CASE STUDY

Staying Ahead of the Technology Curve

AI Technology Evolution at Quill.org

Stephen Rockwell, Sarah Kelley, and Sarah Di Troia

SEPTEMBER 2025

PROJECT EVIDENT

About Project Evident

Project Evident harnesses the power of data, evidence, and technology to achieve greater impact. We believe that by empowering practitioners to drive their own data and evidence building while also strengthening the surrounding ecosystem, we can increase the number of effective solutions in the social and education sectors and scale them faster, ultimately producing stronger, more meaningful, and more equitable outcomes for students and communities.

Project Evident's **OutcomesAI** practice provides consulting, technical assistance, resources, and tools to support practitioners – nonprofits, school districts, and funders. We achieve this by strengthening their ability to utilize AI to enhance their understanding, improve their impact, support informed decision making, advance R&D, and allocate resources toward achieving better and more equitable outcomes. We recognize the potential for misuse of data, evidence, and technology and seek to limit harmful practices. We serve on the EDSAFE AI Steering Committee and strongly recommend the S.A.F.E. Benchmarks Framework for K-12 AI efforts. Project Evident's differentiator is its use of AI to drive outcomes. We support processes to detect and avoid technology overriding our evaluative work in delivering equitable outcomes.

About the Equitable AI Adoption Project

Artificial intelligence (AI) and generative AI hold great promise for helping nonprofits expand their services and achieve more equitable outcomes for the people and communities they serve. Few in the philanthropic, social, and education sectors would claim satisfaction with society's progress in addressing persistent social problems. Grantmakers and nonprofits share the goal of scaling impact, and AI provides new tools to achieve this goal.

A February 2024 working paper by Project Evident and Stanford's Institute for Human-Centered Artificial Intelligence found that approximately 80% of funders and nonprofits believe their organizations would benefit from using more AI, specifically for mission-related work. However, there is a question about "how" – a lack of clarity about how AI will benefit individuals and organizations, as well as a lack of organizational expertise and materials about AI for social and education sector organizations, were the most frequently cited barriers for funders and practitioners, after concerns about bias. With the support of the Gates Foundation, the Equitable AI Adoption (EAIA) project aims to inspire and inform practitioners and educators on how AI can help them achieve their mission. To that end, EAIA is surfacing, creating, and disseminating stories of early adopters to study progress, distill broadly applicable insights, and share findings. At the same time, we are leading a Community of Practice comprising 15 nonprofit organizations in developing a practical and actionable tiered AI adoption framework to support others on their journey.

Acknowledgments

We are incredibly grateful to the teams at Project Evident (Stephen Rockwell, Sarah Kelley, Sarah Di Troia, Kelly Fitzsimmons, Nora Greene, Laura Schoonmaker, Christina Gordon, Meaghan Birnie, and Mariko Tada) and Quill.org (Peter Gault, Maheen Sahoo, Rebekah Bergman, and Daniel Drabik), members of our Equitable AI Adoption Design Committee (Peter Gault, Sasha Rabkin, Sarah Radcliffe, and Lorne Rodriguez), and especially to the Gates Foundation for making this work possible.

This publication is based on research funded by the Gates Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect the positions or policies of the Gates Foundation.

CASE STUDY

Staying Ahead of the Technology Curve

AI Technology Evolution at Quill.org

Introduction: Why Keeping Pace with AI Evolution Matters

Education and social nonprofit practitioners are contending with a new reality: Artificial intelligence (AI) is shifting the dynamics of both organizational life and pathways to impact. This transformation has already begun across the US economy. In 2024, the proportion of survey respondents reporting AI use by their organizations jumped to 78% from 55% in the prior year. Similarly, the number of respondents who reported using generative AI in at least one business function more than doubled – from 33% in 2023 to 71% in 2024.[1] In short, AI evolution is underway and rapidly advancing, and grantmakers and practitioners are innovating on how AI can enhance outcomes. This case study highlights the efficiencies, reach, and impact achieved by Quill.org through its use of AI, while emphasizing the necessity of keeping pace with the AI evolution by centering program strategy, integrating program staff, building adaptive capacity, and accepting sunk costs.

