AI-Powered Innovations in Mathematics Teaching & Learning: Initial Findings

Sierra Noakes, Alison Shell, Parker Van Nostrand, Alexis M. Murillo, Pati Ruiz, and Babe Liberman

September 2024

Digital Promise®
Accelerating Innovation in Education

Acknowledgments

We thank the students, educators, instructional coaches, technical experts, math experts, and equity experts who supported the review of the request for information submissions. We would also like to thank many of our colleagues at Digital Promise for their expertise and contributions to this project, including Shayla Cornick, Nicola Hodkowski, Lauren McMahon, Julie Neisler, and Megan Pattenhouse. This research was supported by the Bill & Melinda Gates Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funders.

Contact Information

Email: sierra@digitalpromise.org

Digital Promise
Washington, DC:
1001 Connecticut Avenue NW, Suite 935
Washington, DC 20036
https://digitalpromise.org/

Suggested Citation

Noakes, S., Shell, A., Van Nostrand, P., Murillo, A., Ruiz, P., and Liberman, B. (2024, September). *AI-Powered Innovations in Mathematics Teaching & Learning: Initial Findings.* Digital Promise. https://doi.org/10.51388/20.500.12265/229

Executive Summary

This report discusses initial findings based on the analysis of responses to a request for information (RFI) led by the Bill & Melinda Gates Foundation and Digital Promise. The RFI was launched in April 2024 and closed in May 2024, receiving nearly 200 responses that described a variety of innovative approaches to leveraging artificial intelligence (AI) for mathematics teaching and learning. The analysis focused on three core questions to drive broader learning across the field, to support both education leaders with decisions about AI as well as providers in learning more about market saturation and strategies to mitigate risks.

As AI becomes increasingly prevalent in education, three key questions often drive conversations around this emerging technology:

1. What does AI in education look like today?

Responses generally fell into categories around teacher support (39.3%)—such as automatic feedback and grading, lesson generation, access to new types of data, and rapid data analysis for real-time instructional recommendations—and student support (52.0%)—including adaptive and personalized learning materials and intelligent tutoring systems. A handful of responses considered the unique opportunity AI brings to creating inclusive learning environments and to scaling evidence-based strategies.

2. What are the risks to leveraging AI in education, and how might those risks be mitigated?

Three primary concerns dominated the risks described in responses. Algorithmic bias was the most common risk articulated by respondents (31.9%), followed by data privacy (27.5%) and technical reliability (27.5%). To mitigate these risks, many described plans to leverage testing and monitoring—including training models with diverse data sets, utilizing human-in-the-loop workflows and co-design, and using technical techniques, like the self-consistency and knowledge graph approaches and building precise prompts or unique models for specific output generation. Additionally, across identified risks, respondents emphasized the importance of professional learning and guidance for ethical and appropriate use.

3. What role should AI play in education?

Leveraging AI-enabled technologies to drive equitable outcomes in education calls for education leaders to hold collaborative and inclusive conversations with their communities to determine how to position AI as a tool toward achieving their shared vision. Driven by district priorities and subject matter experts, the field also needs to establish benchmarks to help recognize and evaluate the ethical design of AI for teaching and learning. AI developers should embrace and center the expertise of teachers and students by prioritizing co-design throughout a tool's R&D processes. The future of AI also calls for innovations in research; there need to be rapid and responsive, yet rigorous and reliable, research methods to meet the rate of change established by AI so the field can understand what works and what doesn't.

Table of Contents

Introduction

What role can artificial intelligence (AI) play in transforming mathematics teaching and learning? Mathematics education holds immense potential to empower students, equipping them with essential skills such as critical thinking, problem solving, and a deeper understanding of the world. By integrating AI, we can further amplify this potential, enhancing mathematics teaching and learning, as highlighted in a recent position paper from the National Council of Teachers of Mathematics (NCTM, 2024). For instance, NCTM expects AI to help by personalizing problems, evaluating learner understanding, supporting student engagement and learning, and reducing workload for teachers. However, the organization cautions that teachers are essential to students' learning and that, in the era of AI, having deep knowledge and expertise becomes even more vital in order to promote critical and creative thinking in mathematics.

AI has the potential to personalize instruction to meet the diverse interests, strengths, and needs of individual learners to ensure that education delivers impactful learning experiences. AI can also empower educators with real-time, data-driven recommendations and grading support, and can offer easy access to evidence-based instructional strategies. The expectations around AI in mathematics and education more broadly, including excitement and concerns, from end users and the field are matched by the ambitions of many developers. As a means of capturing innovative solutions in AI, Digital Promise and the Bill & Melinda Gates Foundation issued a request for information (RFI) to provoke ideas ranging from early-stage concepts to on-market tools to best understand the rapidly evolving landscape, with a focus on the potential for AI to promote equitable mathematics teaching and learning. The Digital Promise team analyzed data from 178 RFI responses. We acknowledge that these responses came from a sample of innovators in networks connected to Digital Promise and the foundation. While they are not necessarily representative of *all* AI innovations, they do represent a wide range of ideas and add up to a valuable first step in understanding the landscape. The goals of the analysis—and this report—are to provide the field with a broader understanding of what to expect from AI, to identify emerging opportunities for AI in education, and to learn more about the types of technologies being leveraged and the strategies being deployed to mitigate the myriad risks associated with AI.

Background

AI systems are typically designed to serve the needs of researchers rather than the local communities they affect. However, community data and knowledge are highly specific to local contexts, which often makes them better suited for addressing community needs. It can be difficult for researchers who are not a part of the community to truly understand the meaning of their data (Hsu et al., 2022). One of the goals of this RFI is to understand where the market is projected to go to identify where, without philanthropic intervention and amplifications of opportunities and successes, these failures might repeat themselves. This

nascent stage of AI in education offers a pivotal opportunity for funders, solution developers, and other education leaders to ensure students and teachers from diverse backgrounds drive solution development so that schools and classrooms can address the full spectrum of learner strengths and needs. Mechanisms like this RFI can help philanthropic organizations understand where organizations who are breaking the status quo need support in their uphill battles to develop edtech solutions with novel approaches to research and development, including co-design with educators and community members.

The RFI was launched as part of a commitment by the foundation to advance research- and evidence-based approaches that leverage AI to enhance mathematics teaching and learning (See Appendix A for the full RFI). [1] We hoped that this RFI would allow the foundation and our partners to gain deeper insights into the uses and aspirations of AI in education. The foundation's K-12 Education team leveraged the findings of this RFI to help guide, develop, and refine their approach to supporting AI innovations. Several of the ideas from this RFI have led to grants that focus on research and/or solutions to meet the needs of students who are Black, Latino, and/or from low-income backgrounds and their teachers. Learnings from these projects are expected to produce evidence for the field about critical features of products and services that lead to improved outcomes for students and teachers.

RFI responses demonstrated a tremendous amount of focus around using AI to personalize learning, feedback, and instruction, as well as inspiring ideas around analyzing student work and providing real-time instructional recommendations. We were also excited to see many thoughtful responses where innovations were being intentionally designed for *and with* priority populations. These findings reinforced the foundation's priorities to invest in tools that meet real needs through co-design to ensure solutions move toward transformative outcomes.

[1] This paper was written through funding from the Bill & Melinda Gates Foundation. Views expressed here do not necessarily reflect positions or policies of the foundation.

