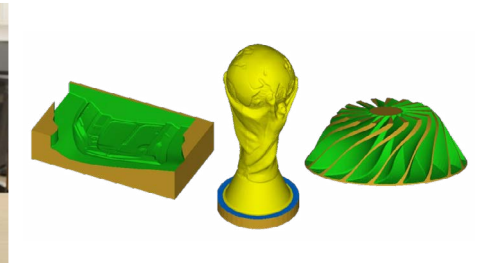
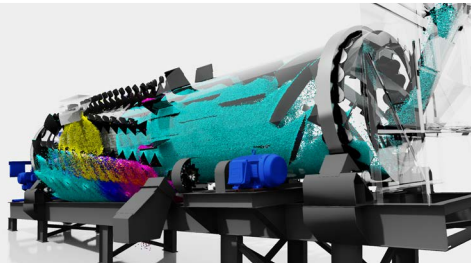
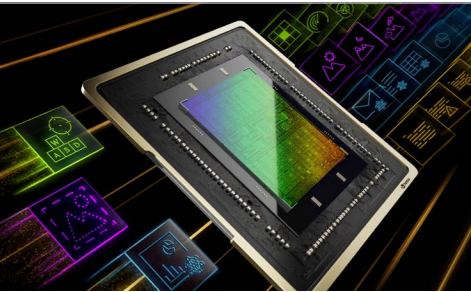


SPECIAL DIGITAL
ISSUE

NEXT-LEVEL ENGINEERING



Advanced
PRODUCT DEVELOPMENT
Resource Center

DE247
Digital Engineering


nVIDIA

DELL
Technologies

Welcome AI in Focus

Contents

- 4 The AI Workstation**
At Dell Technologies World, Dell and NVIDIA outline the future of PCs as more companies adopt on-prem AI.
- 8 GPUs Lift Heavy Simulation Load**
Using Altair® EDEM™ and NVIDIA RTX™ GPUs, Astec Industries has improved and accelerated its DEM simulations.
- 10 CAM Embraces GPU Acceleration**
New CAM software releases leverage NVIDIA RTX GPU power to provide faster results.
- 12 AI Model Integrity**
Ensuring useful AI-based simulation results requires robust data curation and training.
- 14 AI Meets Simulation at NAFEMS Americas**
Technology leaders provide guidance on how artificial intelligence can enhance simulation workflows.
- 16 AU 2024: Autodesk Offers Glimpses of the Future with Project Bernini**
New Proof of Concept Hints at AI Training Based on Proprietary Data.
- 18 Siemens and Sony Showcase NX for VR at Siemens Realize Live 2024**
Cloud-Hosted NX X, Partnerships with NVIDIA and Microsoft, and More.

The Advanced Product Development Resource Center (APDRC) is an ever-growing repository of information to help design engineering teams take advantage of new and innovative technologies that help them create better designs faster.



Artificial intelligence (AI) and GPU acceleration continue to advance in the design and simulation market, opening up new levels of efficiency and enabling advanced design and workflows.

Both NVIDIA and Dell Technologies are at the forefront of integrating these technologies into engineering applications.

In this Special Digital Issue from APDRC, we take a look at how both AI and GPU acceleration are unlocking new avenues of innovation in the engineering world.

You can learn more by visiting the resource center at [APDRC.com](https://www.apdrc.com) for the latest developments on technological advances, videos, white papers and more from *Digital Engineering*, Dell and NVIDIA. We hope you enjoy the digital issue.

Brian Albright

Editorial Director, *Digital Engineering*, Peerless Media, LLC

Comments? E-mail me at: balbright@digitaleng.news

Editorial Staff

Darrell Dal Pozzo
Publisher

Brian Albright
Editorial Director

Paulette Chevalier
Art Director

Peerless Media, LLC

Brian Ceraolo
President & CEO

Editorial Office

Peerless Media, LLC
50 Speen Street, Suite 302 • Framingham, MA 01701 • 508-663-1590
de-editors@digitaleng.news • www.DigitalEngineering247.com

