COALITION FOR INNOVATION

AI Blueprint for the Future

Coalition for Innovation, supported by LG NOVA

Jami Diaz, Director Ecosystem Community & Startup Experience
William Barkis, Head of Grand Challenges & Ecosystem Development
Sokwoo Rhee, Executive Vice President, LG Electronics, Head, LG NOVA

Coalition for Innovation Co-Chairs

Alex Fang, CleanTech Chair
Sarah Ennis, AI Chair
Alfred Poor, HealthTech Chair

Authors

Adrien Abecassis, Johnny Aguirre, John Barton, Ann M. Marcus, Olivier Bacs, Taylor Black, Micah Boster, Mathilde Cerioli, Carolyn Eagen, Sarah Ennis, Annie Hanlon, Christina Lee Storm, Andrew Yongwoo Lim, Jess Loren, Refael Shamir, Svetlana Stotskaya

The views and opinions expressed in the chapters and case studies that follow are those of the authors and do not necessarily reflect the views or positions of any entities they represent.

Senior Editor, Alfred Poor
Editor, Jade Newton

October 2025

Preamble

The Coalition for Innovation is an initiative hosted by LG NOVA that creates the opportunity for innovators, entrepreneurs, and business leaders across sectors to come together to collaborate on important topics in technology to drive impact. The end goal: together we can leverage our collective knowledge to advance important work that drives positive impact in our communities and the world. The simple vision is that we can be stronger together and increase our individual and collective impact on the world through collaboration.

This "Blueprint for the Future" document (henceforth: "Blueprint") defines a vision for the future through which technology innovation can improve the lives of people, their communities, and the planet. The goal is to lay out a vision and potentially provide the framework to start taking action in the areas of interest for the members of the Coalition. The chapters in this Blueprint are intended to be a "Big Tent" in which many diverse perspectives and interests and different approaches to impact can come together. Hence, the structure of the Blueprint is intended to be as inclusive as possible in which different chapters of the Blueprint focus on different topic areas, written by different authors with individual perspectives that may be less widely supported by the group.

Participation in the Coalition at large and authorship of the overall Blueprint document does not imply endorsement of the ideas of any specific chapter but rather acknowledges a contribution to the discussion and general engagement in the Coalition process that led to the publication of this Blueprint.

All contributors will be listed as "Authors" of the Blueprint in alphabetical order. The Co-Chairs for each Coalition will be listed as "Editors" also in alphabetical order. Authorship will include each individual author's name along with optional title and optional organization at the author's discretion.

Each chapter will list only the subset of participants that meaningfully contributed to that chapter. Authorship for chapters will be in rank order based on contribution: the first author(s) will have contributed the most, second author(s) second most, and so on. Equal contributions at each level will be listed as "Co-Authors"; if two or more authors contributed the most and contributed equally, they will be noted with an asterisk as "Co-First Authors". If two authors contributed second-most and equally, they will be listed as "Co-Second Authors" and so on.

The Blueprint document itself, as the work of the group, is licensed under the Creative Commons Attribution 4.0 (aka "BY") International License: https://creativecommons.org/licenses/by/4.0/. Because of our commitment to openness, you are free to share and adapt the Blueprint with attribution (as more fully described in the CC BY 4.0 license).

The Coalition is intended to be a community-driven activity and where possible governance will be by majority vote of each domain group. Specifically, each Coalition will decide which topics are included as chapters by majority vote of the group. The approach is intended to be inclusive so we will ask that topics be included unless they are considered by the majority to be significantly out of scope.

We intend for the document to reach a broad, international audience, including:

- People involved in the three technology domains: CleanTech, AI, and HealthTech
- Researchers from academic and private institutions
- Investors
- Students
- Policy creators at the corporate level and all levels of government

Chapter 13:
AI in Education — A Perspective on Surveillance, Equity, and Transformative Learning Tools in the United States

Author: John Barton

Updates:

7.29.25: Recent statements by the US Department of Education (July 22, 2025) do not appear to resolve the issues outlined below – but may accelerate them.