Responses Overview

The RFI opened in April 2024 and closed in May 2024 after receiving 178 responses focused on innovative approaches to transform mathematics teaching and learning using AI-powered technologies (See Appendix A for RFI). One hundred sixty-nine submissions were designed to support PK-12 mathematics education in the United States. Beyond the United States, a subset of submissions aimed to meet the needs of mathematics education across Asia, Africa, Australia, Europe, North America, and South America. Responses intended to impact a variety of audiences across roles and learner levels (See Table 1).

Table 1. RFI response types by categories (N=178).

Category	Type	Percentage
Beneficiaries	Teachers	33.8%
	Students	32.2%
	The Field (a greater understanding of AI in shaping mathematics teaching and learning)	18.6%
	Instructional Coaches	15.4%
Age Group	Secondary 6-12	46.4%
	Primary PK-5	34.7%
	Early Childhood	10.5%
	Postsecondary	8.5%
Maturity	Early Concept Ideas or Research Projects	37.1%
	Initial Prototyping and Field Testing	35.4%
	On-Market with Feature Development	27.5%
Subject Focus	Mathematics Education	84.8%
	Subject Agnostic	11.2%
	Not Specified	3.9%
Type of Organization	For-profit edtech company	45.0%
	Nonprofit edtech company	17.8%
	Research institution/ University	17.8%
	Non-edtech nonprofit organization	11.8%

	Public school district, charter school, private school	5.4%
	Non-edtech for-profit organization	2.4%

Methods

The RFI was conceived to help the foundation and the broader field to learn about innovative approaches that leverage AI to transform mathematics teaching and learning. Digital Promise and the foundation designed the components of the RFI to elicit information about the organization that submitted, the geography and beneficiaries of intended impact, the approach rationale, the innovation's approach to co-design, the purpose for using AI, as well as discuss the risks and mitigation strategies.

On April 15, 2024, Digital Promise and the foundation announced the launch of the RFI at the ASU+GSV Summit. The RFI opened as a 21-question Qualtrics form and closed May 20. In the intervening period, three "office hours" sessions were hosted where prospective respondents could learn more information about the RFI. Digital Promise and the foundation promoted the RFI to edtech providers, district leaders, and education researchers through conference presentations, newsletters, targeted emails, and social media.

After the submission window closed, the Digital Promise team reviewed, analyzed, and coded all 178 responses. Descriptive statistics were run to analyze submission type, including approach maturity, geographic reach, type of beneficiaries, age group, subject focus, and types of technologies used. Thematic coding was used to identify themes across project topics, the purpose of leveraging AI, potential risks, and how respondents intended to mitigate risks.

Submissions that the Digital Promise team deemed research based and aligned to field priorities were reviewed by a panel of 26 people, including current students (11), teachers and instructional coaches (8), professional learning experts (2), AI experts (4), and an equity expert (1), to offer feedback on the usefulness, potential for equitable impact, and feasibility of the responses. Following the review process, the Digital Promise team held virtual debriefing interview sessions with external reviewers to gauge their overall impressions on the proposals they read and how well these innovations did or did not meet needs or expectations, according to their expertise. These findings along with the responses were shared with the foundation for further review and potential follow-up.

Findings

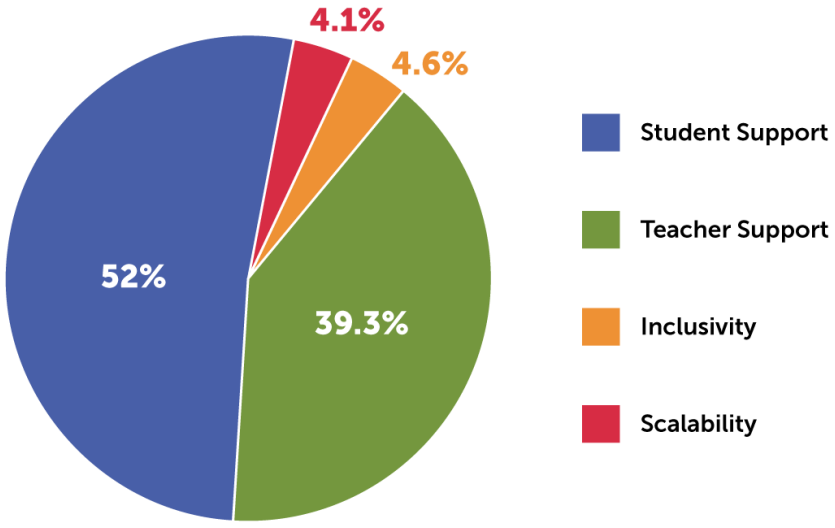
Why Use AI?

In line with the findings from a recent AI Literacy position piece (Mills et. al., 2024), RFI responses generally leveraged AI for teacher support (39.3%)—such as automatic feedback

and grading, lesson generation, access to new types of data, and rapid data analysis for real-time instructional recommendations—and student support (52.0%)—including adaptive and personalized learning materials and intelligent tutoring systems (See Figure 1). A handful of responses considered the unique opportunity AI brings to create inclusive learning environments and scale evidence-based strategies.

On April 15, 2024, Digital Promise and the Gates Foundation announced the launch of the RFI at the ASU+GSV Summit. The RFI opened as a 21-question Qualtrics form and closed May 20. In the intervening period, three "office hours" sessions were hosted where prospective respondents could learn more information about the RFI. Digital Promise and the Gates Foundation promoted the RFI to edtech providers, district leaders, and education researchers through conference presentations, newsletters, targeted emails, and social media.

Figure 1. Percentage of responses for the value of AI by theme (N=283 tagged across 178 submissions).

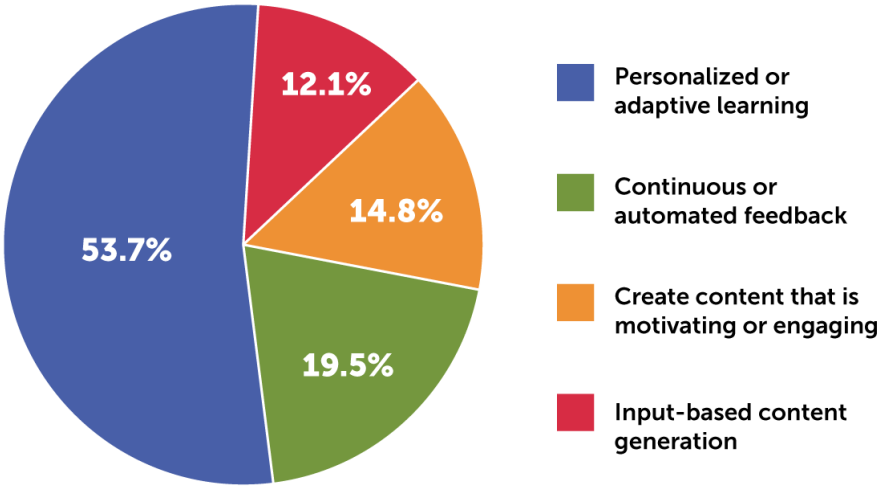


Student Support

Over half of all descriptions for the use of AI considered the technology as essential for developing adaptive, relevant, and personalized content for students. Respondents described the use of AI to:

- Create personalized, culturally responsive, and/or adaptive learning opportunities.
- Generate content in response to inputs, such as individual students' interests.
- Develop content that supports students' motivation and engagement.
- Offer continuous and/or automated feedback or support for students (See Figure 2).