AI models are becoming increasingly cheap to use; the cost of querying a large AI model (such as ChatGPT, Claude, Gemini, etc.) experienced a more than 280-fold reduction in approximately 18 months, while smaller, less expensive AI models are delivering stronger performance.[2] This rapid decrease in costs makes AI more accessible to practitioners in the education and social sectors, unleashing new ways to enhance program outcomes. As Maheen Sahoo, Quill's Managing Director of Strategic Partnerships, says, "As models kept coming out that were better and better and the cost got cheaper and cheaper, it became really clear that Generative AI is something that we should take advantage of. This is the future." However, practitioners seeking to realize the potential of AI for impact must grapple with a fundamental shift: the pace of technology innovation is now measured in weeks, not years. Large technology companies are retiring their models (e.g., ChatGPT-3.5 becoming ChatGPT-4.0) within one to two years, and tasks — such as multi-step reasoning, interactions with software tools, or direct use of browsers — that were previously impossible for AI to complete are now routine. This dynamic environment disrupts conventional program management and execution models, which often rely on advanced planning that makes responding to technical shifts challenging.

To stay ahead of the curve in this rapid evolution, nonprofits using AI to enhance outcomes should embrace four key principles:

[1] Artificial Intelligence Index Report 2025

[2] Ibid.

CENTER PROGRAM STRATEGY: Centering the program logic model can allow nonprofits to make effective technical roadmap decisions on when and how to embrace new technologies in ways that enhance outcomes.

DEPLOY CROSSFUNCTIONAL TEAMS TO SUPPORT RAPID ITERATION: To drive effective change, program staff must be involved at every stage of AI implementation, working in tight coordination with technical and measurement, evaluation, and learning staff.

BUILD ADAPTIVE CAPACITY AND PROCESSES: Continuous learning cycles and iterative design processes allow nonprofits to adapt to the changing technology landscape.

CULTIVATE A RAPID CHANGE MINDSET: Technology evolutions require making time to explore and play with new technology and letting go of legacy systems, even after substantial investments have been made.

> "We are in this period of very rapid iteration and growth. So every six months, it feels like what we were using six months ago is dated."
>
> *—Daniel Drabik, Chief Technology Officer*

Managing the velocity of AI evolution to optimize program impact requires more than routine technical upgrades, and Quill.org is a strong example of smoothly navigating this new world. Quill's leadership was quick to recognize the power of Generative AI, encouraged early experimentation, and decisively shifted to invest in the new technology once it became clear that it could meet their high standards. The lessons from Quill apply to all nonprofits seeking to utilize AI, even those that do not embed technology as deeply as Quill. We are grateful to Quill for sharing their story – and some painful choices – in clear and candid detail that can serve as an exemplar for other social and educational practitioners seeking to surf the big waves of technological change without being overwhelmed by them.

Definitions

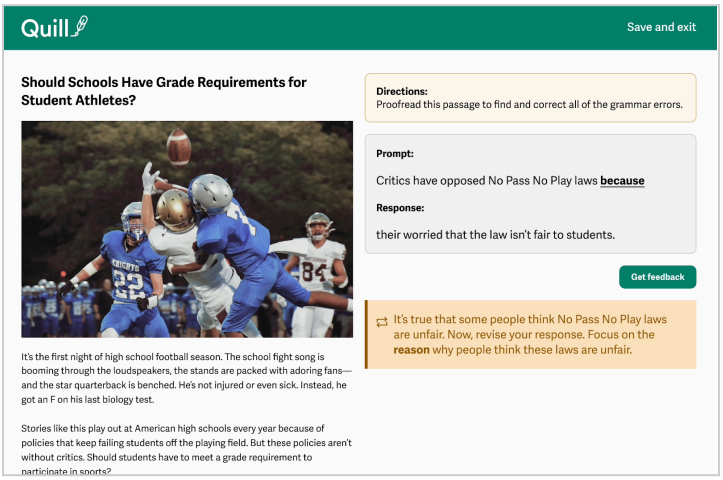
Predictive AI: AI models that predict trends and behaviors based on cause-and-effect patterns in the data. These algorithms use patterns in previous data to make predictions about what will happen. An example of this is an algorithm that predicts the likelihood of a student dropping out based on their engagement with an online learning platform.