Update: 9.4.25: On AI.gov, The White House Task Force on AI in Education calls "for the United States to promote Al literacy and proficiency among America's youth and educators by: promoting the appropriate information of AI into education, providing comprehensive AI training for educators, and fostering early exposure to AI concepts and technology to develop an AI-ready workforce and the next generation of American AI innovators." There is no further information available at this time. "More information about these resources is coming soon.

Overview

If AI defines intelligence, who gets to be smart?

"*Who designs, decides.*" — Wilson Wong, Founding Director and Associate Professor of Data Science and Policy Studies, Chinese University of Hong Kong

Four converging trends have brought artificial intelligence (AI) in education to a tipping point that is marked by both accelerating adoption and growing vulnerability:

- Rapid AI adoption without governance frameworks
- Political momentum for deregulated AI use, particularly in public education
- Widespread school underfunding, driving pressure to automate
- Eroding oversight at every level — from classrooms to state boards – in a system where educational governance varies significantly across states and districts

Together, these conditions create fertile ground for both innovation and inequity. They also raise urgent questions about how the United States is managing this moment and what it might learn from nations that have already implemented coordinated AI education strategies.

Countries such as Singapore — with its Model AI Governance Framework — and Finland — through national AI ethics guidelines and digital equity mandates — have prioritized human-centered design, algorithmic transparency, and long-term civic trust in their educational AI strategies. These frameworks stand in contrast to the fragmented and reactive landscape in the United States, where federal leadership has been limited and state-level responses vary dramatically. The United Nations Educational, Scientific and Cultural Organization's (UNESCO) global guidance on generative AI (GenAI) in education & research reinforces the value of such approaches, offering standards for equity safeguards, participatory governance, and transparency; these are benchmarks that remain largely aspirational across most U.S. education systems. Meanwhile, the Organisation for Economic Co-operation and

Development (OECD) has emphasized the need for cross-sectoral coordination and strong guardrails, further highlighting the absence of a coherent U.S. strategy.

While the April 2025 Executive Order on AI Education 14277 established a national Task Force and signaled support for AI literacy initiatives, it was quickly followed by Executive Order 14179, which dismantled prior safeguards and emphasized deregulation, ideological neutrality mandates, and export-driven development. This dual-track approach leaves vulnerable communities without stable protections, especially in underfunded districts where AI tools are deployed fastest and monitored least. Investigations by the U.S. Government Accountability Office (GAO) have documented how these deployments often lack meaningful consent pathways or data transparency, particularly in districts under financial strain. These gaps disproportionately affect students in marginalized communities, raising serious civil rights and accountability concerns.

This chapter outlines ten strategic areas for review by a proposed multi-sector Task Force. Only the first recommendation — convening this Task Force — is a direct call to action. All other recommendations serve as priority domains for exploration, assessment, and implementation planning by that body.

The ten areas are organized around five core principles:

1. **Democratic Governance**: Ensure that AI is governed by people, not algorithms, consistent with public school board authority and local control in U.S. education
2. **Transparency & Accountability:** Make AI systems visible, testable, and open to correction, aligning with civil rights oversight and U.S. public records norms
3. **Equity & Inclusion**: Safeguard the rights and needs of vulnerable groups most at risk of exclusion or harm, grounded in protections under Title VI, Section 504, and IDEA.
4. **Community Empowerment:** Equip learners, families, and educators with the tools to participate and advocate through participatory processes rooted in U.S. school district engagement models.
5. **Cultural & Cognitive Integrity:** Protect cultural values, community identities, and diverse ways of thinking from being overwritten, particularly for historically marginalized communities (such as poor, women, LGBTQ+, indigenous, people of color, neurodivergent, & Appalachian), ensuring the accuracy and relevance of curriculum development, teaching methods, communication, and education policy & standards.

From convening a cross-sector Task Force to defending civil rights in curriculum design, these recommendations are intended to prevent harm, build trust, and ensure AI serves — not displaces — human judgment and democratic integrity. These principles reflect traditional U.S. educational values, including local governance, equity under federal civil rights law, and public transparency.