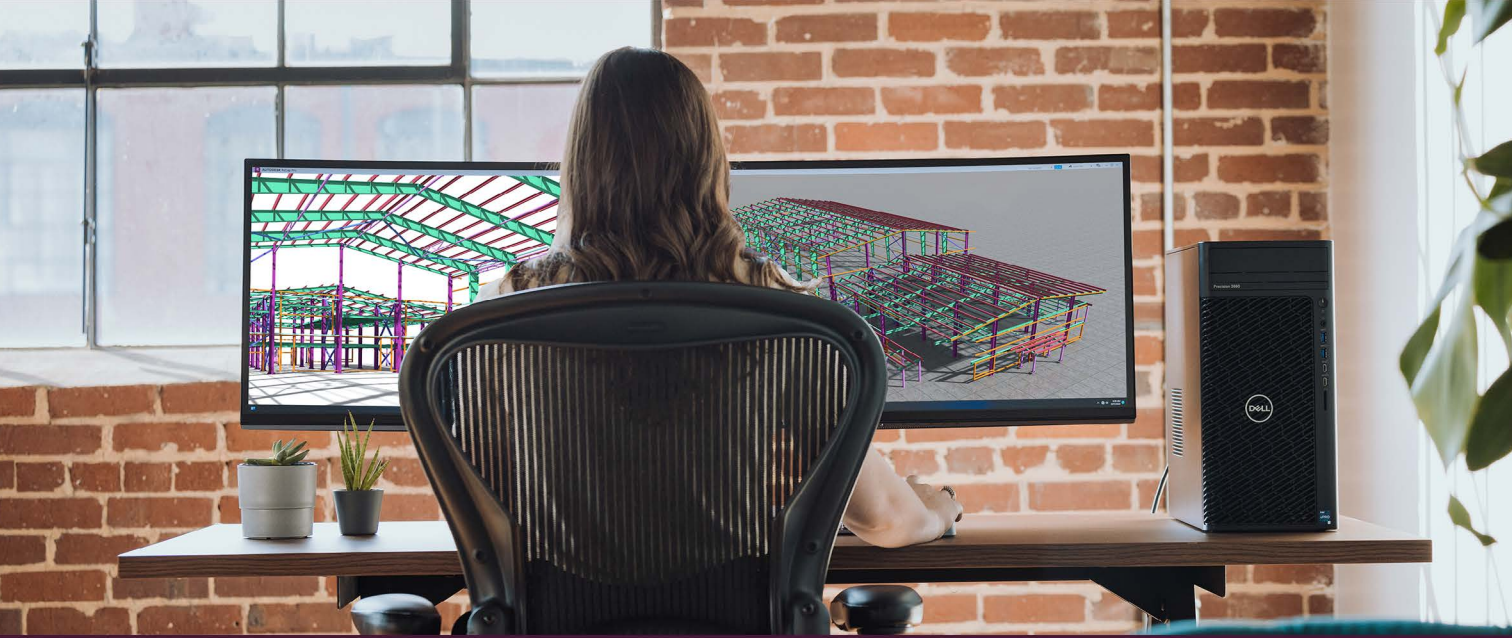
DE247
Digital Engineering



PEERLESS MEDIA



PEERLESS
CONTENT CREATION



AI-enabled technology designed to improve your demanding workflows.

Dell Technologies offers the most powerful and reliable workstations, specifically designed and optimized for demanding workflows in scientific, engineering, and manufacturing fields. These workstations drive the performance of leading professional software applications across industries with advanced AI capabilities.

By partnering with Dell and NVIDIA you will realize unparalleled AI optimization through the Dell AI Factory with NVIDIA, which provides comprehensive solutions harnessing the power of artificial intelligence.

Together, Dell Technologies and NVIDIA can revolutionize your product design, engineering, and manufacturing workflows, significantly enhancing productivity and innovation in your complex projects and business processes.



DELLTechnologies

 **NVIDIA.**

The AI Workstation

At Dell Technologies World, Dell and NVIDIA outline the future of PCs as more companies adopt on-prem AI.



As part of its Dell AI Factory with NVIDIA solution, Dell is offering Accelerator Services to help companies more effectively deploy AI applications with its line of AI-ready workstations.

A lot of attention has been focused on new solutions to enable artificial intelligence (AI) solutions in the data center and in the cloud. Where does that leave the workstation? At [Dell Technologies World 2024](#), held in Las Vegas at the end of May, Dell shifted the focus to the AI PC in an attempt to answer that question.

“The world has capitalized on the industrial revolution and Moore’s Law, and now into the AI revolution,” said Dell founder and CEO Michael Dell during his keynote address. “[It is] powered by an unprecedented magnification of scale that is driving unbelievable GPU acceleration ... for me, every workflow in every enterprise, in every industry, in every corner of the world, will be reinvented with generative AI. That’s why I believe this is a business transformation of a generational magnitude. And just think about this – \$11 trillion in GDP impact will be felt by the global economy in the next three years alone. So this is really big.”

While the conference was in full swing, there was a flurry of activity around the new Microsoft Copilot+ PC, a new class of PCs enhanced for AI. According to a [Microsoft blog](#) announcing the PCs, the new system architecture combines the processing power of the CPU, GPU, and Neural Processing Unit (NPU), large language models running in the Azure Cloud, and small language models (SMLs).

Most of the announcements around Copilot+ were for consumer-grade hardware, and Dell announced [five new Copilot+ PCs](#) as part of its XPS, Latitude and Inspiron lines. Dell Technologies and NVIDIA previously announced [AI-ready hardware in the Dell Precision line](#) of professional engineering workstations powered by the much higher AI processing performance of NVIDIA RTX professional GPUs.

In addition to new hardware, Dell also announced the expansion of the Dell AI Factory with the NVIDIA platform to include new server, edge, workstation, solutions, and services advancements that the company says will speed AI adoption and innovation.

Dell AI Factory with NVIDIA is what the company describes as an “end-to-end AI enterprise solution that integrates Dell’s compute, storage, client device, software and services capabilities with NVIDIA’s AI infrastructure and software suite, all underpinned by a high-speed [Spectrum-X] networking fabric.”

According to NVIDIA CEO Jensen Huang, “A whole new class of data centers will be designed for just one purpose, the purpose of manufacturing

intelligence at scale. There will be a shift to intention-driven computing away from retrieval-based computing that focuses on retrieving existing content and toward these generative AI factories.”

As part of the AI Factory with NVIDIA offering, Dell now provides Accelerator Services that can help companies launch their AI initiatives on AI-ready Precision workstations in a more streamlined fashion. According to this [blog](#) by Bethan Williams, Global Portfolio Lead – Applications and Data Consulting at Dell, this approach “combines retrieval augmented generation (RAG) on a powerful AI mobile workstation running developer workbench software, with new professional services to make it easy for you to take advantage of this innovative approach.”

RAG and the Dell Precision AI-ready workstations with [NVIDIA AI Workbench](#) simplifies these deployments. The Accelerator Services for RAG on Precision AI-ready workstations offering will help further jumpstart GenAI adoption. As part of the service, Dell provides “a ready-to-use mobile lab as a convenient, cost-effective way for customers to explore use cases and improve skills in a low-risk environment. This mobile lab not only enables developers to experiment with and investigate GenAI but also is an ultra-convenient way to demonstrate the effectiveness and outcomes of GenAI.”

Dell consultants set up the lab on Precision mobile workstations, and the service includes installation and configuration of NVIDIA AI Workbench.

Dell and NVIDIA have also announced the availability of [NVIDIA AI Enterprise](#) on select Dell Precision [fixed](#) and [mobile](#) workstation solutions. The software platform is designed to accelerate data science pipelines and streamline the development and deployment of production-grade AI applications.

According to Dell, NVIDIA AI Enterprise is available on Precision workstations with eligible NVIDIA RTX Ada Generation GPUs, including NVIDIA RTX 5000 Ada, [NVIDIA](#)

[RTX 6000 Ada](#), and [NVIDIA A800](#), which have been designed with up to four NVIDIA RTX™ Ada Generation GPUs. This provides the computational power needed for AI development tasks, such as training, fine-tuning, and inferring AI workloads.