Generative AI: AI models that can create convincing or pleasing imagery, compelling text, or coherent audio outputs. Generative AI uses very large neural networks to generate new content based on the patterns it has learned from enormous volumes of text, audio, and video data. ChatGPT, Claude, and Gemini are prominent examples of generative AI.

Who is Quill.org?

In 2014, Peter Gault, Quill's founder and Executive Director, "imagined a world in which all people had the writing and critical thinking skills to engage in active, intelligent debate and discourse."[3] To make this vision a reality, the organization provides evidence-based, AI-powered literacy tools designed to help students become stronger readers, writers, and critical thinkers through its accessible digital platform, which is provided free to students and teachers. By leveraging technology, Quill can scale its impact far beyond what a traditional tutoring approach could; Quill serves about 10% of schools across the country, or about three million students a year, more than 63% of them in Title I schools.[4] The organization has recently surpassed the milestone of $1 billion in impact, providing 40 million hours of free tutoring to date. "Reaching the billion-dollar impact mark proves that advanced AI can advance equity, not just efficiency," says Gault.[5]

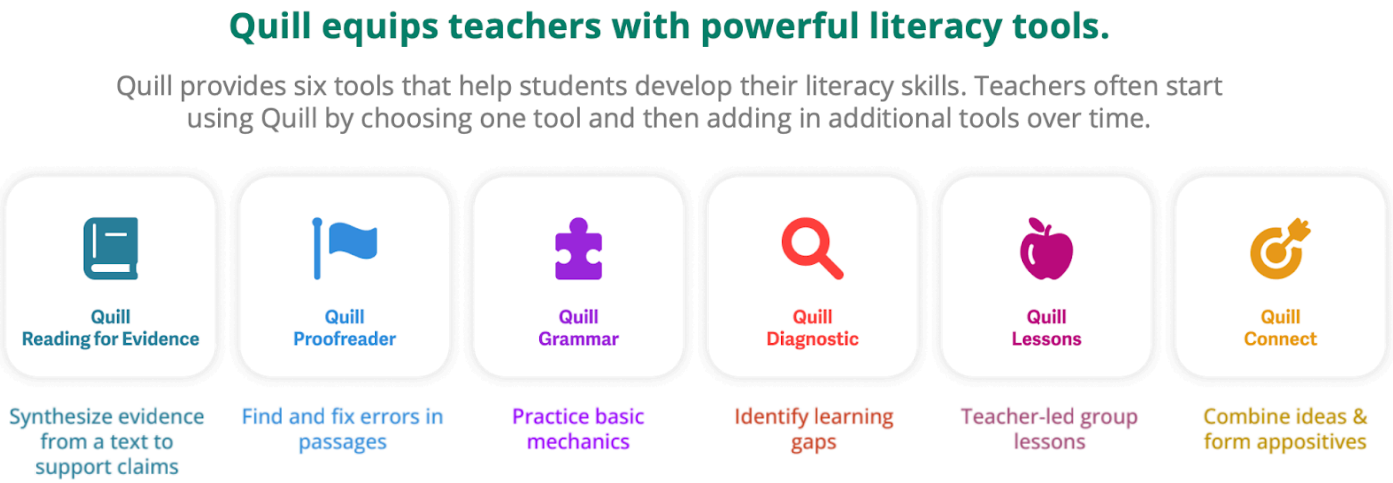
> "The core to our work is: How do you make the AI really fine-tuned toward the needs of a student, in particular, in a low-income, Title I school?"
>
> *—Peter Gault, Founder & Executive Director*