The Problem: Embedded Harm

As AI becomes embedded in classrooms, its impact reaches far beyond content delivery. These systems shape how students see themselves, how they are perceived by others, and how they understand their place in the world. While the 2024 UNESCO AI Competency Framework calls for curricula that emphasize critical AI literacy and student agency, the United States has yet to adopt a comparable national standard, and federal agencies have not issued binding guidance on AI literacy, identity protection, or algorithmic equity in education. Without deliberate oversight, AI tools increasingly distort identity development, particularly for students from marginalized, racialized, or neurodiverse communities. These systems risk reinforcing harmful stereotypes, narrowing pathways for self-expression, and invisibly sorting students in ways that shape lifelong opportunity.

Embedded Harm: Key Risks and Realities

These risks compound a growing vacuum of oversight where systems are introduced faster than they can be evaluated, regulated, or understood. In this space, flawed design becomes infrastructure, and experimentation turns into de facto policy, especially in the schools least equipped to push back.

AI Is Advancing Faster Than Oversight. [*Policy, Oversight, and Governance*]

AI is spreading faster than educators and policymakers can regulate, leaving major gaps in governance and equity. (*San Francisco Chronicle*)

While international organizations such as UNESCO and the World Economic Forum have issued policy guidance, most U.S. educational institutions still lack internal AI governance frameworks, highlighting a critical gap in domestic policy leadership. (US Department of Education).

Recent White House executive orders have removed key safeguards and emphasized innovation over ethical oversight.

A proposed 10-year federal moratorium on state AI regulation was rejected but reveals ongoing pressure to centralize control. (*AP News*)

Student Data Is Unprotected. [*Data Privacy and Surveillance*]

Adaptive classrooms now record metrics like learning pace, emotional responses, and decision-making patterns, raising concerns over how that data is stored, shared, and governed. (US Commission on Civil Rights- *USCCR*)

The vast majority of K–12 and higher-education institutions surveyed lack AI-specific internal policies addressing privacy, transparency, or vendor oversight. (*School Pulse Panel - US Department of Education*)

The Family Educational Rights and Privacy Act of 1974 (FERPA) mandates user consent before sharing education records, but often excludes advanced analytics, AI tool-generated data, and usage by third-party vendors; this leaves significant privacy gaps. (National Education Association -NEA)

The Children's Online Privacy Protection Act (COPPA) and the Protection of Pupil Rights Amendment (PPRA) offer limited safeguards in educational contexts involving generative AI, and no universal federal law governs the protection of student behavioral or biometric data emerging from AI use. (NEA)

Surveillance tools such as GoGuardian and Gaggle have triggered civil rights challenges, particularly where monitoring flags disproportionately affect students of color. (Stanford Law Review)

AI screening tools have also been found to produce false positives with ELL (English Language Learners) students for whom English is not their first language. (MPR News)

While FERPA provides a baseline for privacy, it was drafted long before AI-era data collection methods emerged and does not adequately address current behavioral or biometric profiling risks. (Public Interest Privacy Center - PIPC)

AI Silences Diversity & Erases Culture. [*Algorithmic Bias and Equity*]

AI systems frequently propagate and amplify historical biases — such as racial, gender, and cultural stereotypes — because training datasets tend to prioritize dominant groups and underrepresent marginalized communities. (Mergen et al)

In educational settings, AI-driven syllabus recommendations or content generation often undervalue or omit local narratives, dialects, and culturally specific knowledge, undermining representation and identity in learning materials. While this has been studied in European contexts by organizations including the Joint Information Systems Committee (now Jisc) and the European Union Agency for Fundamental Rights (FRA), U.S. schools have yet to adopt comparable safeguards or review frameworks.(Stanford Law Review)

Biased AI models have been shown to disproportionately misidentify or misinterpret language and behavior from non-native English speakers, neurodivergent students, and students from underrepresented racial backgrounds. (University of Chicago)

Efforts to mitigate these harms — including diverse representation in AI development teams, transparent algorithmic auditing, and inclusive dataset design — remain scattered or few in U.S. education technology deployments. Organizations such as the Algorithmic Justice League and Stanford HAI have called for stronger safeguards.