Eligible Precision mobile workstations include the [Precision 7680](#), [Precision 7780](#), and [Precision 5690](#) workstations with the RTX 5000 Ada. Eligible Precision fixed workstations include the [Precision 7960 Rack](#), [Precision 7960 Tower](#), [Precision 7875 Tower](#), [Precision 5860 Tower](#) and [Precision 3680 Tower](#) with the RTX 5000 Ada, RTX 6000 Ada, and NVIDIA A800 GPUs.

devices will enable AI adoption.

In one presentation, Sam Grocott, Senior Vice President of Product Marketing at Dell Technologies, emphasized the importance of on-premises and edge devices, where most data is stored. “A lot of people thought everything was going to be in the cloud,” he said. “They couldn’t have been more wrong. We can learn from these cloud sins of the past and not do that with AI.”

Grocott said that companies need to be smarter about figuring out how AI can be deployed, and repeated Michael Dell’s assertion that AI should be brought to the



NVIDIA CEO Jensen Huang and Michael Dell at Dell Technologies World 2024.

The Future of the AI PC

According to Michael Dell, desktop workstations and edge devices will play a critical role in enterprise AI. “You don’t want to put your data in a public service,” he said. “Inference for large language models can be 75% more cost effective on premises versus in the public cloud. And 83% of enterprise CIOs plan to repatriate workflows from the public cloud in 2024 [according to a Barclay’s CIO survey]. This is driven by inference and data gravity. You want to bring AI to your data, not the other way around.”

Where does the AI PC fit in? Several sessions focused on how workstations and on-premise

data where it lives.

Jeff Boudreau, Chief AI Officer at Dell Technologies, emphasized the importance of data driving AI strategy. “If you have bad data, you’re going to have bad AI. If you have no data, you’re going to have no AI,” he said. A robust data strategy, including management practices and governance, is essential for ensuring data quality and accurate insights. He also highlighted three critical areas for organizations to tackle during this process: data, talent, and infrastructure.

You can watch replays of some of the Dell Technologies World sessions [here](#).



Image Courtesy of Koenigsegg

Designed to Innovate. Engineered for Speed.

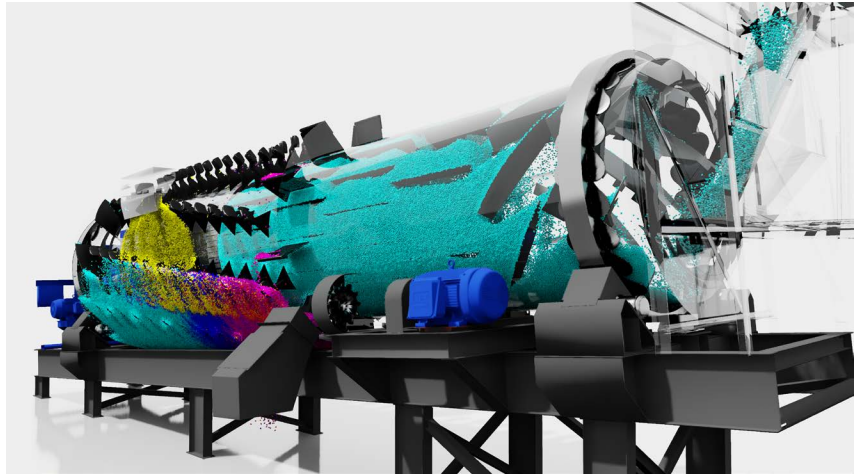
Enhanced with the latest advancements in AI, ray tracing, and physics simulation, NVIDIA RTX™ technology enables incredible 3D designs, and photorealistic visualizations that were once unimaginable.

From grappling with complex simulations to crafting immersive product experiences and more, RTX gives you the power to tackle your most demanding workloads with speed and precision.

Learn how you can accelerate your workflows and bring your innovations to life at [nvidia.com/manufacturing](https://www.nvidia.com/manufacturing)



GPUs Lift Heavy Simulation Load



An Altair® EDEM™ simulation rendered in NVIDIA Omniverse™. Image courtesy of Astec Industries.

Using Altair® EDEM™ and NVIDIA RTX™ GPUs, Astec Industries has improved and accelerated its DEM simulations.

Astec Industries designs and manufactures heavy equipment for the construction and road building industries. Its machines need to crush, move, and mix rocks, and the company relies heavily on fluid, mechanical, and multi-body dynamics simulations, as well as discrete element method (DEM) simulation to model the behavior of gravel, rocks and other particles.

Andrew Hobbs, Director of Advanced Technologies at Astec Digital, spoke to us about how GPU acceleration has helped advance its DEM simulation via its Altair EDEM software, which is part of the Altair® HyperWorks® design and simulation platform. Using NVIDIA's flagship desktop GPU, the [NVIDIA RTX™ A6000](#) GPUs, Astec testing showed a 90x improvement in simulating realistic particle shapes.

Hobbs began his career as a graduate in 2001 with Astec doing FEA and CFD simulations. Since then, the simulation team has grown from one employee to eight. Over more than 20 years, the company gradually expanded its use of simulation to the point that it is now an essential part of the design workflow.

You can learn more about Astec and its DEM efforts in these videos about simulation of [rock crushing](#) and [asphalt mixing](#), as well as a [video](#) interview with Hobbs. You can also read more about NVIDIA RTX GPUs and Altair EDEM here and here.

What are some of the key simulation challenges you face at Astec?

We are a company that makes equipment for every

part of the Rock to Road® process, from getting rocks out of the ground, making them smaller, to moving, sorting, and stockpiling, through to asphalt and concrete plants, all the way to pavers and reclaimers. For asphalt production we design plants that heat and dry rock and mix and coat it with liquid bitumen. These are very complex multi-physics processes with high energy requirements. Simulation provides us with a tool to analyze and optimize our designs for better performance and efficiency.

We started using EDEM, which is now part of Altair, in 2006 in the very early days of commercial DEM codes. Initially the CPU compute power was not there to do everything we wanted to do, but the application was a good fit.

In much of our equipment harsh operating conditions and high temperatures mean direct observation of the internal process is impossible. You can see what is going in and coming out, but changes to internal features have historically been very trial and error. DEM gives us visibility into what is happening inside the equipment, and then the opportunity to very quickly try different design changes to see how they can affect performance. It has been a game changer for equipment design.

How has DEM evolved? How accurately can you predict the behavior of irregular objects like gravel? DEM has traditionally used spheres, correct?

