CASE STUDY

Penguin Solutions and NVIDIA Partner to Build the Georgia Tech AI Makerspace

PENGUIN™ SOLUTIONS

Today's undergraduate and graduate students will soon join a workforce reshaped, in every way, by artificial intelligence (AI). The World Economic Forum estimates 97 million AI jobs will be created by 2025. While students need rigorous, hands-on AI training to thrive in an increasingly AI-driven professional landscape and world, the tools and technology underpinning this real-world experience have been largely out of reach. Advanced AI technologies are typically so complex and expensive that only large corporations can afford to provide access, creating a major barrier to developing and preparing the next generation of AI-ready talent. Georgia Institute of Technology, Penguin Solutions, and NVIDIA are working together to overcome this significant challenge by giving students access to real-world AI experience.

The Opportunity: Harnessing AI's Potential in Higher Education

Through the creation of the AI Makerspace— an innovative AI supercomputer located on the university's campus - Georgia Tech, Penguin Solutions, and NVIDIA have set a new standard for AI education. This first-of-its-kind initiative serves as a sandbox, providing students with hands-on access to cutting-edge AI tools and technologies, allowing them to explore and understand emerging AI models through practical experience. Aimed at democratizing access to powerful computing resources typically reserved for researchers or tech companies, students are now able to integrate this technology into coursework and engage with AI models, applications, and solutions directly. This approach deepens students' understanding and strengthens their practical AI skills, preparing them to be part of the AI workforce and influence the future of technological innovation.

"I can do so much that I couldn't do before on my own computer, like LIDAR mapping, because it's limited by how much memory I have. I only have a measly four gigabytes of VRAM on my computer. With the AI makerspace, I can get hundreds of gigabytes, and it unlocks a whole new possibility."

Min Han, Student,
Electrical and Computer Engineering,
Georgia Tech

"Before I came to Georgia Tech, I didn't have a lot of experience with artificial intelligence. Now that I'm learning, it's fascinating and important to understand how the AI model makes decisions and how it comes up with an answer."

Julia Barghouti, Student,
Electrical and Computer Engineering,
Georgia Tech

The AI Makerspace also offers valuable benefits to faculty by enabling them to harness the potential of AI in an educational capacity. The initiative allows the university to enhance or redesign existing AI courses as well as create new courses—both introductory and advanced—that teach AI principles to students across all majors.

"What we're seeing on college campuses right now is the dawn of what can rightfully be called 'Generation AI' —the first generation that will experience its entire professional life in a world reshaped by machine learning and AI technologies. The AI Makerspace, in this sense, is an invaluable incubator, and we're thrilled to be collaborating with Georgia Tech and NVIDIA to make it a reality."

Pete Manca, President,
Advanced Computing at Penguin Solutions

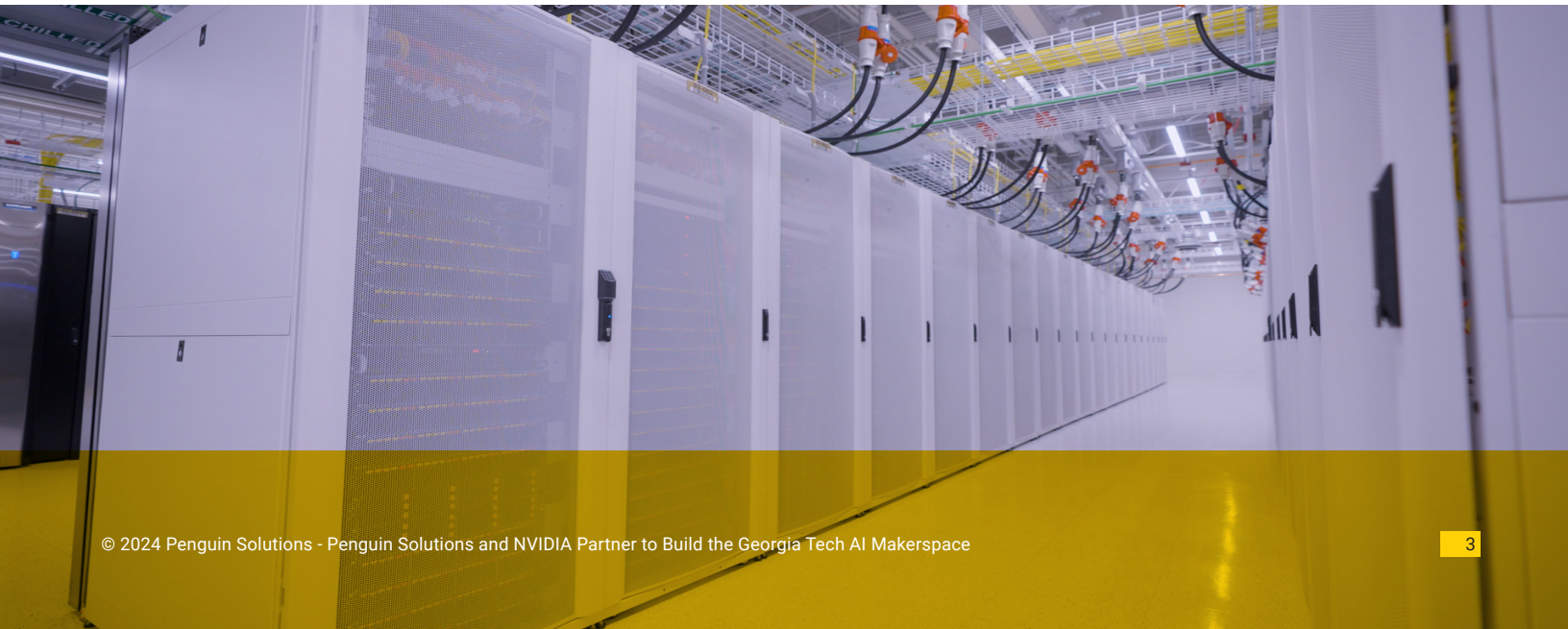
The Challenge: Bringing A "First-of-Its-Kind" AI Supercomputing Hub in Higher Education to Life