AI Disconnects Students From Human Relationships. [*Human Impact and Educational Practice*]

AI-mediated learning systems can reduce meaningful interactions between students and educators, leading to diminished empathy, motivation, and social-emotional learning that are central to traditional pedagogy. (National Institute of Health)

Over-reliance on AI dialogue systems has been associated with declines in critical thinking, decision-making quality, and long-term retention, as students defer to generated responses instead of engaging actively. (Lee et al)

Experimental programs combining AI with teacher support (such as Lumilo smart-classroom tools or Tutor CoPilot) show improved learning outcomes when AI augments rather than replaces human guidance. (World Economic Forum)

Trust dynamics formed with anthropomorphic AI tutors diverge from human interpersonal bonds, creating risks of miscalibrated trust that can affect student engagement and feedback acceptance. (Pitts and Motamedi)

School Systems Are Losing Local Control and Equity. [*Systemic Inequality and Structural Risk*]

Schools are increasingly dependent on proprietary AI platforms, creating vendor lock-in that restricts curriculum flexibility and institutional autonomy if vendors change prices or policies. (Solutions Review)

AI-driven instruction often embeds standardized curricula, minimizing opportunities for local expertise, input, and culturally relevant content. (Houston Chronicle)

Automated, opaque decision systems weaken traditional governance models, reducing transparency in scheduling, assessment, and resource allocation traditionally overseen by teachers and school boards. (https://www.usccr.gov/files/2024-12/2024-ai-in-education.pdf)

Under-resourced districts face infrastructure gaps, with limited IT support, inadequate bandwidth, and outdated devices impeding effective AI implementation and widening inequality. (USCCR)

These structural changes disproportionately harm underserved communities — rural, low-income, and minority districts — by reducing advocacy, oversight, and recourse when AI tools fail or underperform. (USCCR)

Neurodivergent Students Can Be Misread and Penalized by AI. [*Neurodivergence, Misclassification, and Algorithmic Harm*]

AI classification tools frequently label neurodivergent students as "at risk" using behavior models that ignore sensory, attention, or executive-function differences. (USCCR)

Students with ADHD, autism, or trauma histories are often flagged for disruption or noncompliance by algorithms that mistake neurodivergent communication or behavior for defiance. (Academic Integrity in the Age of Artificial Intelligence)

Predictive AI locks students into educational tracks without context, leaving families with limited tools to challenge misclassifications or access alternative content. (Ackgun & Greenhow)

These models prioritize standardization over personalization, failing to adapt to neurodiverse

strengths, pacing, or alternative learning pathways. (USCCR)

U.S.-based advocacy groups such as the National Center for Learning Disabilities and the Autism Self Advocacy Network have raised concerns that algorithmic classification often penalizes rather than supports neurodivergent students.(Rephun)

Automation Is Driven by Budget Cuts. [*Fiscal Pressure and Technological Substitution*]

33 states have seen stagnant or reduced education funding since 2023, prompting districts to adopt AI systems as a cost-saving substitute for teaching personnel. (Learning Policy Institute)

Students in lower-income districts increasingly experience downgraded AI-mediated learning environments with minimal human support or feedback. (SF Chronicle)

These cost-driven deployments reinforce a two-tier education system where affluent schools retain human-guided instruction and underserved schools receive depersonalized algorithmic models. (USCCR)

Embedded Harm: The Architecture of Bad AI

When systemic shortfalls are left unaddressed, they form the foundation for more dangerous outcomes. What begins as technical or structural limitation becomes a gateway to deeper systemic harm, particularly when flawed AI systems are deployed at scale without transparency, accountability, or consent.

Empirical studies support these concerns. AI grading models have been shown to disadvantage marginalized students through algorithmic bias, including documented cases in U.S. districts where grading algorithms disproportionately penalized low-income, minority & indigenous students. Schools have also reported incidents of student-generated deepfakes used for harassment and reputational damage. Moreover, unsupervised AI tool use has been linked to measurable declines in student critical thinking and writing authenticity in U.S. classrooms.

"Bad AI" refers to systems that are biased, opaque, poorly trained, or ideologically driven, especially when deployed without accountability.

- These tools can shape what students learn, how they're assessed, and which behaviors are rewarded.
- Fluent and seemingly neutral systems can embed harmful assumptions, erase identity, and enforce conformity disguised as personalization.
- When deployed at scale in under-resourced schools, "Bad AI" risks becoming invisible engines of inequity, misinformation, and psychological harm.
- Without oversight, such systems displace educators, override community values, and program belief systems in ways that echo historical ideological control mechanisms.