Spheres are computationally a very efficient way to track contacts, and that is why most DEM codes used spheres. By adding more complex physics to the contact models that determine particle behavior it's possible to represent bulk flow of non-spherical materials like rock with sphere or clumped spheres. For simulating the flowability of aggregates, we've found that a properly calibrated material model using clumped spheres is accurate and fast. You can go for hyperrealism in terms of shape, but that comes at a cost.

For applications where interlocking forces are important or there are orientation specific apertures, shape is more of an issue, so you can use polyhedral particles, which Altair supports. There is a computational cost to that, and for irregular shaped materials like rock you still need to decide on a representative shape or set of polyhedral shapes.

As with all simulation, decisions about methodology including how you approach particle shape depends on what question you are trying to answer, the level of accuracy required, and how much time you have to answer it.

What types of software and hardware are you using for simulation, other than Altair?

Most of the team have Dell Precision 7920 workstations with NVIDIA RTX A6000 GPUs for day-to-day simulation projects. The technology both from solver and from the hardware perspective has advanced significantly, and with access to cloud compute instead of just being able to do a handful of simulations to provide an engineering designer with an answer, we can now look at hundreds of simulations. We can run simulations in the cloud to build a reduced order model (ROM), or use machine learning to create an optimization routine to narrow a very wide initial design space and identify an optimum we might not have found through an iterative design process. We do not use the cloud for every problem, but it has changed the landscape of what is possible with DEM.

We still use local compute for most of our work, but large design of experiments containing hundreds of variations can be drastically accelerated with cloud GPU instances of multiple NVIDIA A100 Tensor Core GPUs.

We've developed strong relationships with our CAE vendors over the years. For CFD, FEA, and some multi-body (MBD) simulation we are predominantly using ANSYS tools. If we are coupling DEM with MBD, we use Altair® MotionView® and Altair® MotionSolve®. When doing large scale design of experiments (DOEs), we use Altair HyperStudy to optimize designs and train ROMs. If we are coupling MBD with FEA, or feeding inputs into Ansys Mechanical, we use Ansys Motion. Astec uses SOLIDWORKS for CAD.

How have the more powerful GPUs and workstations affected your DEM workflows?

GPU and DEM are made for each other because of the nature of contact detection algorithms. It has been a game changer in terms of the number of particles we can simulate and the computational speed. When we started out with DEM, we were looking at 25,000 particles in our drum, with an 8-core workstation, which was as good as you could get back then. It would take a day to solve 60 seconds of simulated time. Now in the same amount of time we can simulate 60 seconds of a million particles quite easily with the RTX A6000, and we can extend the physics models to include liquid coating and cohesion, heat transfer, and couple to CFD and MBD.

Currently we are using a single GPU in the workstations, but the next step for us is probably going to be looking at multi-GPU workstations. As we look ahead towards real-time digital twins, using simulation to supply the data to train models will require us to run hundreds of simulations.

What are your plans for artificial intelligence and digital twins?

We are still in the process of understanding where AI can help us. There are great use cases for machine learning and optimization studies and using simulation to create synthetic data. We've already seen the benefits of this approach. Simulation gives us a completely controlled, virtual environment to try out lots of things.

Astec Digital's cloud platform that connects Astec equipment will enable more data visibility and a step closer to real-time digital twins. As more data is available from physical equipment, we can compare the physical systems performance to the predicted performance from a simulation produced ROM, and then close that loop to provide better inputs to our simulations from physical sensors. We are making good progress in this area and the next few years will be really exciting. The technology is there, but we have some work to connect equipment data to simulation and machine learning models to deliver real value to our customers.

I don't believe AI will replace simulation anytime soon. Most of our applications are very niche and not easy to generalize. Any kind of AI model we use is something we will have to train ourselves. We are looking at that and exploring where it would make sense to do this.

We have been using NVIDIA Omniverse, initially as a post-processing tool. Simulation is great as a design tool, but also valuable for sales and marketing. Our equipment is a black box, and being able to show the advantages of our design using simulation results and nice post processing is helpful.

We hope NVIDIA Omniverse will be the 3D platform we use for the digital twins connecting CAD, simulation, IoT data, and machine learning. We are using NVIDIA Omniverse and NVIDIA Isaac Sim™ to produce synthetic data to train computer vision models. NVIDIA Omniverse excels at creating realistic virtual worlds. Combining this with the physics models in our simulations and IoT sensor data will be a very powerful tool to build intelligent and autonomous controls to improve the performance and sustainability of our equipment.

CAM Embraces GPU Acceleration

New CAM software releases leverage NVIDIA RTX GPU power to provide faster results.



GPU acceleration can enable up to 10x faster material removal simulations. Image courtesy of NVIDIA.

Engineers have benefited from the use of GPU acceleration across a number of workflows, including CAD, rendering/visualization, and engineering simulation/analysis. At the IMTS 2024 conference in Chicago in September, the benefits of GPU were on display with a new group of software vendors – those in the CAM sector.

Both ModuleWorks and MachineWorks have announced GPU acceleration within their computer aided manufacturing solutions this year, and both were on hand at the IMTS show to discuss how GPUs (specifically, [NVIDIA RTX GPUs](#)) can save time and boost productivity.

ModuleWorks [announced](#) GPU acceleration features in August, which is expected to be made available for testing with the ModuleWorks 2024.12 software release. The new feature will be offered as part of the ModuleWorks Cutting Simulation software and can be integrated directly in existing CAM systems, enabling solution providers to upgrade to GPU-accelerated material removal simulation while retaining the full feature-set of the [ModuleWorks](#) simulation products.

In benchmark tests conducted at the ModuleWorks facilities in Aachen, Germany, the ModuleWorks Cutting Simulation with GPU processing was up to 10x faster than CPU-based simulations on mid range NVIDIA graphics cards, whereas more than 20 times acceleration has been observed with high-end

NVIDIA graphic cards.

According to ModuleWorks Founder & Managing Director, Dr. Yavuz Murtezaoglu, the company is responding to the need for faster simulations that were constrained by the capabilities of the CPU. Rather than porting the CPU version of the software to support GPU compute, the company took the most time-consuming part of the simulation and developed new software code to natively calculate that on the GPU using OpenGL commands.

