capable ones, but subordinate to the human purposes of the inquiry.

A Framework We Can Actually Teach

Drawing on research in epistemic cognition (Barzilai & Chinn, 2018), automation trust (Lee & See, 2004), and the urgent evidence from frontier AI capability (Anthropic, 2026), I propose a tiered framework of intelligence management competencies organized into three layers: epistemic, operational, and civic. Like the upper levels of Bloom's revised taxonomy, each layer presupposes the one before it. This is a developmental sequence, not a checklist, and it should be treated as such in curriculum design.

Layer One: Epistemic Competencies

1. **Epistemic Authority Under Delegation.** The capacity to evaluate AI outputs critically, even when the student cannot independently reproduce the underlying computation. This is the intelligence manager's most foundational skill: knowing when to trust a system's output, when to contest it, and how to distinguish confident confabulation from reliable inference (Barzilai & Chinn, 2018). It is developed through explicit instruction, structured practice with AI outputs that are wrong in instructive ways, and repeated modeling by teachers who demonstrate contestation rather than deference.

2. **Question Quality and Problem Framing.** The capacity to formulate problems in ways that extract genuine value from capable AI systems: knowing what you do not yet know and asking for it precisely, in a way that exposes rather than obscures your assumptions. Poor problem framing is among the most consequential and least discussed failure modes of human-AI collaboration, and teaching it requires disciplinary contexts where the stakes of framing choices are visible and debatable.

Layer Two: Operational Competencies

3. **Calibrated Skepticism.** Not reflexive distrust of AI outputs, but trained sensitivity to the specific failure modes of highly capable systems. Research on trust calibration in automated systems consistently shows that both over-trust and under-trust degrade human-AI team performance; the goal is appropriate reliance, calibrated to the actual reliability of the system in the specific context (Lee & See, 2004). Calibrated skepticism is best developed through structured encounters with AI outputs that are fluent, confident, and instructively wrong.

4. **Computational and Algorithmic Reasoning.** The capacity to understand how algorithmic systems make decisions, where their assumptions are embedded, and what data and design choices produce what kinds of outputs, at a level sufficient to interrogate those outputs and participate meaningfully in governance decisions. This is a reasoning literacy cultivated across disciplines, not only in computer science classrooms.

Layer Three: Civic Competencies

5. **Moral Non-Delegability.** The recognition that certain decisions must not be handed to machines, not because machines lack relevant information but because accountability is intrinsically human and cannot be technically offloaded without democratic consequences. Who is responsible when an AI tutoring system systematically underestimates a student's capacity? Who answers when an automated grading engine encodes a teacher's historical biases at scale? Teaching students to identify and protect the boundary of non-delegable moral agency is among the curriculum's most important civic functions.

6. **Societal Impact Reasoning.** The capacity to trace the consequences of AI-enabled decisions through social systems: to ask not only whether an algorithmic outcome is accurate but whether it is just, whose interests it serves, and what institutional arrangements would make it accountable (Eubanks, 2018). This competency sits at the intersection of computational literacy and civic education. It is the domain where the teacher as ethical mentor is most irreplaceable, because the questions it raises cannot be answered by any system and must be worked through in the company of people prepared to disagree.

A Note on Assessment. A complete intelligence management curriculum must eventually grapple with how these competencies are measured, and that is a genuinely hard problem. Standard testing approaches, built around correct answers and demonstrated retention, are structurally inadequate for competencies that are relational, contextual, and expressed through the quality of contestation rather than the correctness of recall. What principled assessment looks like, whether through portfolio-based evidence, performance tasks, or structured argumentation rubrics, remains an open question. It is not peripheral to this framework. It is one of the most important things we do not yet know how to do.

The Objections I Take Most Seriously

A proposal that does not reckon with its own vulnerabilities is not a scholarly proposal. It is advocacy. Two objections therefore deserve direct attention.

Objection 1: The Domain Expertise Problem. You cannot govern what you do not understand. The physician who contests a diagnostic AI recommendation brings deep domain knowledge to that contestation: knowledge of pathophysiology, clinical presentation, and the specific ways diagnostic tools fail. Strip away that domain knowledge and what remains is not intelligent management. It is uninformed skepticism, which may be worse than deference.