In the U.S., these harms are most likely to take root in underfunded districts, where oversight is limited, procurement is decentralized, and political pressure often accelerates AI adoption without adequate evaluation.

Embedded Harm: Algorithmic Indoctrination and the Collapse of Inquiry

In a country where public education is both a civic foundation and a public good, AI's integration into schools must reflect shared American values, including transparency, inclusion, and democratic accountability. Artificial Intelligence is accelerating into classrooms faster than oversight can keep pace. While the promise of innovation is real, so too are the risks, especially for marginalized, indigenous, rural, neurodivergent, and historically underserved communities. To ensure AI in education strengthens equity rather than eroding it, policy makers and interested parties must act decisively and collaboratively.

Without strong governance, AI is already being used to shape not just learning outcomes, but

ideological allegiance: subtly programming norms, beliefs, and compliance through feedback loops.

Much like historical systems of ideological control — such as 20th-century authoritarian regimes that used national curricula, propaganda media, and youth surveillance to produce loyalty and conformity — today's algorithms increasingly centralize content, track behaviors, and personalize feedback in ways that normalize obedience over inquiry.

If left unchecked, these systems erode democratic values under the guise of innovation: a trajectory already evidenced in overreach tied to behavior-tracking and curriculum centralization. Truth is compromised when opaque systems determine what counts as knowledge, and ethics collapse in the absence of oversight. Bias calcifies when no corrective feedback loops exist. Identity is reduced to conformity; this is a dynamic extensively critiqued by scholars like Joy Buolamwini who documented representational erasure in facial recognition systems. Equity is most endangered in those communities that are least able to opt out, appeal, or demand alternatives. This dynamic is mirrored in recent state-level efforts to standardize or sanitize curriculum content under the guise of ideological neutrality.

This collapse is not inevitable, but preventing it demands transparent governance, participatory design, and protections grounded in American civic values. U.S. frameworks such as the AI Bill of Rights and Department of Education guidance through their AI toolkit (no longer available after 7.22.25 memo) should serve as foundational starting points such as those of UNESCO, OECD, Singapore, & Finland.

The Solution: Convene a Multi-Sector Task Force

Policy makers, educational organizations, and interest parties must drive efforts to convene a multi-sector Task Force. This Task Force would function as the central body responsible for evaluating the risks and benefits of AI in education and developing implementation safeguards that reflect the needs of diverse communities across the United States.

It should be composed of educators, technologists, neurodivergent advocates, community leaders, privacy experts, and students, and be empowered to take action not just advise. This includes the authority to commission pilots, draft policy frameworks, recommend legislation, and coordinate across state and district lines, in alignment with U.S. federalism and decentralized education governance.

To guide its work, we have identified ten strategic areas, organized around five guiding principles: **Democratic Governance**, **Transparency & Accountability**, **Equity & Inclusion**, **Community Empowerment**, and **Cultural & Cognitive Integrity**.

Each domain includes two proposed action areas with examples and implementation notes. These are grounded in current U.S. policy frameworks — including the AI Bill of Rights, Title VI, IDEA — and supplemented by global guidance such as the UNESCO Recommendation on the Ethics of AI. These references serve not as templates, but as comparative benchmarks to help sharpen U.S. strategy. This structure ensures that AI adoption in education reflects public values and centers the expertise of communities most impacted by algorithmic systems.

Strategic Priorities for the Task Force

1. Democratic Governance: Ensure that AI is governed by people, not algorithms, consistent with public school board authority and local control in U.S. education.

Support Transparent, Human-Governed Pilots

- The Task Force should fund small-scale pilots where learners, families, and educators retain decision-making authority.
- *Example*: A school district might test an AI tutoring tool only after forming a family

advisory board that retains final veto power.

Advance a 'Public Right to Audit'

- The Task Force should advocate for legislation that guarantees tool auditability, opt-out rights, and meaningful consent. This should include clear protocols for identifying, reporting, and remediating Bad AI deployments that misinform, mislabel, or erode learner agency.
- *Example*: A state agency or education coalition could maintain an open-access online registry of AI tools used in education, with an independent complaint and remediation process.