“We realized we could benefit from heavily optimized OpenGL [graphics standard] drivers and hardware from companies like NVIDIA, to convert our problems in a way that they can work on the GPU,” Murtezaoglu said. “We get the same accuracy, quality, and all of the features we have developed over 15 years, so we do not sacrifice any functionality, but we can run the simulations faster.”

Customers can enable the new features without any complex retrofitting; they just need to pay an additional license fee and make sure they have the right GPU.

ModuleWorks conducted initial testing with an NVIDIA GeForce GPU on a desktop system, and has plans to perform benchmarking on the professional-grade NVIDIA RTX GPUs. Both ModuleWorks and NVIDIA have facilities in Aachen, Germany, which has facilitated close collaboration.



ModuleWorks Founder & Managing Director, Dr. Yavuz Murtezaoglu, at the IMTS conference in Chicago.

“NVIDIA supports OpenGL and works out of the box with our software,” Murtezaoglu said. “The NVIDIA cards scale extremely well, and we expect additional boosts in performance as we move up to more powerful GPUs.”

NVIDIA continues to work with its OEM partners to help enable this type of performance on engineering workstations. “Combining professional NVIDIA RTX graphics with purpose-built Precision workstations from Dell demonstrates real-world customer performance benefits as ISVs like ModuleWorks and MachineWorks create new functionality leveraging GPUs to accelerate workflows.” says Matt Allard, Director of Strategic Alliances at Dell Technologies. [MachineWorks](#) announced [MachineWorks GPU](#) in early September, a new

software release that takes advantage of the high numbers of cores in modern GPUs to accelerate overall simulation performance. It is targeted for release in 2025 as part of MachineWorks 9.0.

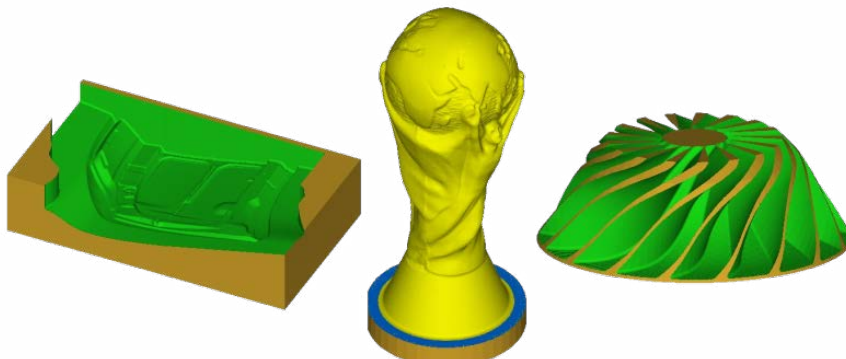
“Having developed early prototypes more than a decade ago, we have been watching the development of GPU hardware closely,” said Dr. Fenqiang Lin, managing director of MachineWorks, in a press release. “GPU has become more mature and, with advances in AI LLMs [artificial intelligence large language models], it is clear that massively parallel hardware will become a commodity item.”

According to Richard Baxter, head of sales and marketing for MachineWorks, the company has adopted GPU-enabled CNC simulation based on customer demand and because there is a clear benefit in speed. The software uses simulation to produce a stock model that customers can then utilize for the toolpath generation process, and Baxter says that the company has seen between a 10X and 25X speedup (depending on which GPU is being used) of those processes compared to using the CPU alone.

“We can see a reduction from several minutes down to just a few seconds, subject to the constraints of accuracy,” Baxter said.

The company has transitioned one of its geometry engines to GPU compute and has benchmarked performance on the NVIDIA RTX™ 2000 Ada Generation GPU as well as NVIDIA GeForce RTX products.

At IMTS, there was also significant activity around the use of artificial intelligence (AI) in the CAM and CNC programming space, and Murtezaoglu at ModuleWorks says he hopes that the work NVIDIA is doing in both AI and engineering will benefit the CAM software market. “We hope that it will drive technology even faster and further because of the big investments being made in AI, and it will raise the bar even higher,” he said.



Examples of MachineWorks GPU simulation results. Image courtesy of MachineWorks.

AI Model Integrity

Ensuring useful AI-based simulation results requires robust data curation and training.



Artificial intelligence is rapidly expanding into engineering and simulation applications, with software tools from vendors like Ansys and Altair incorporating the technology to provide rapid simulation and automation capabilities. The use of AI and machine learning in simulation presents opportunities, but many end users are concerned about the risks of the technology.

In more mainstream use cases of AI, such as ChatGPT-type systems that leverage large language models (LLMs), the technology can potentially hallucinate — in other words, it can generate results that are false or, in some cases, absurd. Another potential pitfall is the concept of model collapse, which can occur if an AI model ingests a significant amount of AI-generated data. This can also degrade the performance of the model.

For AI-based simulation use cases, model reliability is paramount. Unlike consumer-style AI applications, engineering organizations will leverage internal (usually proprietary) physical test and simulation data to train the models, often informed by physics, which provides an advantage when it comes to model integrity. To find out more about data and model integrity in this context, we spoke to Erick Galinkin, AI Security Researcher at NVIDIA.

What are the primary differences between the data considerations in a ChatGPT-type environment, versus a more enterprise-specific use case like simulation?

Erick Galinkin: I think the big difference is that with ChatGPT or other models you might see in academia or at big labs, they are typically dealing

with pretty clean, well-understood data formats. I come from a cybersecurity background, so I have a lot of experience working with data that is not well-formatted or heterogeneously formatted. In industrial applications, we are often dealing with outputs of sensors, and when you are dealing with that, getting it into a format that is reliable for training a model can be challenging.

How do firms ensure data integrity for AI-based simulation and other business-focused applications?

You want the data to be as close to its original form as possible, which creates a fundamental tension with the fact that the models want the data to be in a certain shape. LLMs want flat text. Convolutional neural networks want a rectangle. It can be hard to get data into those shapes without inherently destroying some of the information that is encoded in data's structure.

Depending on the specific application, you have to try and find the right balance. Keep it as close as possible to the original data while putting it in a form that can be used by the model architecture you have decided is right for the problem. Sometimes that means picking a different model architecture than you started with.

What are the risks of AI model collapse or hallucination, and how do you mitigate against that?