Designing, deploying, and operating AI clusters is a complex endeavor, requiring specialized skills, tools, and knowledge to mitigate or overcome the challenges involved. Success hinges on a tightly integrated, finely tuned AI infrastructure specifically designed for unique workloads and environments. AI platforms need to achieve an optimal balance among compute, storage, and network performance to deliver optimal performance and results.

Penguin Solutions has worked with Georgia Tech for more than two decades. Its experience in high performance computing deployments and connections with other industry partners, like NVIDIA in this project, is what led to success in delivering on the vision and technological requirements of the AI Makerspace.

Technology and Expertise: Design, Build, and Deploy Advanced AI Infrastructure at Scale

With more than 25 years of HPC experience and 75,000 GPUs deployed and under management, Penguin Solutions has deep experience and technology expertise in AI infrastructure. Penguin's experts ensured that the AI Makerspace solution was tailored to Georgia Tech's needs, built the system in its factory, and deployed the cluster to ensure operational readiness. Penguin's experts invested the time to understand Georgia Tech's vision and technical requirements, and then designed and deployed the AI infrastructure to meet their unique needs to maximize GPU availability and utilization. As a result, Penguin Solutions has delivered a comprehensive solution that features tightly integrated, end-to-end compute, data management, networking, software, and infrastructure, providing the AI cluster with the ability to process intense amounts of data with ultra-low latency. Penguin Solutions, with more than two decades of serving customers, provides professional services and support for the AI Makerspace ensuring maximum performance and stability.



The vision for AI Makerspace is to grow the solution to serve the entire Georgia Tech Campus. The initiative will be made available to additional colleges within Georgia Tech and rolled out in phases.

Phase 1 of the AI Makerspace serves the Georgia Tech College of Engineering. Built on an NVIDIA AI platform, AI Makerspace includes 20 NVIDIA HGX H100 systems, housing 160 NVIDIA H100 Tensor Core GPUs, interconnected by an NVIDIA Quantum-2 InfiniBand networking platform with in-network computing, which permits extreme performance for cloud-native supercomputing at any scale. To illustrate the immense computational power of this system, a single NVIDIA H100 GPU can complete a multiplication task in one second that would take all 50,000 Georgia Tech students 22 years to finish.

Phase 2 envisions an additional 18 NVIDIA HGX™ H200 servers in the fall of 2024. The NVIDIA HGX H200 supercharges AI workloads and features the NVIDIA H200 Tensor Core GPU with advanced memory to handle massive amounts of data for generative AI and high-performance computing workloads.

NVIDIA accelerated computing combined with NVIDIA AI Enterprise Software provides universities with a comprehensive, turnkey AI platform, allowing students to gain hands-on experience with the world's most powerful AI tools and creating endless possibilities for innovation.

The AI Makerspace expands Georgia Tech's foundational, theory-focused AI curriculum by offering students a hands-on platform to tackle real-world AI challenges, develop advanced applications, and present their AI-driven ideas at scale. The initiative, currently available to students, has expanded since its April 2024 launch and has been incorporated into the curriculum of all eight engineering schools. Students are accessing the AI Makerspace to learn, experiment, prototype, and showcase their AI-driven ideas at scale.

> "People are doing machine learning everywhere, but I think offering a tool and infrastructure that allows our students to really work on problems that are meaningful – that's unique. All these tools are in the hands of our students and I'm expecting them to come up with ideas that we haven't thought about yet."
>
> Matthieu Bloch, Associate Dean of Academic Affairs, College of Engineering and Professor, Electrical and Computer Engineering, Georgia Tech

By spring 2025, all Georgia Tech engineering students — both undergraduate and graduate — will have access to non-instructional learning. In 2026, Georgia Tech plans to set up the AI Makerspace Omniverse, a sandbox for augmented reality (AR) and virtual reality (VR). The education and research hub is based on NVIDIA Omniverse, a platform for connecting and developing 3D tools and applications and will be available to all students.

Learn More

To learn more about our artificial intelligence solutions visit https://www.penguinsolutions.com/solutions/ai

PENGUIN™ SOLUTIONS