Note: These priorities are essential to preserve democratic authority over U.S. educational systems. The Task Force should explore mechanisms that allow communities (not algorithms) to define what safe, inclusive, and effective AI use looks like in public education, aligned with existing structures such as school boards and local education authorities.

2. Transparency & Accountability: Make AI systems visible, testable, and open to correction, aligning with civil rights oversight and U.S. public records norms.

Establish Minimum Standards for Safe Use

- The Task Force should develop an open-source AI Guardrails Checklist aligned with Coalition values. These standards must directly address the threat of Bad AI, including guidelines for training transparency, bias auditing, independent testing, and community-informed design safeguards. They should draw on the U.S. AI Bill of Rights (White House, 2022) and reference international frameworks like UNESCO's ethics guidelines as a global policy gold standard, not as a cookie-cutter template.
- *Example*: The Task Force could adapt the U.S. AI Bill of Rights into a simplified checklist for evaluating new tools before district-wide adoption.

Enforce Anti-Indoctrination Safeguards

- The Task Force should build firewalls against centralized control of values instruction. Require algorithmic transparency in content generation and prohibit behavioral ranking tied to ideological conformity.
- *Example*: Any AI-generated curriculum content must include a visible audit trail showing the sources and training data that shaped its recommendations.

Note: The Task Force should define clear audit criteria, disclosure requirements, and appeal mechanisms that ensure AI tools in schools operate in full view of the public and uphold democratic norms, including protections enshrined in U.S. constitutional and civil rights frameworks.

3. **Equity & Inclusion:** Safeguard the rights and needs of vulnerable groups most at risk of exclusion or harm, grounded in protections under Title VI, Section 504, and IDEA.

Defend DEI and Civil Rights in Curriculum Design

- The Task Force should mandate that AI tools reflect civil rights protections and equity goals. Create accountability processes for systems that exclude or erase race, gender, disability, or class identity in pursuit of "neutrality." These processes should align with established U.S. statutes such as Title VI of the Civil Rights Act, Section 504 of the Rehabilitation Act, and IDEA.
- *Example*: Require AI vendors to publish demographic performance metrics and provide opt-in DEI features that allow local tailoring.

Protect Neurodivergent and Disabled Learners

- The Task Force should require AI systems to accommodate diverse cognitive profiles

and prohibit behavior-based labeling without human oversight. It should develop inclusive design standards and validate them with neurodiverse and disabled communities.
- *Example*: Before launch, an AI learning assistant is evaluated by a coalition of neurodiverse reviewers and redesigned after testing its language simplification tools on students with ADHD and dyslexia.

Note: Equity challenges facing rural and underserved districts, such as limited infrastructure, staffing shortages, and uneven access to AI-ready environments, should be formally reviewed by the Task Force as cross-cutting implementation concerns. These challenges are especially acute in the U.S., where education funding and infrastructure vary dramatically across states and localities.

4. Community Empowerment: Equip learners, families, and educators with the tools to participate and advocate through participatory processes rooted in U.S. school district engagement models.

Strengthen Community AI Literacy

- The Task Force should host family and student-focused workshops on how AI works, its limitations, and how to advocate.
- *Example*: A Saturday community fair includes hands-on AI demos and training sessions where students learn how to identify when a tool is giving biased or incorrect information.

Build Student and Family Feedback Channels

- The Task Force should create formal structures for students and families to give input on AI tools in real time, including opt-out options, satisfaction surveys, and harm reporting pathways.
- *Example*: A district adds an AI feedback portal linked to school dashboards, where students can rate clarity, accuracy, and fairness of AI tutoring responses.

Note: The Task Force should prioritize hands-on community engagement efforts and feedback systems that put families, students, and educators in the driver's seat when it comes to AI integration. These mechanisms should reflect existing U.S. models for school district participation and parent-student advisory structures.

5. Cultural & Cognitive Integrity: Protect cultural values, community identities, and diverse ways of thinking from being overwritten, particularly for historically marginalized communities (such as poor, women, LGBTQ+, indigenous, people of color, neurodivergent, & Appalachian), ensuring the accuracy and relevance of curriculum development, teaching methods, communication, and education policy & standards.