The Quill.org platform helps students develop the skills needed to effectively share their thoughts in writing through a variety of structured writing activities. Quill is designed for use in the classroom as a supplement to regular instruction. Across their different offerings, the platform promotes student learning through cycles of writing, feedback, and revision. Quill offers various specific writing activities, including drills focused on sentence connection, grammar drills, a tool for practicing proofreading, whole-class lessons, and exercises that encourage students to respond in writing to provided texts (the newest addition). When using the latest tool, called "Quill Reading for Evidence," students begin each activity by reading a short (600-800 words) nonfiction source text. Then, they write about what they read, completing a series of open-ended writing prompts to demonstrate their understanding of the source text. This tool emerged because teachers "want[ed] a rich way to teach [their] students critical thinking, reading and writing," explains Maheen Sahoo, Quill's Managing Director of Strategic

[3] Quill.org website, About
[4] "Quill.org Becomes a "Non-profit Unicorn," Generating $1 Billion in Social Impact," Quill.org press release
[5] ibid

Partnerships, and represents an important evolution of Quill's work in integrating writing with academic content. Across all of Quill's products, students receive immediate, responsive coaching from the AI to improve their writing, rapidly strengthening key literacy skills through response and feedback cycles. Evidence suggests that this works: in a recent evaluation, students who used Quill's sentence combining tool had a statistically significant improvement in their ability to revise paragraphs after using the tool for four weeks, an improvement that was sustained in a two-month follow-up.[6]



A Mission-Oriented Approach to Technology

Gault's vision for Quill.org centered around a mission to "help all students become strong writers, readers, and critical thinkers."[7] "Our mission across our six tools, if summed up in one sentence, is to help students learn to write and write to learn. That's the motto going forward," says Sahoo. To achieve this, Quill delivers evidence-based writing instruction that leverages digital technology to provide real-time feedback. Essential to Quill's success in delivering sustained impact at scale is its grounding in pedagogy. While many AI-powered education technology tools promise to revolutionize education, they often lack a nuanced understanding of how to apply technology to actually meet students' needs and deliver on educational outcomes. Being grounded in pedagogical best practices is essential to Quill's success in delivering sustained impact at scale. "Our theory of action is that when kids get really good feedback on their writing, they're better able to learn and build their skills," explains Gault. This commitment to improving educational outcomes guides Quill's AI strategy. Every investment in AI is filtered through their theory of action to create a technical roadmap that reflects and reinforces the organization's purpose. When deploying AI to enhance outcomes, the program logic model and the technology roadmap are foundational plans that must work in tandem.

[6] Chang, B. & Lazowski, R. (2023). Effects of a Quill.org intervention on paragraph revision. College Board Report.
[7] Quill.org website, Mission statement

This emphasis on mission impact motivates investment in and deep engagement with a team of former educators who serve as the subject-matter experts behind the artificial intelligence. Quill's Generative AI Playbook, which guides their use of technology, notes, "ten full-time curriculum developers at Quill – primarily former classroom teachers – manually evaluate over 100,000 student responses each year." And Quill's bridging of subject matter expertise and technical roles goes beyond their internal teams to include extensive user engagement. Their Teacher Advisory Council comprises over 300 educators who review every AI-powered writing prompt. "They don't just look at the AI once: we complete three rapid-cycle evaluations," Quill's Playbook notes. This advisory structure ensures that classroom realities inform every stage of AI development. Quill embeds educators so deeply into its AI workflow because this drives impact. Pedagogical best practices help students learn to write, enabling them to learn more effectively.