Figure 2. Data-driven content: Percentage of responses for value of AI by sub-topic (N=149 tagged responses for purpose of using AI for data-driven content).

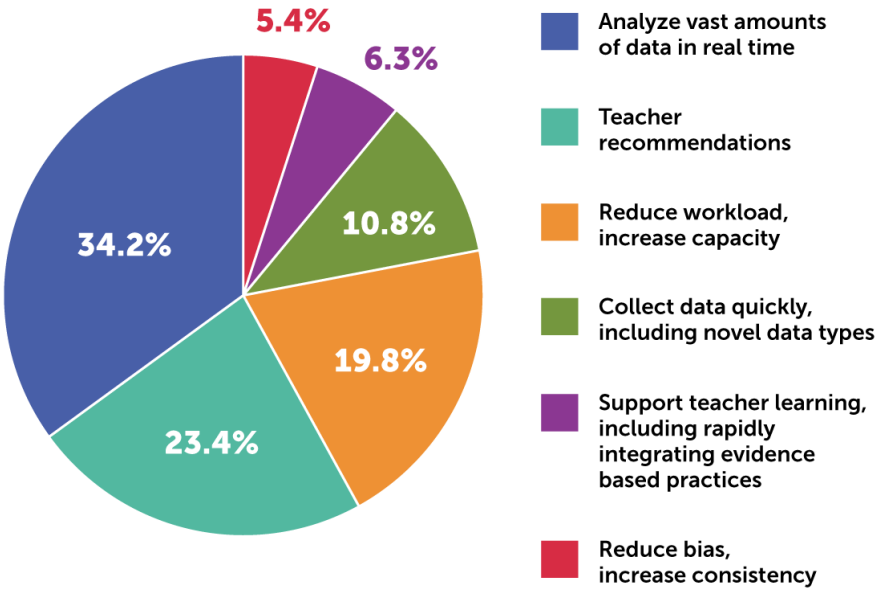


Teacher Support

Nearly 40 percent of respondents named AI's ability to support teachers as a key aspect of their innovative approach. These responses considered the ways AI was uniquely positioned to:

- Reduce teacher workload and increase their capacity to build relationships with students and personalize instruction.
- Reduce teacher bias and increase consistency in grading and assessment.
- Provide teachers with real-time recommendations for instruction.
- Analyze vast amounts of data in real time.
- Collect data, including new types of data like images, handwriting, and discussions.
- Support teachers' learning and efficacy, including rapidly integrating evidence-based practices (See Figure 3).

Figure 3. Support teachers: Percentage of responses for value of AI by sub-topic (N=111 tagged responses for purpose of using AI for supporting teachers).



Inclusivity

While fewer than five percent of submissions described the use of AI as opening new opportunities for inclusivity, those that did identified valuable tactics for the education field to consider. A few submissions described the ways AI could be leveraged to lower costs and thus increase access to resources and supports. Others described the ways the translation capabilities of AI could enable learners to engage with content in their native language, or even develop connections with others through the use of their primary language. Some described using AI to build community and relationships. One submission aims to explore the use of AI for students with disabilities to create accessible learning spaces. Finally, a small number of submissions considered the unique ability AI brings to both collect data on students' social-emotional learning and, through the availability of this data, be able to focus efforts on these outcomes, such as reduced anxiety for learners.

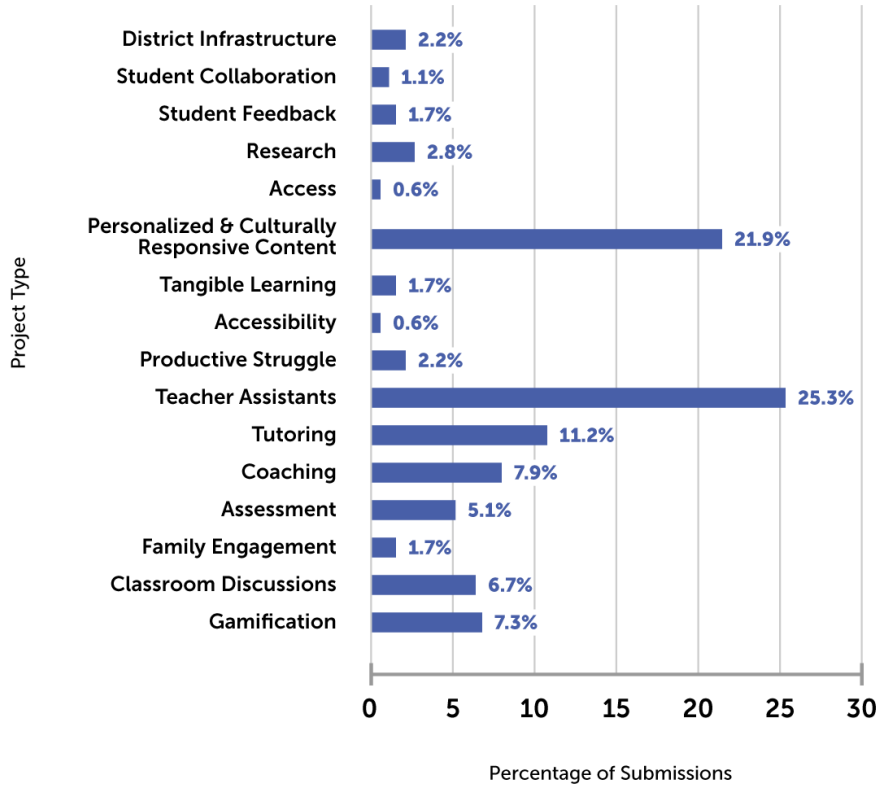
Scalability

A small number of submissions named the use of AI as critical to increasing scalability of evidence-based strategies. For example, several submissions described the value of using AI as allowing effective mathematics practices to occur at scale through amplifying best practices to teachers, coaches, and/or tutors. In other cases, respondents described the use of AI as enabling the team to increase ease of use to accessing content in a toolkit.

Project Topics

Through an analysis of the submissions, two topics dominated as the focus: 1) Using AI to personalize learning, especially with a focus on making content culturally responsive (22%) and 2) providing teachers with assistance (25%) to reduce workload, provide real-time data to improve instruction, and/or responsive content generation to support with the development of materials such as lesson plans. A few submissions specifically named the use of AI to support learners through productive struggle (2.2%) or with tangible learning (1.7%). A handful of submissions described research projects (2.8%), considering ways to conduct research or rapid research and development (R&D) with the use of AI or research on the impact or effectiveness of AI itself in education (See Figure 4). This analysis of the responses to the request for information is not a landscape scan of existing products that use AI in education – with only 27.5 percent of responses discussing solutions currently on market. It should be noted that these findings are in part aspirational, and that many of the concepts highlighted may not be feasible as imagined. However, the findings here are indicative of market saturation and the direction of future development.

Figure 4. Percentage of submissions by project type (N=178 submissions).



Several unique innovations stood out from the more common project topics (See Figure 4). A handful of submissions aimed to use AI for student feedback (1.7%) and to create space for students to collaborate with each other, their teacher, or peers outside of their classroom (1.1%). Only one submission considered access to devices that would empower students to

become engineers technically able to build their own LLMs. Similarly, only one submission focused exclusively on accessibility and using AI in a novel way to support students with disabilities to engage in learning in a transformatively inclusive way. A few submissions also described AI as a unique opportunity to improve family engagement, including offering educators research-based templates to communicate with families and/or translation services to communicate across native languages. Four submissions also sought to prioritize district infrastructure and policy development to support sustainable decision making approaches around the procurement and effective use of AI.

