

ROSEN CENTER

FOR ADVANCED
COMPUTING



20
24

HIGH-PERFORMANCE COMPUTING AT THE HIGHEST PROVEN VALUE

2024

ANNUAL REPORT



Rosen Center for
Advanced Computing

Director's Welcome



We've wrapped up the 2024 fiscal year here at Purdue with another long list of achievements enabling science for researchers at Purdue and nationally. In these pages you will hear the stories behind the impact of the campus cyberinfrastructure to researchers, students, and society.

- Anvil has completed its second year of production operations, and with an additional \$5M supplemental funding from the National Science Foundation has joined the National AI Research Resource (NAIRR) as a resource provider.
- Our Research Software Engineering (RSE) efforts throughout the organization are unified into a new Center for Research Software Engineering, which will provide a single front door, and smoother processes for RSE services.
- Resources for AI continued to grow, with significant expansions of GPUs and a new parallel filesystem in the Gilbreth system, which is optimized for throughput AI.
- RCAC PIs are leaders on several research awards supporting CI professionals, campus networking enhancements, digital forestry, and sustainability.

Most importantly, while it's the machines and facilities that get all the press, it's the people who truly provide transformative impact. Powerful resources are important, but it is our research scientists, research software engineers (RSEs), and domain expert computational scientists who facilitate the use of the cyberinfrastructure who really make the difference.

Programs like coffee hour consultations make it easy for graduate students to work with domain experts, and RCAC scientists are always available to help you benchmark for a proposal, meet a deadline, provide training for your students, or work on text for that upcoming grant submission.

I'm looking forward to the initiatives that we have already in flight for the next year, where you can expect to see the results of continued investment in support of AI, protected data, and more.

Thank you again for your continued support and making us your provider of choice.

Preston Smith, PhD
Executive Director, Rosen Center for Advanced Computing
Purdue IT

TABLE OF CONTENTS

05	Anvil Science Highlights
15	Envision Center
23	Research Software Engineering (RSE)
25	Science Highlights
33	Partner Awards and Successes
35	External Funding
37	Education and Outreach
39	Student News
43	Staff News
47	Purdue's Women in High-Performance Computing

Mission & Vision

Mission

Purdue IT provides the technology infrastructure, services, solutions, and information security that support teaching and learning, enhance research, and enable faculty and staff to achieve their objectives while providing a positive student experience.

Vision

To empower giant leaps across Purdue by providing safe, efficient and reliable services in our pursuit to become the benchmark for IT in higher education.

"Since 2000, RCAC has accounted for \$118M of sponsored research awards either as lead department or as a participant."



ANVIL ENTERS YEAR THREE OF PRODUCTION

Anvil, Purdue's most powerful supercomputer, continues its pursuit of excellence in HPC as it enters its third year of operations. Funded by a \$10 million acquisition grant from the National Science Foundation (NSF), Anvil began early user operations in November 2021 and entered production operations in February 2022. After two years online, Anvil has more than proven its value. The supercomputer has been used to help nearly 6,000 researchers push the boundaries of scientific exploration in a variety of fields, including artificial intelligence, astrophysics, climatology, and nanotechnology.

Anvil at a Glance— Two Years of Operations

Over the past two years, Anvil has had a significant impact on scientific research and student development. With nearly 6,000 total users thus far, of which over 3,000 were undergraduate students, Anvil is not only helping meet the growing need for high-performance computing (HPC) within the realms of research, but also actively assisting with the development of cyberinfrastructure professionals of tomorrow. Overall, Anvil has allowed users access to 1.006 billion CPU hours and 1.196 million GPU hours, supporting research across 62 diverse scientific domains. In 2023 alone, 72 research publications cited Anvil usage. Aside from the supercomputer itself, the Anvil team has been hard at work promoting the benefits of HPC and ensuring the nation has a workforce trained in the use, operation, and support of advanced cyberinfrastructure. In the two years of operations, the Anvil team has participated in 20 outreach events and conducted 24 training sessions, with more on the horizon. These training sessions are designed to deliver working knowledge of HPC systems and teach users how to get the most out of their research time on Anvil. The team also provided hands-on training to students through initiatives such as the Anvil Summer REU program, which allowed the students to gain much-needed knowledge and experience in the field of HPC.



Anvil Tech Specs

Anvil is a supercomputer deployed by Purdue's Rosen Center for Advanced Computing (RCAC) in partnership with Dell and AMD. The system was created to significantly increase the computing capacity available to users of the NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS), a program that serves tens of thousands of researchers across the United States. The system consists of 1,000 Dell compute nodes, each with two 64-core third-generation AMD EPYC processors, and will deliver over 1 billion CPU core hours to ACCESS every year. Anvil's nodes are interconnected with 100 Gbps Nvidia Quantum HDR InfiniBand. The supercomputer ecosystem also includes 32 large memory nodes, with 1 TB of RAM per node, and 16 GPU nodes, each with four NVIDIA A100 Tensor Core GPUs. These GPU nodes provide an additional 1.5 PF of single-precision performance to support machine learning and artificial intelligence applications.

