

# what is the EDSAFE AI SAFE Framework?



**EDSAFE AI**  
ALLIANCE

Founded in 2020, the EDSAFE AI Alliance is a global initiative coordinated by InnovateEDU and powered by a coalition of organizations representing stakeholders across the education sector to provide global leadership for developing a safer, more secure, more equitable, and more trusted AI education ecosystem through a focus on research, policy, and practice.

By joining forces and complementing rather than competing with stakeholders in the space, we can address one of the most pressing educational policy challenges of our time: how to build and develop an ecosystem that reflects the best practices for AI use in education.

With a shared mission to leverage AI to facilitate better, more equitable student outcomes, save time for teachers and other educators, and increase efficiencies for those invested in teaching and learning ecosystems, ED SAFE AI is an uncommon alliance dedicated to furthering safe, accountable, fair and transparent, and efficacious AI use within the K-12 education space through the SAFE framework.

We believe that AI should not be used to replace educators in classrooms or in decision-making processes: instead, AI can be used to reduce educator paperwork and help to synthesize and make suggestions based on learner data. In this fashion, AI can be leveraged to allow teachers more time for direct instruction with their learners, using technology as a “co-pilot” that supports teachers in doing the human, emotional, affective work that is at the heart of the learning process.

Led by a Steering Committee of organizations and their leaders who are already working at the intersection of AI and education, EDSAFE AI aims to bring diverse voices to help educate policymakers and shape an equitable and safe AI education ecosystem.

## STEERING COMMITTEE

**AASA, The School Superintendents Association**

**AI in Education**

**aiEDU**

**American Association of Colleges for Teacher Education (AACTE)**

**American Federation of Teachers (AFT)**

**CAST**

**Code.org**

**Consortium for School Networking (COSN)**

**Council of Great City Schools**

**Data Science 4 Everyone**

**Digital Promise**

**Dxtera Institute**

**EdTrust**

**Education Counsel**

**Federation of American Scientists (FAS)**

**ISTE/ASCD**

**National Center for Learning Disabilities (NCLD)**

**National Digital Inclusion Alliance (NDIA)**

**National Education Association (NEA)**

**National Parents Union**

**Opportunity Labs**

**Project Evident**

**SETDA**

**Software & Information Industry Association (SIIA)**

**The New Teacher Project (TNTP)**

**Transcend**

## THE FRAMEWORK

The work of the EDSAFE AI Alliance centers on the SAFE Benchmarks Framework. The framework creates a policy process and roadmap for the essential issues in creating a SAFE AI ecosystem. The framework was built starting in 2021 and brings together more than 24 global AI safety, trust and market frameworks. Frameworks and benchmarks are essential to innovation as a means of targeted guidance, focusing disparate efforts towards shared language, objectives, and outcomes and ensuring the development of appropriate guidelines and guardrails for use. By working together through the Framework, EDSAFE aims to accomplish two things: achieve equitable outcomes for students and improve working conditions for teachers.

### SAFETY

Safety is a primary consideration for ensuring edtech users can be active in a digital environment that prioritizes protecting their data and privacy while managing cybersecurity risks. At the same time, solution providers have a shared commitment to responsibly building innovative education solutions. Solution providers must be able to continue developing and deploying tools and the evolution of their product roadmaps while responsibly building innovative learning solutions.

### ACCOUNTABILITY

Accountability is a cornerstone for establishing benchmarks that are collaboratively defined by a diverse group of constituents, encompassing subject matter experts, edtech solution providers, educators from Pre-K through higher education, and learners of all ages. This approach ensures the creation of standards that not only foster accountability but also bolster it through the integration of pertinent existing policies and regulatory undergirding. Concurrently, all parties involved are dedicated to the transparent development and implementation of these standards, ensuring they are in line with the evolving educational landscape and effectively address the needs and expectations of all stakeholders.

### FAIR and TRANSPARENT

Achieving ethical, unbiased, and equitable learning opportunities necessitates a conscious and deliberate effort from both solution providers and users to scrutinize the quality of data being utilized, including the acquisition of datasets, the application of their products, and the monitoring for any inadvertent biases. AI products and experiences must be accessible for all individuals and there needs to be strategies for creating standards, guidelines, and/or quality indicators to gauge this accessibility. This endeavor extends to ensuring fairness and transparency, particularly in the procurement of materials produced by AI outputs, underlining the importance of vigilance and intentionality in every aspect of educational technology engagement.

### EFFICACY

Deliberations are essential regarding the effectiveness of applying AI in specific scenarios, coupled with the commitment by edtech solution providers to integrate comprehensive and transparent evaluation tools within educational technologies. Such integration aims to precisely measure advancements and provide educators and learners with insightful feedback on usage and progress. Efficacy is understood to be deliberately tied to an equity in student experiences as well as outcomes. This approach underscores the significance of carefully considering the utility of AI and the importance of inclusivity and clarity in the mechanisms used to gauge and communicate educational outcomes.

## EDSAFE AI SAFE Framework

**S**

**SAFETY**  
Security, Privacy, Do Not Harm

**A**

**ACCOUNTABILITY**  
Defining Stakeholder Responsibilities

**F**

**FAIRNESS**  
Equity, Ethics, and Mitigating Bias

**E**

**EFFICACY**  
Improved Learning Outcomes