Types of Technologies

Respondents were asked to select the technology(ies) they are using or anticipate using in the described project (See Figure 5). Nearly a quarter (23%) anticipate using Large Language Models (LLMs), followed by Natural Language Processing (NLP) (17%), predictive AI (14%), chatbots (13%), and Augmented Speech Recognition (ASR) (10%). Only five submissions indicated that they anticipated using robotics for the project. Over 40 respondents also indicated using technologies not included on the list, a description of which is provided in *Figure 6*.

Figure 5. Percentage of technologies respondents anticipate using for their described projects (N=666 selected responses).

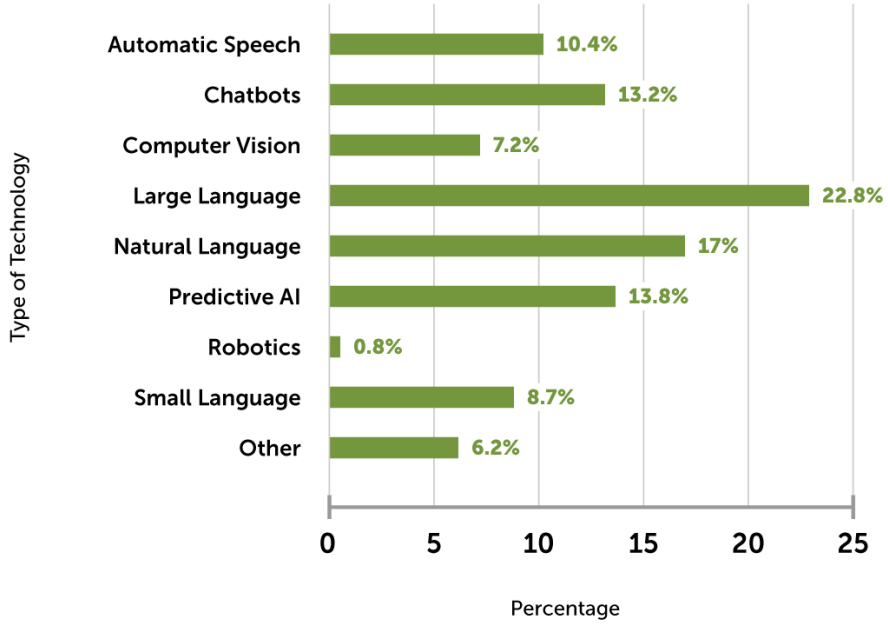
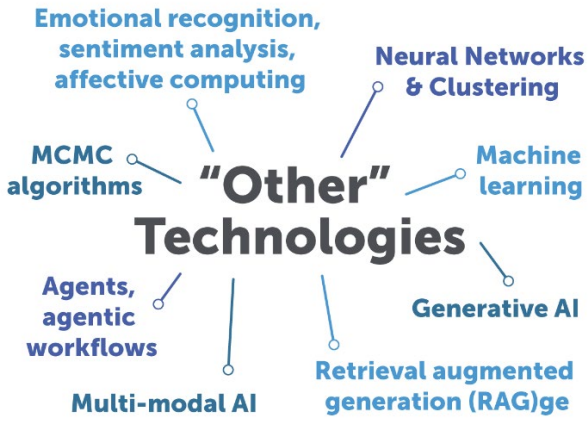


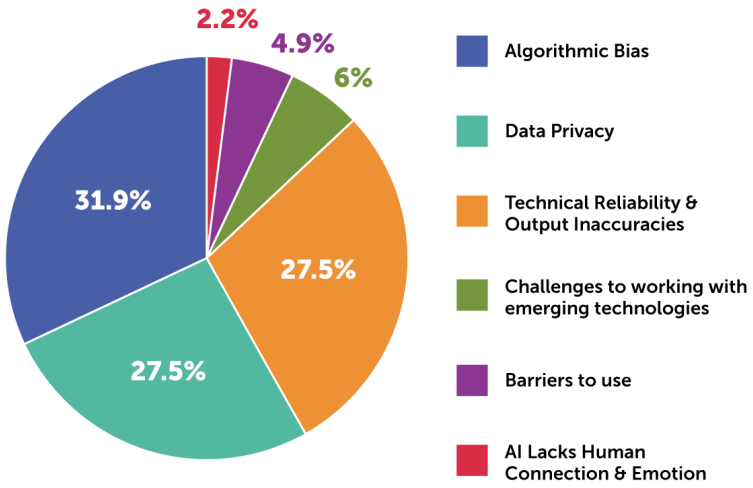
Figure 6. "Other" technologies that respondents anticipate using for their proposed projects.

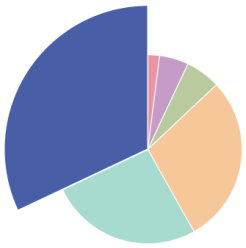


Risks and Mitigation Strategies

Respondents were asked to consider the risks of leveraging AI for the project and how the respondent planned to mitigate those risks. A large number of themes emerged from this qualitative analysis, where risks were thematically coded, followed by a similar process for coding mitigation strategies for each unique risk described (See Figure 7). Respondents most commonly identified algorithmic bias (31.9%) as a risk, followed by data privacy (27.5%) and technical reliability (27.5%). A few respondents also shared risks that come from working with emerging technologies (6.0%), barriers to use (4.9%), and the reality that AI lacks human connection and emotion (2.2%).

Figure 7. Percentage of risks described in submission responses (N=185 tagged responses).





31.9% Algorithmic Bias

Mitigation Strategies:

Testing and Monitoring, Human Involvement, Technical Techniques, Guidance and Professional Learning

Algorithmic Bias

Nearly one-third of respondents noted a risk of algorithmic bias (31.9%) when considering the use of AI for their described projects. A variety of strategies were described to mitigate the inevitability of bias, including testing and monitoring techniques (61.9%), human involvement (22.7%), technical strategies (8.2%), and supporting users through guidance and professional learning (7.2%).

Mitigation Strategy: Testing and Monitoring

Among the testing and monitoring strategies described, respondents discussed the need to conduct ongoing monitoring and audits of outputs, including output testing and validation, to enhance reliability and accuracy. Some even called out the need for this testing to be mixed methods to better understand user experience. Others described the importance of training and testing with diverse data sets and continually vetting and refining the training data based on output testing results to mitigate bias.

Mitigation Strategy: Human Involvement

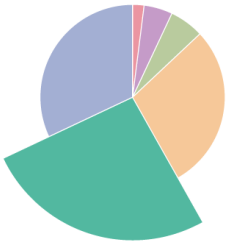
Many respondents described strategies that involve humans to mitigate algorithmic bias. The majority of these strategies involved human-in-the-loop testing, where experts and users are involved in evaluating the quality of outputs. Others described the necessity of co-designing studies with AI-enabled technologies to center users' needs and priorities in the development of an ideal output or outcome, thus ensuring alignment in goals across developers and those who use or are impacted by the technology.