This objection is serious. My response is that it identifies a genuine constraint rather than a fatal flaw. Intelligence management competencies must be developed alongside domain knowledge, not instead of it. The curriculum being described does not replace disciplinary education; it reorganizes what disciplinary education is for. A student studying biology who also learns to interrogate the algorithmic assumptions of a diagnostic model is developing both domain knowledge and governance capacity at the same time. The challenge, and it is real, is designing learning environments where these two things reinforce each other rather than compete for instructional time.

Objection 2: The Privilege Problem. A second objection holds that this curriculum is implicitly designed for students who already have advantages: good schools, sophisticated teachers, reliable technology, and the cultural capital to imagine themselves as legitimate contesters of institutional AI systems. For students in under-resourced settings, whose encounter with AI is more likely to be an automated benefits denial than a collaborative classroom simulation (Eubanks, 2018), teaching epistemic authority without addressing the structural conditions that deny them actual authority is at best naive.

This objection also has real force. But one implication should be resisted: the suggestion that teaching governance capacity to under-resourced students is inappropriate because structural inequality has not yet been resolved. The students most likely to have AI systems making consequential decisions about their lives without their meaningful input are precisely the students who most need the intellectual and civic tools to contest those decisions. Delaying that education until structural conditions improve is not a justice position. It is an abandonment of the students who cannot afford to wait.

What I Do Not Yet Know. We do not have a validated account of what intelligence management development looks like across grade levels, or which pedagogical interventions reliably produce it. We do not know at what age students can meaningfully begin practicing epistemic authority over AI systems. We do not have large-scale classroom evidence. The epistemic cognition literature (Barzilai & Chinn, 2018) provides a theoretical foundation, but it was built primarily in non-AI contexts and requires substantial empirical extension. These gaps are named not to undermine the proposal but because naming them is the only honest scholarly posture available.

Questions That Should Keep Us Experimenting

I do not want to end this paper with a confident conclusion. I want to end it with the questions I find most alive, because I believe these are the ones the field needs to work on together.

- **What does development actually look like?** At what age, and with what prior knowledge, can students meaningfully begin practicing epistemic authority over AI systems? What does intelligence management look like at the elementary level, the middle school level, the secondary level? We need longitudinal research that does not yet exist.
- **How do we assess what cannot be scored?** If intelligence management competencies are relational, contextual, and expressed through the quality of contestation rather than the correctness of recall, what would a valid and equitable assessment of those competencies actually look like? Can portfolio-based or performance-task approaches scale to the level of accountability systems demand?
- **Can domain knowledge and governance capacity grow together?** The domain expertise objection is real. What instructional designs allow students to develop disciplinary knowledge and intelligence management competencies at the same time, rather than trading one for the other?
- **How do we reach the students who need this most?** What policy mechanisms, procurement standards, and infrastructure investments would ensure that on-device intelligence management education reaches students in the least-connected communities before the curriculum diverges irrevocably along lines of connectivity?
- **Who decides what good intelligence management looks like?** The competency framework I have proposed reflects particular assumptions about what counts as principled authority, appropriate skepticism, and non-delegable moral agency. Whose assumptions are those? How do we build a framework that is genuinely responsive to the communities and contexts in which it will be implemented?

Managing intelligence is already the defining skill of some of the most consequential professional and civic roles in society. A system like Mythos Preview is not the end of that story. It is a signal of how quickly the story is moving, and how much catching up our educational institutions still have to do (Anthropic, 2026). The question is whether our schools will respond to that reality deliberately and equitably, or whether they will continue preparing students for a world that no longer exists.

I am curious, and a little impatient, to find out what we will choose. I hope you are too.

References

Anthropic. (2026). *Claude Mythos Preview: Security capabilities and Project Glasswing*. Anthropic Red Team Blog. https://red.anthropic.com/2026/mythos-preview/

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences, 27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Curi, M., & Sabin, S. (2026). Scoop: NSA using Anthropic's Mythos despite blacklist. *Axios*. https://www.axios.com/2026/04/19/nsa-anthropic-mythos-pentagon

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Gattupalli, S. (2024). *The art and science of promptgramming*. Scholarworks@UMass. https://doi.org/10.7275/mk13-q340

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80.

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

UNESCO. (2021). *AI and education: Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. https://doi.org/10.54675/PCSP7350