> "In building out the Quill.org Reading for Evidence tool, we found that while our first five tools teach students how to write, we now had an opportunity to help them *write to learn*."
>
> *—Maheen Sahoo, Managing Director of Strategic Partnerships*

Surfing Big Waves of Technology Change

Quill.org originally used predictive AI – which we refer to here as earlier generations of artificial intelligence technologies focused on using patterns in data for prediction or classification – to provide students with feedback on their writing, investing many years in learning the existing AI technology. "[W]hen Reading for Evidence launched, the feedback was powered by predictive AI, which meant that the team started with the sentence stem, wrote out the article, and then was tasked with figuring out what our students were going to write in response to this," describes Rebekah Bergman, Vice President of Curriculum. Curriculum team members, she explains, had to identify all the ways students might respond to a given sentence and "look at thousands of pieces of feedback so that we could come up with the categories." All of this 'pre-work' was required to develop the categories into which predictive AI would sort student responses, so the pre-work of category creation had to be done by subject matter experts (former teachers) for it to be accurate. Through Quill's work with these earlier technologies, they developed deep expertise in how to get AI systems to deliver usable, pedagogically sound feedback.

In 2022, ChatGPT-3.5, the first widely available generative AI, was released to the public. Generative AI represents a new wave of deep-learning-based models designed to generate new text, images, and audio based on learning from massive amounts of input data. Amid the initial wave of excitement following the launch of ChatGPT-3.5, Quill.org did not rush to integrate generative AI. As Quill Chief Technology Officer Daniel Drabik puts it, "We had to wait for the models to catch up." Quill requires that its AI solutions produce the 'correct' feedback 95-99% of the time. In this case, correctness was measured against a benchmark of teacher-created feedback. In contrast, the first wave of generative AI models was only able to match human

feedback 60-80% of the time. Quill also avoided the early ChatGPT version because “it felt very laggy, and we didn't feel comfortable launching with it,” as Drabik states. Quill's emphasis on its program logic model – which includes delivering timely, accurate feedback to students – guided the organization in avoiding the implementation of generative AI until the technology had caught up to its needs in terms of speed, cost, and performance.

> “Eventually, the models caught up and got good enough and cheap enough.”
>
> *—Daniel Drabik, Chief Technology Officer*

Despite ChatGPT-3.5’s early, insufficient performance, Quill.org knew generative AI would eventually be an option, so VP of Curriculum Rebecca Bergman was "given the directive just to start experimenting with prompts." She spent weeks iterating on prompts and bringing them back to the team, and this extensive exploration catalyzed organizational experimentation, building on Quill's culture of continuous learning. Gault champions this mindset, explaining, “[G]enerative AI is brand new. The way to learn about it is to use it and to try it.” Technology continues to evolve rapidly, so Quill has not stopped experimenting. Drabik explains that his technology staff is "spending 10% of our research time looking at what is possible with the newest things.”