Mitigation Strategy: Technical Techniques

Some respondents described technical techniques for mitigating algorithmic bias. These approaches included using anti-bias mitigation strategies in development and iteration of the technology. Some respondents described using a closed model to control training data. Others described building precise prompts and algorithms or unique models for specific questions rather than using a broad large language model (LLM) for the entire user experience. Some also identified the knowledge graph approach as a mitigation strategy toward more fair and equitable algorithmic outputs and recommendations.

Mitigation Strategy: Guidance and Professional Learning

Several respondents also emphasized the importance of providing guidance and professional learning to support the critical consumption of outputs and ethical use of the technologies. Some noted the importance of transparency around the processes that generate outputs, including recommendations generated by the model, to ensure users clearly understand the technology and can use human judgment to drive final decisions about how to utilize synthetically generated content. Others noted the importance of creating clear policies around AI being used as a supplementary tool to support human teachers in their work to make explicit the necessity of teachers in students' learning. Some responses also advocated for clear guidelines around best practices for AI in education.



27.5% Data Privacy

Mitigation Strategies:

Data Governance, Technical Strategies, Human Involvement and Guidance

Data Privacy

Over one-quarter of respondents named data privacy as a risk for using AI (27.5%). A variety of strategies were described to mitigate the inevitability of bias, including data governance and monitoring (68.9%), technical strategies (18.0%), and involving humans and providing users with guidance (13.1%).

Mitigation Strategy: Data Governance

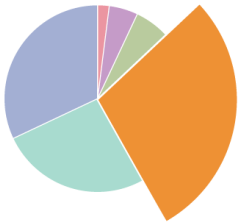
More than two-thirds of those who indicated data privacy as a risk to using AI stated data governance and monitoring as their mitigation strategy. Most commonly, respondents shared that they planned to use strict data governance or protocols (31.1%) to mitigate the risk of data privacy breaches. Others described a commitment to complying with relevant laws and regulations, including FERPA. Many also described a plan for ongoing monitoring and audits to continuously check for data breaches.

Mitigation Strategy: Technical Strategies

Respondents described several technical strategies to mitigate the risk of data privacy breaches. Some indicated that they would use a closed model for their solution. Others described anonymizing data to minimize the risk of a data breach that would expose personally identifiable data. Similarly, several respondents intend to limit access to data to reduce the risk. Many also aim to use data encryption to protect personally identifiable data.

Mitigation Strategy: Human Involvement and Guidance

Some respondents described relying on humans and professional learning to mitigate the risk of data privacy breaches. Several laid out a plan to involve users in defining privacy policies, including requiring explicit and readable consent. Others emphasized the significance of professional learning for safe data use. Similarly, some underscored the need for teachers and students to be involved in setting guidelines to ensure alignment to their intended usage and to confirm the guidance is clearly articulated for safe use.



27.5% Technical Reliability

Mitigation Strategies:

Human Involvement, Testing and Monitoring, Technical Techniques, Professional Learning and Intentional Use

Technical Reliability

Over 25 percent of respondents considered the risk of technical reliability when using AI, including inaccurate outputs and latency problems (27.5%). Most commonly, respondents that noted this risk described the necessity for human involvement (36.9%). Many also described strategies that leverage testing and monitoring (26.2%), technical techniques (18.5%), and professional learning and intentional use of AI-enabled technologies (18.5%) to mitigate concerns around technical reliability.

Mitigation Strategy: Human Involvement

Among the respondents who described technical reliability as a risk to leveraging AI, over one-third identified human involvement as a key strategy to mitigate this risk. Many respondents described a plan for human evaluation processes or human-in-the-loop workflows. Some described an approach that centers co-design with users and involves users directly in testing the accuracy of outputs. Through this process, some emphasized the opportunity to turn inaccurate output moments as knowledge sharing and collaboration opportunities to support students and teachers in being critical consumers of synthetically generated content. Other respondents considered the importance of creating straightforward and easy-to-use teacher reporting mechanisms to make sure practitioners could quickly report inaccuracies back to the developers.

Mitigation Strategy: Testing and Monitoring

Many respondents described a plan to use testing and monitoring strategies to mitigate the risk of technical reliability issues. Several stated that they would use rigorous testing of outputs, in some cases defining target outputs and holding AI-generated outputs against these ideal outputs to check for correctness. Many also claimed to use ongoing monitoring and auditing. Some also plan to leverage guardrails, safeguards, or content moderation systems to mitigate the risk of technical reliability issues.

Mitigation Strategy: Technical Techniques

Respondents described several technical techniques to mitigate the risk of technical reliability issues. Some described inaccurate output mitigation techniques, namely a self-consistency technique that combines chain-of-thought prompting with pre-trained large language models, which researchers have seen lead to promising results (Wang et. al., 2022). Several

respondents stated that they plan to continually refine algorithms based on identified errors with outputs, while others plan to continually refine training data to result in increased output accuracy. A few respondents also described techniques with weighted training content or dual agent design to monitor for errors.

Mitigation Strategy: Professional Learning and Intentional Use

Some respondents advocated for professional learning and clear, intentional use of AI as strategies to mitigate technical reliability concerns. Several stated the importance of making intentional decisions about where AI should be used, and where it should not be used. A few respondents also articulated the importance of AI literacy and user awareness to ensure teachers and students could identify inaccurate outputs and maintain skepticism about synthetically generated content. Some also described the importance of positioning AI-enabled technologies as a tool for teachers and ensuring it never oversteps its role as a supplementary tool for learning.

Additional Risks

Some respondents also articulated risks from working with emerging technologies (6.0%), barriers to use (4.9%), and AI lacking human connection and emotion (2.2%).

Working with Emerging Technologies

Some respondents described challenges that come with working with new and emerging technologies. Namely, these respondents described the difficulty of working on a large number of hypotheses, making outputs and outcomes risky. Others described challenges with technical decisions, such as rule complexity. Some also identified cost as a challenge to working with new and constantly evolving technologies. These respondents suggested a variety of strategies to mitigate the risks, including a call for philanthropists to invest in building out datasets that can be leveraged by a variety of developers to reduce costs. Similarly, creating reusable elements of models that can be shared across developers would not only support lowering costs, but also inspiring improvements and new approaches as developers could build on each other's work rather than requiring each developer to start from scratch.

Barriers to Use

Several respondents named challenges that serve as barriers to use, including general fear around AI, access to the newest and best models, and ineffective use of AI, including over-reliance on AI's decisions or recommendations. Most of these respondents suggested a variety of training and testing strategies to mitigate these risks, such as A/B testing on large data sets to validate patterns and outputs, refining AI algorithms to ensure alignment to educator objectives, and guardrails and moderation processes. Others underscored the importance of professional learning to ensure intentional and ethical use of AI. Similarly, several respondents emphasized the need to engage with established partners in designing

and implementing the AI-enabled technologies to ensure shared goals and understanding, as well as to establish transparency and clear communication channels with users.

AI Lacks Human Connection and Emotion

A handful of respondents articulated challenges to using AI given its lack of emotion and concerns around it reducing connection and relationship development between teachers and students and among students. These respondents recommended clearly establishing AI as a supplementary tool and incorporating positive social and emotional learning supports.

Additional Strategies

Respondents also shared a variety of strategies that were not necessarily tied to a specific risk to ensure effective and appropriate use of AI-enabled technologies:

- Collaborate with AI experts.
- Allow educators to adjust and have final say on all content.
- Continuously engage with users throughout implementation to iteratively improve the tool.
- Ensure a diverse developer team.
- Only use well-vetted models.
- Prioritize teacher and administrator professional learning.
- Collaborate with district and school leadership, teachers, and the community to develop fair and transparent policies around ethical use.