I will tackle those two separately because they have different causes and outcomes.

Hallucination is endemic to LLMs, because they are just next-token generators. They are trying to predict the next token in a sequence. Once a token is generated, it is not going to look back for consistency. You can end up with cases where the model decides the next best token is just not correct, and it will keep generating tokens based on that incorrect assumption.

To prevent hallucination, you can use things like guardrails to do consistency checks, or you can check if the output is contained in the source documents. You can take steps to mitigate it, but the existence of hallucination is baked into this type of model.

Model collapse is something that the AI research community has struggled with for a really long time. If you change the distribution the model is training on, it can forget the stuff it used to be good at. You can think about it in terms of bell curves. If you are rolling two fair dice, the most likely outcome is seven. If you make those dice weighted and update those weights after rolling them repeatedly then eventually, you will get enough sevens that you only ever roll a seven.

You pull from the bell curve and update the curve with what you just sampled, and as you do that more and more tails disappear and it will collapse to the mean.

Ensuring enough diversity in the data is really crucial, because ultimately what you want is heterogeneous variance that still kind of reflects what the real world should look like.

In simulation, you have an advantage. If you are simulating physical systems, there is a lot of existing research in physics we can use. So within some error boundaries, you can make sure the synthetic data and simulated data do not look crazy to the model or in the physical world. If you have a differential equation, a heuristic, that can be a good bounding function.

We have seen use cases for synthetic data in engineering. How do you ensure integrity of AI-generated synthetic data in these scenarios?

Like any tool, you need to use it when it is appropriate, but not over use it. AI can be useful

for data that is difficult to collect. It is better to have real-world data, because you cannot perfectly simulate it, but it is still worth augmenting the data if you are lacking certain cases.

You should keep an eye on how the overall model performs on real-world data, as a kind of subset. That is something that I have done. I use my real-world data, not as a hold-out set, but as a sanity check to compare to the performance on synthetic data.

Does NVIDIA offer tools or services that help with model integrity?

I think that the NVIDIA NeMo is an excellent platform for building custom generative AI models with integrity. It offers NeMo Guardrails, which allows developers to add programmable guardrails, ensuring trustworthiness, safety, security, and controlled dialog while protecting against common LLM vulnerabilities

Model collapse is harder, because it is more domain specific. It is more of a data curation problem, where you want to make sure you have a validation set that is representative and may contain some outliers. Using NeMo Curator, you can efficiently handle such data curation challenges by ensuring that your dataset is diverse and representative. It offers various out-of-the-box functionalities such as deduplication, text cleaning, quality filtering, privacy filtering, and domain and toxicity classification.

No model is ever going to be perfect. One of the first things you learn when training models is that if you ever have a perfect score, you probably did something wrong. You are going to miss things. You need that validation data as a sanity check. NeMo offers an end-to-end platform for building custom generative AI, allowing you to continuously customize and evaluate the models to find the best one for your needs.

Conclusion

Engineering organizations are just beginning to grapple with the promise and challenges of adding AI to their simulation toolkits, and data management and curation will play a large role in the success of those efforts. NVIDIA and Dell have partnered to create the [Dell AI Factory with NVIDIA](#), to assist enterprises with implementing AI in their workflows. NVIDIA also offers a number of AI tools to assist developers, which you can learn about [here](#).



NAFEMS CEO
Tim Morris
welcomes
attendees to
Louisville.

AI Meets Simulation at NAFEMS Americas

Technology leaders provide guidance on how artificial intelligence can enhance simulation workflows.

At the 2024 [NAFEMS Americas](#) conference in Louisville, KY in July, simulation users, software providers, integrators, and academics converged to discuss the latest developments in engineering simulation.

NAFEMS is the international association for engineering modeling analysis and simulation, with chapters around the globe. To kick off the conference, Tim Morris, NAFEMS CEO, outlined the agenda for the three-day event, and highlighted some of the major developments in the industry – with artificial intelligence (AI) at the top of the list.

“AI is particularly relevant to our world in simulation,” he said. “We are learning a lot about how it can be applied in our environments.” Other notable developments included the growth of systems simulation and model-based systems engineering; increased access to high-performance computing; CAE vendor consolidation; and simulation certification efforts.

But AI came up again and again across a number of sessions. Keynote speaker Anthony Petrella, director of the Computational Biomechanics Group at the Colorado School of MINES, was asked about AI in the context of discussion of the future of engineering education.

“Using AI is something that every student is doing,” he said. “I think we should avoid viewing it as a threat. We should view it as a tool that we should leverage to its maximum extent. We are all still mid-stream in discovering the right ways to do that.”

Accelerated Computing Enables AI-based Simulation

Ian Pegler, Global Business Development Manager for CAE at NVIDIA, led a standing-room-only session on Demystifying AI for CAE, during which he explained basic AI concepts.

According to Pegler, AI can potentially help solve some of the speed and computational demand challenges faced by companies that want to expand their use of simulation. “Companies want to use simulation to work faster and make better products,” he said. “They want to run more simulations and do it a lot quicker, but with limited resources. We can use AI to get faster feedback early in the design process.”

In most of the emerging products offered by simulation software vendors, AI models are trained on existing test and simulation data to help provide faster analysis of a high number of design iterations to help engineers narrow their options before running more comprehensive simulation studies. The use of NVIDIA RTX™ GPUs provides the speed and computational horsepower to do this efficiently, both at the workstation level and in the data center.

During the presentation, Pegler explained the history and structure of the neural networks that power AI solutions, as well as new frameworks like TensorFlow and PyTorch that help developers build models without having to code everything from scratch. NVIDIA also offers the [Modulus](#) framework, which is an open-source library that helps companies build AI models.

He also explained the role of the GPU in accelerating AI adoption because of their ability to rapidly perform computations like mixed-precision matrix multiplication that are critical for AI applications. The fact that NVIDIA GPUs excel at such mixed-precision operations, with stronger integer and floating point performance, makes the GPU an ideal choice for these applications. Combined with the NVIDIA software toolchain and features for training and developing AI models, this puts the NVIDIA GPUs at an advantage, compared to AI PCs that rely on neural processing units (NPUs),

NVIDIA was also part of a workshop, led by Mustafa Kaddoura, Senior Application Engineer at Ansys, titled Leveraging AI to Boost Computation-Intensive Simulations.