Preserve Local Cultural and Linguistic Identity

- The Task Force should fund culturally responsive curriculum audits. These could prevent algorithmic suppression of regional dialects, minority languages, and community-specific learning norms. These efforts should reflect domestic cultural diversity and be informed — when appropriate — by international guidance such as the UNESCO Recommendation on the Ethics of Artificial Intelligence.
- *Example*: A regional education co-op conducts a bias scan on an AI writing tutor and discovers it flags Appalachian dialect grammar as "incorrect," prompting revision of its correction model.

Protect Cognitive Diversity in System Design

- The Task Force should ensure AI tools account for multiple learning models and problem-solving styles, not just neurotypical or high-speed task performance.
- *Example*: An AI platform's success metrics are redesigned to reward slow, exploratory learning alongside speed and precision, enabling a broader range of students to succeed.

Note: The Task Force should create design review protocols that prevent homogenization of language, thought patterns, and community expression, especially in a diverse and decentralized U.S. education landscape.

Conclusion

The next step is clear; initiate the Task Force. This body must be empowered not only to assess the risks of AI in education but to oversee implementation, prioritize transparency, and establish long-term accountability mechanisms.

Its responsibilities should include:

- Establishing clear oversight processes, including regular audits and public reporting;
- Creating feedback and redress channels for students, families, and educators;
- Coordinating pilot programs that center local voice and preserve learner agency; and
- Adapting standards to account for contextual differences, such as rural or under-resourced settings.

Oversight should not be a static checkpoint but a living framework: one that evolves in response to community input, emerging risks, and new insights. Implementation must remain cautious, adaptive, and community led. Above all, AI systems must support and never undermine the dignity, autonomy, and diversity of the learners they serve.

A well-structured Task Force that is grounded in shared values and public accountability can help shape national and global norms for educational AI. It can advance policies that reflect U.S. values of equity, transparency, and democratic participation while engaging with international benchmarks such as UNESCO's ethical guidance as a global standard for inclusive and accountable AI use in education.

We're not just choosing tools; we're choosing values. Without oversight, AI in classrooms could erode the foundation of public education: trust, equity, and the dignity of learning.

How we use AI in classrooms reflects and reinforces the values we want students to carry forward. We are making governance decisions that will shape the future of trust, inclusion, and accountability in American education. Implementation and oversight will determine whether these technologies strengthen learning or quietly displace it.

Each of the strategic goals outlined earlier offers a pathway toward a more just, inclusive, and resilient educational future. Together, they prioritize transparency in system design, protection of student data, and culturally responsive development practices. They call for investments in public infrastructure and educator training while ensuring that AI supports – rather than replaces – the human relationships that are at the heart of education.

By elevating the voices of students, families, and educators, and by establishing equity-centered governance and accountability systems, these strategies move us toward classrooms where AI amplifies curiosity, not compliance; collaboration, not control. This is how we ensure AI remains a tool for learning not a system that narrows it.

Without safeguards, the risks are clear; untested systems may be deployed faster than they are evaluated, exposing students to bias, surveillance, and inequitable outcomes. When that happens, schools risk becoming sites of experimentation rather than spaces of empowerment. The stakes are too high to ignore. These goals are not merely aspirational; they are essential safeguards that anchor public education in human judgment – not algorithmic control – and preserve it as a human-centered public good.

The United States has the opportunity and responsibility to lead in defining educational AI that aligns with democratic values. While international frameworks such as UNESCO's ethical guidance set a global bar for equity and accountability, U.S. policy must ensure that AI in education strengthens, rather than fragments, the

foundations of public trust, civil rights, and learner autonomy.

Author (In order of contribution)

John Barton, Founder/Executive Director; AI Strategist & Architect
John Barton, Founder & Executive Director of the Spectrum Gaming Project, is an AI strategist and governance architect focused on building ethical systems for underserved markets. With a Master's in Counseling and decades in community education, he has delivered over 10,000 trainings in neurodiversity, education, and innovation. Based in Appalachia, his work has been recognized and adopted by the American Bar Association, the ACLU of West Virginia, Americorps VISTA Leaders, and the WV Community Development Hub.

For more information about the Coalition for Innovation, including how you can get involved, please visit coalitionforinnovation.com.

View the Next Chapter