Discussion

With the incredibly fast emergence of AI to market, vast opportunities to create solutions arise. Yet it becomes increasingly critical to consider who we are developing solutions for and what problems actually need to be solved. In learning technologies, students, parents, teachers, and education leaders have myriad wants and needs that AI has the opportunity to support. The 178 responses to this RFI represent a broad swath of innovative approaches to challenges that exist in mathematics teaching and learning, including increasing differentiation and personalized content, data-informed instructional recommendations, and access to novel data for professional learning. These findings represent the hopes and aspirations of innovators, and we were thrilled to see where these projects indicate the field is heading.

What Does AI in Education Look Like?

The personalization of learning through AI-embedded tech—representing a major proportion of the responses to this RFI—indicates a current and emerging trend on the use of AI to

support individual needs.[2] Individuals benefiting from personalization can include learners, educators, and even administrators making decisions about how to best support those in their schools. Personalization in the RFI responses took a variety of forms, from structuring learning to respond to perceived student skills gaps, to crafting content that aligns with student interest and background, and providing coaching for educators through analysis of their instruction and classroom discussions. Over half of respondents focused on student supports, within which personalization was a major theme (22% of all responses).

Several studies concur with this common trend, also finding an increased interest in adaptive or personalized learning with advances in technology (e.g., Maghsudi et al., 2021; Tetzlaff et al., 2021). Discussions from Stanford's AI + Education Summit also indicate that AI will play an important role in personalized support for teachers at scale and improving assessment quality via tools that help identify student strengths and competencies and make data-based predictions about learners and institutions (Mouta et al., 2023). Moreover, the World Economic Forum anticipates that AI in education will support educators by streamlining administrative tasks, providing them more time to engage with students, support assessment and feedback, and personalize learning. Further, research with teachers on their use of AI confirms the value teachers see in the role of AI in edtech to support student learning, including in differentiating and translating materials, evaluating student needs to support instructional decisions, and providing feedback on student work (). Teachers most often report use of AI for generating course content and creating new materials to support personalization of content, and to engage students, like through educational games. Additionally, they use it to meet a wide range of other needs, like to diagnose student needs and determine data-driven instructional decisions, providing feedback, and communicating with caregivers (e.g., Dilberti et al., 2024; Hamilton, 2024; Krall et al., 2024; Zafari et al, 2022). Krall et al., 2024). Teachers most often report use of AI for generating course content and creating new materials to support personalization of content, and to engage students, like through educational games. Additionally, they use it to meet a wide range of other needs, like to diagnose student needs and determine data-driven instructional decisions, providing feedback, and communicating with caregivers (e.g., Dilberti et al., 2024; Hamilton, 2024; Krall et al., 2024; Zafari et al, 2022).

By relieving some of the weight on overburdened instructors and instructional coaches, leveraging AI would also enable educators to receive more professional learning and focus on tending to specific needs of individual students and groups of learners. This is evident in the RFI responses that describe approaches that use AI-enabled tools to foster student collaboration with peers and instructors. Moreover, student reviewers of the RFI responses also noted a desire to see more explicit use of AI to build community and relationships with others in their school community and beyond. This is in line with research suggesting the power of strong social connections and interactions in not only the engagement of students

[2] It is important to note that this analysis came from the responses to an RFI that focused on using AI to improve math teaching and learning; this is not a field landscape analysis. See the full RFI in Appendix A.

(Jansen et al, 2022) but even in the effectiveness of education technology, above many other predictors (Ran et al, 2021). The student supports mentioned across the RFI responses mostly align with the literature on AI applications in education, including improving metacognitive awareness, supporting students with learning disabilities, promoting STEM and writing skills, and supporting virtual and collaborative learning (Mouta et al., 2023).

Students are generally hopeful about the impact that generative AI might have at school and work, making learning and working more efficient and faster, and about the accessibility of vast amounts of information. One of the college student reviewers for this RFI, Anthony Guerrero Zotea, shared that AI can be helpful in organizing his thoughts and support proofreading. He said that leveraging these tools "helps reduce stress and improves my writing i[by providing me with] the ability to take risks at helping to create my own style of writing." In line with what this project's student reviewers shared about their uses of AI, researchers are finding that students are most likely to use generative AI (genAI) to assist with brainstorming and to find new information (Hopelab, 2024). Some students also use AI to support them emotionally; they feel less alone, or experience less pressure when asking questions when interacting with AI-enabled tools. Students overall report being excited about using AI to help generate ideas, support creativity, and advance science and medicine.

Concerns and Risks with AI in Education

While there are tremendous opportunities for AI to transform education, respondents shared some deeply held concerns about the risks. The RFI explicitly asked respondents to reflect on areas of potential risk, and intended strategies to be used to mitigate those risks. A comparison with the literature shows overlap in key areas of expected risks and methods to combat these potential harms.

Analysis of the RFI responses revealed that over a quarter of submissions highlight student data privacy as a major concern. These responses are consistent with global findings, such as a 2023 survey of over 17,000 parents, educators, and leaders globally to which 71 percent of respondents expressed concerns about potential risks of AI tools, including those regarding student data and privacy (WEF, 2024). Additional literature supports these concerns, including that teachers have fears about who can access their data, as well as of being noncompliant with policies by using specific AI technologies (Krall et al., 2024). Teachers, developers, and professional learning (PL) coaches show shared concerns about equity, access, privacy, and data privacy and ownership. In addition, professional learning organization leaders have concerns about their intellectual property rights, while district leaders express concerns about affording tool subscriptions and data privacy rights. There is an overall concern about professional learning needs with adopting and implementing AI tools, including limited resources and cost, as well as awareness and concerns around policy (Krall et al., 2024).

While responses to this RFI primarily center on AI supporting individual educators to accomplish tasks that are already on their plates – including lesson planning, assessing student work, or identifying promising interventions – a common concern about AI in education, which did not explicitly come through in RFI responses, is that it may replace

human workers, including educators and district or school staff (WEF, 2024). And while AI tutoring, the third most common response topic, may seem like a threat to human jobs, many of the responses highlight its potential to support students where human teachers, tutors, and counselors are unavailable. Rather than invariably replacing educators, AI has the potential to enhance student learning by transforming, not eliminating, the role of educators. Intelligent tutoring systems may also be able to support students at a lower cost than hiring additional staff members; however, many RFI responses note concerns about students' social and emotional needs, which cannot be addressed by AI tutors.

The distrust of AI held by many educators, caregivers, and students, along with concerns about losing the human element of education, could prove a challenge to embracing the possibilities of leveraging AI to respond to student needs. However, there are thoughtful and intentional strategies emerging to mitigate these risks, as described in the Risks and Mitigation Strategies section of this report. At the core, transparency, AI literacy, and collaboration with end users from concept and throughout implementation are key to building a foundation of trust when integrating new uses of AI into educational settings.