The workshop explored how AI is applied to engineering simulations across industries, and the

challenges posed by high computational demands. The workshop also included a CFD demonstration using Ansys SimAI, the company's new cloud-based generative AI platform.

Kaddoura said that the AI models can be efficiently trained on NVIDIA RTX GPUs with large datasets to provide faster insights to engineers.

"AI can be very powerful for repeated simulations where there are lots of design iterations," Kaddoura said.

However, because the application of AI in these workflows is relatively new, most customers will need to investigate the size of the data set needed to train these models for their specific use cases. Kaddoura said Ansys recommends testing the model against an actual Design of Experiments (DOE) to ensure accuracy.

Earlier this year at its GTC event, NVIDIA announced its new Blackwell chip for AI workloads. The company was also involved in the [Dell Technologies World 2024 event](#), where Dell and NVIDIA discussed AI-enabling hardware innovations. NVIDIA RTX GPUs are also part of the Dell AI-ready Precision engineering workstations. Dell and NVIDIA have also collaborated on the [Dell AI Factory with NVIDIA](#) solution for enterprise AI deployments.

Ford also presented a use case that leverages AI for creating surrogate models for CFD simulations during the GTC conference. You can watch a replay [here](#).

For more information on GPU-accelerated CAE, download the AI-Powered Engineering white paper, and check out the NVIDIA [Accelerating and Advancing CAE E-Book Series](#).



NVIDIA's Ian Pegler explains the use of AI for simulation at NAFEMS Americas

AU 2024:

Autodesk Offers Glimpses of the Future with Project Bernini

New Proof of Concept Hints at AI Training Based on Proprietary Data



The giant AI icon on the exhibit floor offered hints on the focus of the conference.

In October 2024, in the sunlit Convention Center a few blocks away from San Diego’s popular Gaslamp Quarter, Autodesk CEO Andrew Anagnost delivered the keynote to Autodesk University, the annual Autodesk user gathering. The event drew an estimated 12,000 in-person guests, according to Autodesk PR.

Considering the AI-related acquisitions Autodesk has recently made (such as that of Blank AI) and the AI-powered features it’s been touting, the year’s AU keynote is no surprise to anyone. “It’s still really early, despite what a lot of technology companies are saying. The great sorting out ... is just beginning,” said Anagnost.

Bernini Paves the Way

For Autodesk, [Project Bernini](#) exemplifies how design programs like Autodesk Fusion might incorporate AI into the conceptual phase. The proof-of-concept technology developed by Autodesk Research allows users to generate 3D objects using text prompts, hand-drawn sketches, and point clouds. The application replicates in 3D the easy content-generating method pioneered by Generative AI text-to-image models like Midjourney and Dall-E, which are trained on NVIDIA GPUs.

Raji Arasu, Executive Vice President, Chief

Technology Officer, Autodesk, revealed a few caveats with the preview technology: “First, Bernini is trained on publicly licensed data, so it’s not available for commercial use. Second, it’s open to the AI community, so we can make it smarter together.”

The Autodesk Assistant is also evolving, bolstered by AI- and language-driven features. Arasu said it “gives timely insights, whether in our docs, autodesk.com, or in our products. It works as a system that understands you, your industry, and the context of the product.”

Later, during his meeting with the industry press, Anagnost said, “AI is going to move faster for manufacturing. First reason is, we’ve been in the cloud longer ... The other thing is, we think AI has a lot of value to add to the manufacturing space ... The manufacturing space in general has shown a great willingness to adopt digital technologies, if they provide real benefit.”

AI Transparency

Arasu revealed Autodesk plans to introduce an AI transparency card, similar to the nutrition and ingredient cards that accompany grocery items. “We think it’s a great way to show at least the bare minimum of info people need to know about the AI used in our products,” said Arasu.

During his meeting with the industry press, Anagnost said, “The training methods [used on Bernini] are data-independent, so we could go to customers A, B or C, and say, would you like to participate in helping us make this model better? There’ll be an evolution as we start partnering with certain customers to train the model on their data.”

This business model opens the possibility that many Autodesk software users in automotive, aerospace, and consumer goods may further refine the foundational model from Autodesk with their own proprietary data and industry-

specific rules to develop variants of Bernini capable of generating designs that conform to a user’s brand aesthetics and compliance requirements. Such workflows benefit from NVIDIA RTX™ GPUs with Tensor cores for AI-powered applications found in Dell Precision workstations.

Hardware for AI

On the exhibit floor, Himanshu Iyer, Senior Industry Manager, Manufacturing, NVIDIA, showcased a new offering from the stock-art merchant Shutterstock. “Imagine you’re creating a virtual photoshoot taking place in an office environment, and you need to populate it with tables, chairs, computers, and typical office furniture. Shutterstock, with their millions of licensed images, has launched a [3D AI Generator based on NVIDIA Edify](#) where you can create the 3D models you need just by entering text or image prompts,” he said.

NVIDIA Edify is described as “a multimodal architecture for developing visual generative AI models for image, 3D, 360 HDRi, physically based rendering (PBR) materials.”

At the conference, Dell showcased its lineup of AI-targeted Dell Precision workstations. “It’s a combination of workflows. It’s never just one system,” said Ken Flannigan, AECO Alliances and Solutions, Dell. “We have from laptops to towers to rugged systems, with neuro-processing units onboard the chipset.”

For AI developers, Dell recommends the Dell Precision 5860 Tower workstation with NVIDIA

RTX 6000 Ada Generation GPU; and the Dell Precision 7000 Series mobile workstations with NVIDIA RTX 5000 Ada Generation Laptop GPU.

For data scientists working in AI, Dell recommends the Dell Precision 7680 mobile workstation with up to NVIDIA RTX 5000 Ada Generation Laptop GPU; and the Dell Precision 7780

mobile workstation with NVIDIA RTX 5000 Ada Generation Laptop GPU.