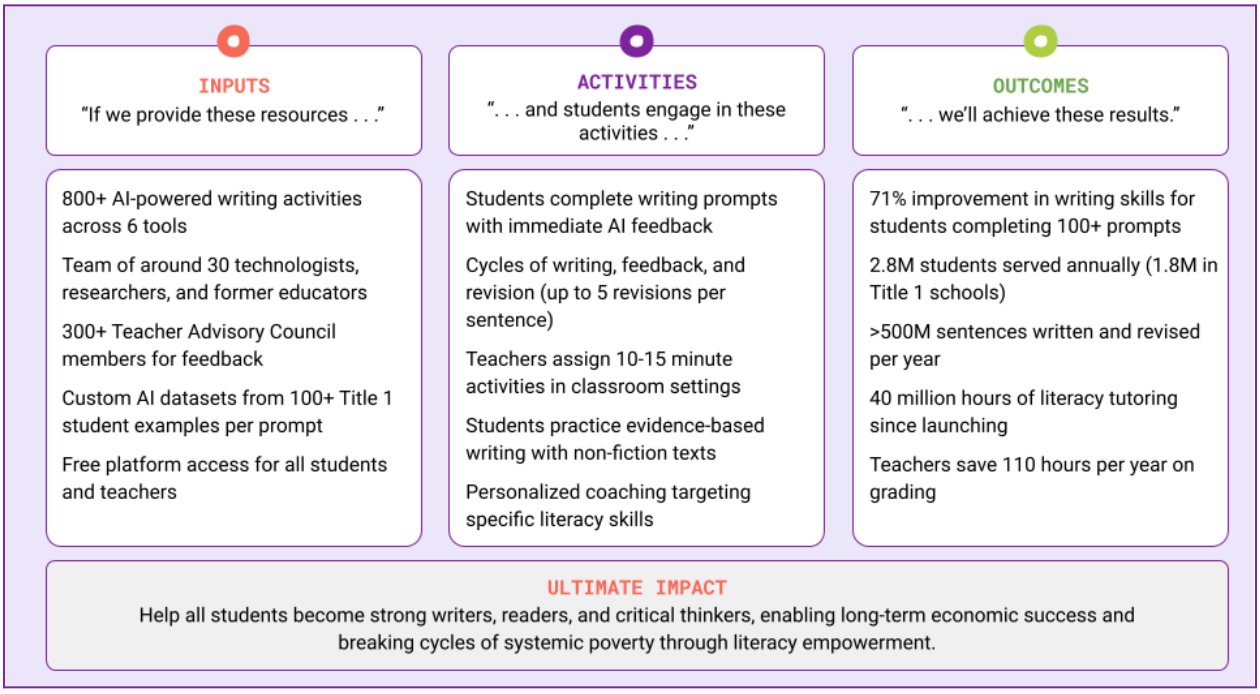
This experimentation positioned Quill to act quickly once the AI met its speed, performance, and cost thresholds. This enabled Quill to pivot decisively despite the fact that they "had put $5 million of R&D into building out the predictive AI system. Essentially, in the generative AI transition, we were rebuilding from scratch,” explained Gault. But Quill realized that “generative AI would be the future and. . . that it would unlock future products as well,” says Gault. So, rather than clinging to legacy systems, Quill made the difficult but necessary decision to undergo a technology transformation.

Leveraging capacity to fund impact, not hype.
Grantmakers aiming to fund impactful AI projects should ask “Could you do this without AI?”, Sahoo explains, in order to identify projects where AI can drive real impact rather than those that are just following the hype. “Our funders really valued that we weren't just creating an AI tool for the sake of creating an AI tool,” she continues. Quill.org’s funding journey reflects the need for grantmakers with the knowledge to evaluate whether proposed AI work reflects a genuine opportunity. Quill has found that grantmakers who have “started to build capacity or already have a team dedicated to thinking through AI strategy, they were more likely to partner with us because they'd already built the confidence that this is something they could see themselves supporting,” Sahoo explains.

> "It's pretty painful to spend a lot of time building something and then having to let that go and having to rebuild."
>
> *—Peter Gault, Founder & Executive Director*

As Quill.org shifted to generative AI, the curriculum team continued to have a pivotal role "because, as an engineer writing AI prompts, you actually can't build AI-powered educational tools to the best of your abilities without a strong pedagogical backing and educators' input," as Sahoo explains. "[T]he change has been a greater collaboration, and an increase of involvement of our curriculum team in the technical input of the actual prompt engineering," Sahoo continues. Bergman adds, "We're still going to do everything in our power to train it [generative AI systems] in the ways that we've learned that good feedback, strong feedback in the writing classroom needs to be."

Quill.org Theory of Action



Importantly, these collaborative efforts occur within a tightly coordinated structure in which the curriculum team is directly embedded in the software development process.

Additionally, they employ a product manager responsible for coordinating priorities across programs and technical teams. As Sahoo describes, this staff member "sits on the curriculum team in her role as AI curriculum product manager. And she is the connector between the product/technical side and the curriculum side." This bridge role translates program imperatives into technical requirements and vice versa.