Among young adults who have not used AI, about one-third say they haven't used it because they don't think it will be helpful, while other reasons for avoiding this technology include data privacy concerns, lack of awareness of the tools, association with cheating, and concerns around bias (Green, 2024). Notably, one survey found that Black and Latino students are more likely than students of other races to use AI for most activities (Hopelab, 2024). But RFI college student reviewer, Amara Santos, shared that she personally uses it more for appealing to others than feeling as if it's truly supporting her own learning: "I know a few Black and Brown youth who use AI, but I don't think many of us actively enjoy using it. A lot of times, it's because we're in a pinch. Maybe we should use Grammarly so our professors don't think we just don't care or so our co-workers don't think we're being too aggressive. Maybe we use Chat GPT because we just became so busy with school and work, we didn't have time for the assignment."

There is a related need for more effort from the field in the area of inclusivity, which was described by less than five percent of respondents in their answers to the question on why the use of AI would drive greater impact. As classrooms become more dependent on technology, the question of who can afford devices with sufficient power to run AI-powered technologies and to fully access this technology looms. As such, edtech developers must counteract these potential inequities by making sure that any edtech available to learners prioritizes supporting learners with differing needs (OET, 2024). This includes, among others, issues of accessibility, cultural responsiveness, and primary language. Despite nearly every submission mentioning equity, rarely were submissions intentional and thoughtful about leveraging meaningful co-design processes to center the voices and perspectives of diverse students and educators to build the tools. Utilizing AI-enabled technologies to drive equitable impact and outcomes in education will call for funders and product developers to center co-design in the R&D process.

What Role Should AI Play in Education?

The vision for AI's role in education feels limitless at this moment, as we are witnessing leaps in technical capabilities on a quarterly basis – if not more frequently. Utilizing AI-enabled technologies to drive equitable outcomes in education requires the humans tasked with driving the market to lead with transparency and intention. We can start by asking (1) what role we want AI to play in our school systems, (2) what benchmarks or standards should we aim for when incorporating AI into our schools and classrooms, and (3) how we'll utilize research to rapidly determine if a tool is on the right track toward realizing our vision.

The first step is for districts to collaborate with community members to develop a shared vision for their students: What learning experiences do they want their students to walk away with each day, week, and year? Based on these desired experiences, where might AI be useful and where might it be harmful? Each community's unique priorities and needs will lead to a vision that is specific to the local context, which will enable communities to clearly articulate their needs from an AI-enabled tool.

Districts are in the early stages of developing evidence-based systems to evaluate and procure AI-embedded edtech. Commonly cited risks with AI, such as data privacy, intellectual property, and inequitable access make developing district AI policies a challenge. This challenge is only heightened in the context of technology evolving at a pace we've never previously encountered. Districts need more support in accessing industry benchmarks or standards that ensure the technologies being used in schools have been ethically and thoughtfully designed. School leaders hold immense, yet underutilized, power in this market: When districts develop and uphold policies responsive to AI's risks, edtech developers are incentivized to create the safest, most ethical versions of their solutions (OET, 2024). The dearth of submissions centered on district infrastructure and policy development for decision making highlight an unmet need in this area. Fortunately, more initiatives to support district decision making would also benefit edtech developers who are seeking best practices and ethical approaches to the complex challenges they face in creating AI-enabled tools.

Success also calls for evidence, which will require us to explore new innovations in efficacy research that are more timely and responsive to the rapid changes in AI. In addition to elevating the voices of those impacted by AI-enabled technologies in education in the design and development of these tools, so, too should we elevate their voices in determining measures of success for students and teachers. As a field, it's critical to innovate and create rapid testing that holistically considers the well-being and experiences of learners to determine if each tool is adding value toward the intended vision. Through the exploration of innovations in research, our field can collectively emphasize the importance of impact data to determine if we are going in the right direction when it comes to equity, learning, and community goals.

Finally, this work calls for the investment into technical infrastructure. Districts need to develop infrastructure for evaluation and decision making, and developers need support with access to the highest quality models and datasets to test and explore new ideas. This

infrastructure would create opportunities for teachers and developers representing diverse lived experiences to jump in and explore the ways they believe AI could transform our education system.

We encourage anyone considering building, using, or investing in AI to continue investigating and innovating, while recognizing that AI is not the solution for every problem. We recommend a leveled approach, prioritizing areas where the field is finding the most success with AI (e.g., data analysis and aggregating large amounts of data) and being mindful of the limitations and ongoing challenges to using AI.

Conclusion

Digital Promise and the foundation led this RFI to understand where innovators were most excited around the potential for AI to transform mathematics teaching and learning, and how innovation could be paired with awareness, reflection, and flexibility. The findings from the submissions highlight areas where there are concentrations of exciting ideas – such as creating adaptive, personalized, and engaging content for students and providing teachers with access to classroom discussion data and increased resources for professional learning. Furthermore, themes across the responses represented attentiveness to a core set of risks of using AI, and in several instances surfacing specific mitigation strategies that can be readily utilized by the field. We were inspired by the creativity in the strategies described to address critical risks that come from using AI. In part, we hope that this analysis helps innovators in the thick of these uphill battles to see that there are many enthusiastic leaders in the field working toward the same goals.

While the submissions highlighted exciting ideas, it is important to acknowledge the gaps in the field. For example, very few submissions aimed to offer increased accessibility in the classroom or sought to empower students to become engineers over their own algorithms that will drive their learning. We hope that by identifying these market gaps, it will inspire innovators to begin developing AI-powered solutions in these directions.

The ultimate goal of the RFI was to learn how organizations currently intend to use AI to dramatically improve mathematics instruction and learning experiences so that all students are able to deeply understand, use, and enjoy mathematics. While the submissions represent a small sample of innovative concepts, they offer insights into a potential projection on what is possible for each student and every teacher. Additionally, these responses demonstrate how advances in AI can accelerate progress toward a future where teachers use digital tools to personalize learning, the classroom is an inclusive environment where all students see their real-world interests reflected in the work they are doing, and teachers receive valuable feedback, coaching, and professional learning aligned to the instructional materials.

We encourage developers to begin working more closely with teachers and students to co-design solutions that meet real needs – a recommendation emphasized by the fact that human-in-the-loop workflows were highlighted as one of the most meaningful risk mitigation strategies.

We recommend education leaders, educators, students, and their communities ask questions of the tools they are using and demand products that are truly valuable to them.

Finally, we advocate for more philanthropic organizations to follow the foundation's approach to identify where the market will need support and invest in organizations developing open source solutions who are willing to share learnings in order to transform the education system. In fact, most respondents to this RFI expressed an openness to sharing their responses with other funders.[3]

[3] Most respondents to this RFI expressed an openness to sharing their responses with other funders. While we have shared RFI responses that have given us permission to do so with several other funders, we encourage more funders to reach out to us to get access: productcertifications@digitalpromise.org.

References

Diliberti, M., Schwartz, H. L., Doan, S., Shapiro, A. K., Rainey, L., & Lake, R. J. (2024). *Using artificial intelligence Tools in K-12 Classrooms*. RAND.

Elhussein, G., Hasselaar, E., & Lutsyshyn, O. (2024, April). Shaping the future of learning: The role of AI in education 4.0. *World Economic Forum.*

Green, A., Tench, B., Weinstein, E., & Lara, E. (2024). Teen and young adult perspectives on generative AI: Patterns of use, excitements, and concerns. Common Sense. https://www.commonsensemedia.org/sites/default/files/research/report/teen-and-young-adult-perspectives-on-generative-ai.pdf

Hamilton, I. (2023). Artificial intelligence in education: Teachers' opinions on AI in the classroom. *Forbes Advisor*.