Autodesk to Join the LA Olympics 28

The AU keynote wrapped up with the announcement that Autodesk has been selected as the Design and Make Platform for the LA28 Olympic and Paralympic Games. The company revealed, “A hallmark of the city, the LA Memorial Coliseum will be the first stadium ever to host three Olympic Games. Autodesk’s Design and Make Platform will be used to reimagine the iconic stadium for the LA28 Games’ Track and Field events, while preserving its past, present, and future traditions.”

Anagnost said, “I can confidently say that LA will be ready. They are reinventing how the city can prepare for this moment and taking a completely unique approach to change—one that has never been done before.”



Autodesk CEO Andrew Anagnost delivers the keynote to AU 2024.



At Siemens Realize Live 2024, Siemens and Sony jointly introduced NX Immersive Design on Sony Head-Mounted Display. Image courtesy of Siemens.

Siemens and Sony Showcase NX for VR at Siemens Realize Live 2024

Cloud-Hosted NX X, Partnerships with NVIDIA
and Microsoft, and More

In May 2024, Siemens Digital Industries Software (Siemens) held its annual user conference, Siemens Realize Live, at the convention center in Mandalay Bay (Las Vegas, NV). During the keynote Tony Hemmelgarn, Siemens's President and CEO, revealed his vision: comprehensive digital twins in the Metaverse. "We have to change the way people work, the different ways of collaborating ... We're talking about the industrial metaverse," he said.

To realize this vision, Siemens had been aligning the necessary building blocks: partnerships with NVIDIA, Microsoft, and Sony, among others.

Historically, design and simulation software programs are not known for ease of use, but Hemmelgarn believes AI is about to change that. "Things like constraints, loads, forces, pressures, and temperatures can be really complex. If you don't put these constraints right, it's garbage in, garbage out, so you'll have a problem. With AI, you can enhance this process," he said. With an injection of natural language processing and reduced order models, Hemmelgarn believed, "We can start putting [these tools] in the hands of the generalist engineer."

Like many of its competitors, Siemens is now beginning to view AI as part of its strategy to expand its

user base. AI-powered features have become part of its Xcelerator products: for example, Teamcenter Assistant, NX Generative Design, Solid Edge Generative Design, Simcenter Reduced Order Modeling, and more. The company promises to deliver Copilot for Siemens Xcelerator, Teamcenter Semantic Search, NX Material Prediction, Solid Edge Automated Drawings, and more in the near future.

Sony and Siemens to Bring NX into VR

One critical piece in Siemens' Metaverse strategy is Sony. At Siemens Realize Live, the two companies revealed what they'd been working on: [NX Immersive Design](#), delivered through the Sony Head Mounted Display (HMD). During the keynote, Sony's Deputy

GM for XR, Henry Kondo, showcased the XR headset.

“Our head mounted display has a stunning 8K video fidelity—4K per eye. Also, you can wear it in comfort all day. It’s armed with this unique controller that provides precise control, to interact with digital objects, almost like you are in the real world,” said Kondo. “This tight integration of software and hardware is a new thing in digital twins.”

Many VR applications allow you to view and collaborate around life-size digital twins, but in the Siemens-Sony implementation, the NX Immersive Designer software allows you to launch NX, access its design tools, and modify the design in VR. In the [press announcement](#), Siemens said, “Sony’s XR head-mounted display is a fundamental part of the forthcoming NX Immersive Designer, an integrated solution that combines Siemens’ industry leading product engineering software, NX, exclusively with Sony’s breakthrough XR technology to deliver immersive design and collaborative product engineering capabilities.”

At the event, Seiya Amatatsu, head of Sony’s XR Technology Development Division, described how his team relied on Siemens’ NX software to help develop its new HMD. The SRH-S1 includes an XR head-mounted display with 4K OLED Microdisplays and controllers for intuitive interaction with 3D objects. Sony has been using the Siemens Xcelerator portfolio as its primary product engineering software since 2015.

“Sony succeeded in implementing NX in the fastest schedule in the world and we were able to improve our design process productivity by 25%. Although significant advancements have been made with NX, I wanted to achieve further evolution with a product that brings innovations allowing direct editing in an immersive environment. I believed that could be realized with Siemens and as a result, the XR head-mounted display was developed that could bring innovative products to market more rapidly,” Amatatsu said.

Siemens’ NX Immersive Designer is expected to launch at the end of 2024.

“The tight integration of the hardware and software gives designers and engineers all new, incredibly natural ways to interact with the Digital Twin,” said Hirohito Kondo, deputy general manager, XR Business Development Division, Product Management Department of Sony. “This is about doing meaningful hands-on engineering. That is why the image quality is so important—as well as the precision of the controls and even the comfort of the headset—because together, it lets you collaborate more, engineer better, and innovate faster. It lets you do real, meaningful engineering—

and without ever having to build a physical prototype.”

Hardware to Power NX Immersive Design

For real-time interactive experience with NX Immersive Design, the Sony HMD can be powered by a professional grade GPU-accelerated workstation. In the solution center at Siemens Realize Live, users had the chance to try out NX Immersive Design, powered by a workstation with dual [NVIDIA RTX 6000 Ada](#) GPUs. Hardware partners such as Dell offer workstations certified to run Siemens NX and other packages from the Siemens portfolio.

Collaboration With Microsoft

During the conference, Siemens and Microsoft announced that, “In response to growing customer demand, the Siemens Xcelerator as a Service portfolio will become available on Microsoft Azure, beginning with Siemens’ Teamcenter X software for Product Lifecycle Management (PLM).” Previously the two companies worked together to



Siemens and Sony partner to bring NX Immersive Designer, including Sony’s XR head-mounted display (HMD), SRH-S1, to market later this year. Image courtesy of Siemens.

enable field technicians to report issues in natural language, via the [Teamcenter app](#) in Microsoft Teams.

According to the latest announcement from Siemens Realize Live, “Siemens is also developing a Copilot for Microsoft 365 plugin for the Siemens Teamcenter app on Teams ... They will be able to ask Microsoft Copilot for help with summarizing outstanding tasks and workflows out of the comfort of their Teams app.”

Hemmelgarn said, “Together, we’re going to make sure that our AI capabilities are coupled together with PLM on Microsoft Azure, and as a start, we’re going to have Siemens Xcelerator as a Service run on Azure to give our customers a choice. It’s so important because now we can integrate what we do with the data we have in Teamcenter with Microsoft’s AI.”