> “We do a lot of testing and iterating on a prompt. That's where the bulk of the work is.”
>
> *—Rebekah Bergman, VP of Curriculum*

This integration between software and curriculum teams facilitates rapid iteration in the software development cycle. "By far the most important part of this playbook. . . is the iteration cycle that gets us from an AI feedback loop that works some of the time to a feedback loop that is consistent and reliable," as Quill's AI playbook states. This iteration involves testing and experimentation from prompt development through beta launch. After iteration on prompt development, AI outputs are tested by “manually grading another set of at least 300 pieces of student response for the first round of iterating,” Bergman says. Then, the process continues with “a lot of time-consuming manual checks and the teacher team looking at it for a general feel,” Drabik explains. After these rounds of iteration, the new lesson is released on the platform to student beta testers. As Bergman notes, "Getting a view of actual student sessions is probably the most insightful thing that we do." The tight connection between the curriculum team and the engineering team enables the engineering team to utilize the insights generated from these iterations to identify key areas for improvement or new features to incorporate into the AI tool itself and to deploy these changes in the next cycle. Because program and technical staff work together in a coordinated manner, changes to the process can be made and tested quickly.

As the pace of AI evolution continues to increase, Quill.org stays current by deploying robust systems that enable the team to verify that new models are functioning as expected. "When we have to upgrade, we must run more experiments with a newer model," Drabik explains. Quill has invested in infrastructure to support this benchmarking, which also ensures that the release of newer models does not compromise the live product (for example, if a new model requires different prompting, it wouldn’t work well with an existing prompt). This infrastructure gives the team confidence in a technology landscape that is changing at an unprecedented speed. “When Google releases a new model of Gemini, the day they release it, they release the deprecation date as well,” notes Drabik. So, Quill builds its systems expecting that the model it is using today will not be the same as the one it will be using next year.

> “The core thing that is unique to Quill versus a lot of other folks in the space has been our emphasis on using custom training datasets to shift the AI so that the AI isn't just on its own. . . grading the students, but it's the teachers grading the students with the AI, using their judgment calls.”
>
> *—Peter Gault, Founder & Executive Director*

How to Stay Ahead of the Technology Curve

Quill.org is a technology-powered nonprofit that delivers program impact through AI-driven feedback. Nonetheless, the methods they use to keep pace with technological change – centering program strategy, leveraging program staff, building adaptive capacity, and embracing rapid change – can inform the journeys of other technology-powered nonprofits.

CENTER PROGRAM STRATEGY: Nonprofits' adoption of new technology must be fundamentally centered on their core missions to be effective. Quill's program logic model detailing how to provide pedagogically sound high-quality feedback on students' writing has driven their technical decisions, from not using the earliest generation of generative AI (which did not meet the high accuracy standards the organization embraces) to transitioning their entire technical infrastructure to the latest generative AI models despite their $5 million investment in earlier technologies. Program logic model clarity enables organizations to evaluate new AI capabilities and develop technology roadmaps within the context of delivering on programmatic goals, rather than pursuing AI capabilities for their own sake.

> "Clear programmatic goals provide the benchmarks and guide rails for rapid AI experimentation."
> *– Quill's Generative AI Playbook*

DEPLOY CROSSFUNCTIONAL TEAMS TO SUPPORT RAPID ITERATION: Cross-functional teams are needed for the rapid cycle learning and tight feedback loops that AI implementations require to effectively drive programmatic value. Quill.org achieves this through strong coordination across program and technical teams: engaging educators in their AI development and evaluation at each step in the process, both through its in-house curriculum developers (former classroom teachers) and through engagement with an external advisory committee of 300 teachers. This engagement allows them to develop pedagogically sound products that drive real learning, enabling them to maintain their social impact mission while embracing rapid technological change.