Impact Research (2024). *AI chatbots in schools: Findings from a poll of K-12 teachers, students, parents, and college undergraduates*. Retrieved from https://8ce82b94a8c4fdc3ea6d-b1d233e3bc3cb10858bea65ff05e18f2.ssl.cf2.rackcdn.com/bf/24/cd3646584af89e7c668c7705a006/deck-impact-analysis-national-schools-tech-tracker-may-2024-1.pdf

Jansen, T., Meyer, J., Wigfield, A., & Möller, J. (2022). Which student and instructional variables are most strongly related to academic motivation in K-12 education? A systematic review of meta-analyses. *Psychological Bulletin*, *148*(1-2), 1.

Krall, Fagan, Schwartz, & Boguslav (2024). *AI in professional learning: Navigating opportunities and challenges for educators.* RPPL, Annenberg: Brown University.

Maghsudi, S., Lan, A., Xu, J., & van Der Schaar, M. (2021). Personalized education in the artificial intelligence era: What to expect next. *IEEE Signal Processing Magazine*, *38*(3), 37-50.

Mills, K., Ruiz, P., Lee, K., Coenraad, M., Fusco, J., Roschelle, J. & Weisgrau, J. (2024, May). AI literacy: A framework to understand, evaluate, and use emerging technology. https://doi.org/10.51388/20.500.12265/218

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2023). Uncovering blind spots in education ethics: Insights from a systematic literature review on artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 1-40.

NCTM (2024). *Artificial intelligence and mathematics teaching: A position of the National Council of Teachers of Mathematics.* Retrieved from https://www.nctm.org/uploadedFiles/Standards_and_Positions/Position_Statements/NCTM_AI_Position_Statement.pdf

Ran, H., Kim, N. J., & Secada, W. G. (2022). A meta-analysis on the effects of technology's functions and roles on students' mathematics achievement in K-12 classrooms. *Journal of computer assisted learning*, *38*(1), 258-284.

Stanford University (2023). *AI will transform teaching and learning. Let's get it right.* Retrieved from https://hai.stanford.edu/news/ai-will-transform-teaching-and-learning-lets-get-it-right.

Tetzlaff, L., Schmiedek, F., & Brod, G. (2021). Developing personalized education: A dynamic framework. *Educational Psychology Review*, *33*, 863-882.

U.S. Department of Education (2024). Designing for education with artificial intelligence: An essential guide for developers*. Office of Educational Technology,* Washington DC. Retrieved from https://tech.ed.gov/files/2024/07/Designing-for-Education-with-Artificial-Intelligence-An-Essential-Guide-for-Developers.pdf

Appendix A

The Request for Information was created in Qualtrics, where all responses were collected.

Information About the Respondent

1. What is your name (first, last)?

2. What is your email?

3. What is your role/title?

4. What is the name of your organization?

5. Where is your organization's mailing address?

6. Which best describes your organization?
    a. Public district or school
    b. Charter school
    c. Private school
    d. For-profit edtech company
    e. Nonprofit edtech company
    f. Research institution/University
    g. Non-edtech nonprofit organization
    h. Non-edtech for-profit organization
    i. After-school program
    j. Not listed (please specify):

7. How does your organization (staff, board, advisory groups, partners, etc.) represent the perspectives of those you are serving? (100 word maximum)

8. Are you submitting this RFI in partnership with another organization or individual?
    a. If yes: Which organization(s), university(ies), district(s), and/or individual(s) are involved? Please provide a contact person (name, role, email) for any partners described.

9. Have you or your organization or any of the affiliated responding organizations previously received funding from the Bill & Melinda Gates Foundation?

Information About the Submission

10. What is the geography of your current impact?
    a. United States: Nationwide
    b. United States: Specific state(s) (please specify):
    c. United Kingdom
    d. Sub-Saharan Africa
    e. India
    f. Not listed (please specify):

11. Age group
    a. Early Childhood
    b. Primary (Grades K-5)
    c. Secondary (Grades 6-12)
    d. Post-secondary
    e. Not listed (please specify):

12. Who is the primary beneficiary for this project?
    a. Teachers
    b. Instructional coaches
    c. Students
    d. The field (e.g. research findings will support the field in better understanding AI in math education)
    e. Other primary end users (please specify):

13. Do we have your permission to share your submission with others? Select all to whom you consent to accessing your submission.
    a. The submission can be shared in its entirety with other interested funders.
    b. The submission can be shared in its entirety with other relevant respondents.
    c. The title of the submission and contact information can be shared with other interested funders.
    d. The title of the submission and contact information can be shared with other relevant respondents.
    e. No, please do not share any information.

14. How did you hear about this RFI?
    a. Direct outreach from a partner
    b. ASU+GSV
    c. Newsletter
    d. Social Media
    e. Not listed (please specify):

15. Approach maturity category
    a. Concept has not been prototyped or tested, or is a research project.
    b. Concept has been prototyped and initial field tests have been conducted.
    c. Approach is in use in multiple educational settings with teachers and/or students, is being tested for impact and there is a focus on new feature development and/or scale.

Descriptive Questions About the Approach

16. Title of your research-based approach: (10 word maximum)

17. In plain language, describe your approach. Please incorporate answers to the relevant sub-questions in your response. (500 word maximum)
    a. What is the existing challenge that your approach addresses?
    b. What evidence demonstrates there is a need or demand for this approach?
    c. What is your theory of change for how your approach impacts math teaching and/or learning?
    d. What are the evidence or research underpinnings behind your approach?
    e. If your project is successful, what difference will it make? What are the measures of success?
    f. Do you have existing evidence of impact through related projects?
    g. What is the current reach of your approach (e.g., in the number of teachers, students, schools, or districts served)? If applicable, please list the top three largest districts you partner with. If you are at the early concept stage, what is your intended reach and how do you plan to get there?

18. We believe that co-design or participatory design that actively involves all constituents (e.g., users, clients, designers, researchers) – particularly Black and Latino students and students impacted by poverty and their teachers, as these communities have historically been excluded in education design decisions – leads to better approaches that meet real needs and are usable. What did/will the design process look like for this approach? Please incorporate answers to the relevant sub-questions in your response. (300 word maximum)
    a. Who do/will you co-design with and why?
    b. How are intended end-users driving design and development in an iterative way?
    c. How is impact shared back with co-designers?

19. How does the research-based approach leverage AI in innovative ways? Why do you think the use of AI would drive greater impact? Please incorporate answers to the relevant sub-questions in your response. (100 word maximum).
    a. If this is an early stage concept, please share any current hypotheses for how to either identify the appropriate partners to support the project or develop the technology.

b. What are the risks to leveraging AI for this project? How do you plan to mitigate these risks?

20. What type(s) of technology(ies) do you anticipate using in this project? Select all that apply.
    a. I'm not sure yet
    b. Natural language processing (NLP)
    c. Large language models (LLM)
    d. Small language models (SLM)
    e. Chatbots
    f. Robotics
    g. Automatic speech recognition (ASR)
    h. Computer Vision
    i. Predictive AI
    j. Not listed (please specify):

21. We believe successful partners know their strengths along with the areas where they will need support. Our intention with this question is to understand your organization's self-awareness around the types of supports you need to be most impactful. What do you need to make this idea happen (e.g., expertise in development and engineering, research and measurement, scale and marketing, access to districts/schools, etc.)? Where are your strengths? What gaps do you have that resources could support you in achieving? (100 word maximum)