> "Working with teachers is not a nice-to-have. It's absolutely essential to being successful with this work."
> *– Peter Gault, Founder & Executive Director*

BUILD ADAPTIVE CAPACITY AND PROCESSES: Quill.org can engage in continuous adaptation because of its purposeful embrace of ongoing iteration and testing. As they develop new tools and content, their development processes support continuous iteration, building space for both exploration and systematic testing of new technologies and approaches before releasing them to their users. These processes create space for innovation and also the structure needed to innovate safely.

> "Let's play with this. Let's try things out. Let's run experiments. Let's run tests."
> *– Peter Gault, Founder & Executive Director*

CULTIVATE A RAPID CHANGE MINDSET: The pace of AI evolution means that new versions of large language models are being released in 12- to 18-month cycles and that new technologies (like Generative AI) need time for exploration and play by staff and could require a significant overhaul of a nonprofit's existing tech stack. Before adopting Generative AI, Quill.org invested in time for staff to experiment and learn about the technology. When Quill moved forward with GenerativeAI, they abandoned years of development work, representing a tremendous financial investment. Only by acknowledging the reality that the fast pace of change made these models obsolete and embracing the need to reinvest in new technologies was Quill able to keep pace with the AI revolution. Quill seeks out funders with a deeper knowledge of AI, who thus understand the need to reinvest in new technologies as AI evolves.

> "One thing I would advise nonprofits if you're getting into this world is really invest in your tooling to build these things. . . [B]uilding out your experiment infrastructure is really important."
>
> *– Daniel Drabik, Chief Technology Officer*

Recommendations

Both grantmakers and nonprofit practitioners should plan for rapid technological change, recognizing that AI investment often has a shorter shelf life. For nonprofit practitioners, a focus on driving program outcomes leads organizations to upskill program staff to guide implementation and success metrics, partnering with technical teams to determine whether and when to leverage the latest AI models or AI functionality. For grantmakers, this means rethinking technology investments from one-off projects to ongoing integration with program funding.

FOR NONPROFIT LEADERS

- Have a well-defined program logic model to serve as the foundation for your technical roadmap.
- Prepare for technology as an ongoing investment, not a one-time expense.
- Equip program staff with the skills to lead projects using AI to enhance outcomes in order to set and maintain success metrics amid ongoing technological evolution.
- Employ cross-functional teams comprising program and technical staff to create the spaces for learning whether and how best to apply new AI models and tools.
- Create space for play and experimentation with new technologies.
- Adopt continuous learning processes to allow for ongoing learning, testing, and deployment.
- Be comfortable with the fact that the evolution of AI will sometimes require rebuilding technical systems and sunsetting legacy tools.

FOR GRANTMAKERS

- Keep abreast of AI evolution in order to understand the ramifications for nonprofit organizations.
- Reorient budget analysis to include technology as a program cost, not administrative overhead.
- Be comfortable with sunk costs, as AI models and tools deployed today will likely need to change within the next year. Hold accountability for the program outcomes with the understanding that the how to achieve the outcomes may change in the grant period.
- Fund the ongoing costs of AI evolution for programs, including the upskilling of staff and slack for learning and incorporating new technology.
- Champion innovation by providing general operating support or requesting that tech-oriented R&D be a line item in program-related grants.

> "Generative AI is unlocking a new chapter of Quill's roadmap. I think all of us in the space did not expect how quickly generative AI would be introduced into the world and how much smarter it would get in such a short window of time. For us, it's a really exciting opportunity for deeper learning and deeper impact for students."
>
> *—Peter Gault, Founder & Executive Director*

Appendix

Resources

- Quill's Designing Ethical AI for Learners: Generative AI Playbook for K-12 Education
- Funding the Future: Grantmaker Strategies in AI Investment
- Inspiring Action: Identifying the Social Sector AI Opportunity Gap
- Equitable AI Adoption (EAIA): Highlighting AI in Action
- Sustaining Scaled Impact, AI & Technology at Crisis Text Line