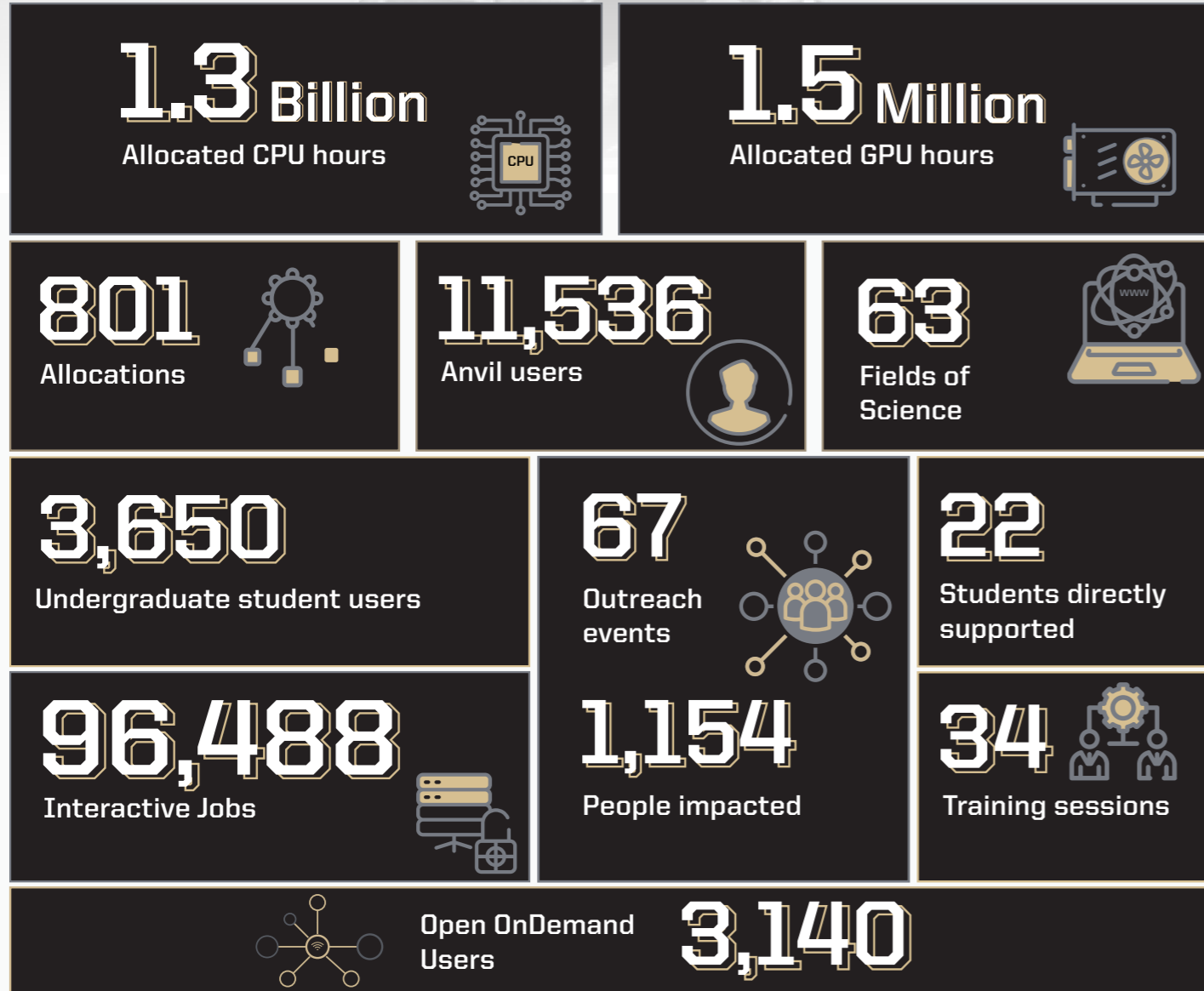
In 2023, GPU capabilities were added to the Anvil Composable Subsystem of the Anvil supercomputer. The Anvil Composable Subsystem hosts eight composable nodes, each with 64 cores and 512 GB of RAM, and a composable GPU node, with 4 NVIDIA A100 80GB GPUs. The Anvil Composable Subsystem is a Kubernetes-based private cloud managed with Rancher that provides a platform for creating composable infrastructure on demand. This cloud-style flexibility allows researchers to self-deploy and manage persistent services to complement HPC workflows and run container-based data analysis tools and applications. The composable subsystem is intended for non-traditional workloads, such as science gateways and databases, and the addition of the composable GPU node supports tasks such as AI inference services and model hosting.



ANVIL

FORGING THE FUTURE OF COMPUTING

1/1/2022 - 8/31/2024



Enabling science through advanced computing

Because of its configuration, Anvil is able to reach a peak processing speed of 5.3 petaFLOPS, making it one of the most powerful academic supercomputers in the US. When it debuted, the Anvil supercomputer was listed as number 143 on the Top500 list of the world's most powerful supercomputers. This advanced processing speed and power has allowed researchers to save hours of time on computations and simulations, enabling innovative scientific research and discovery. The highlights given below are but a few of the hundreds of use-cases stemming from Anvil:

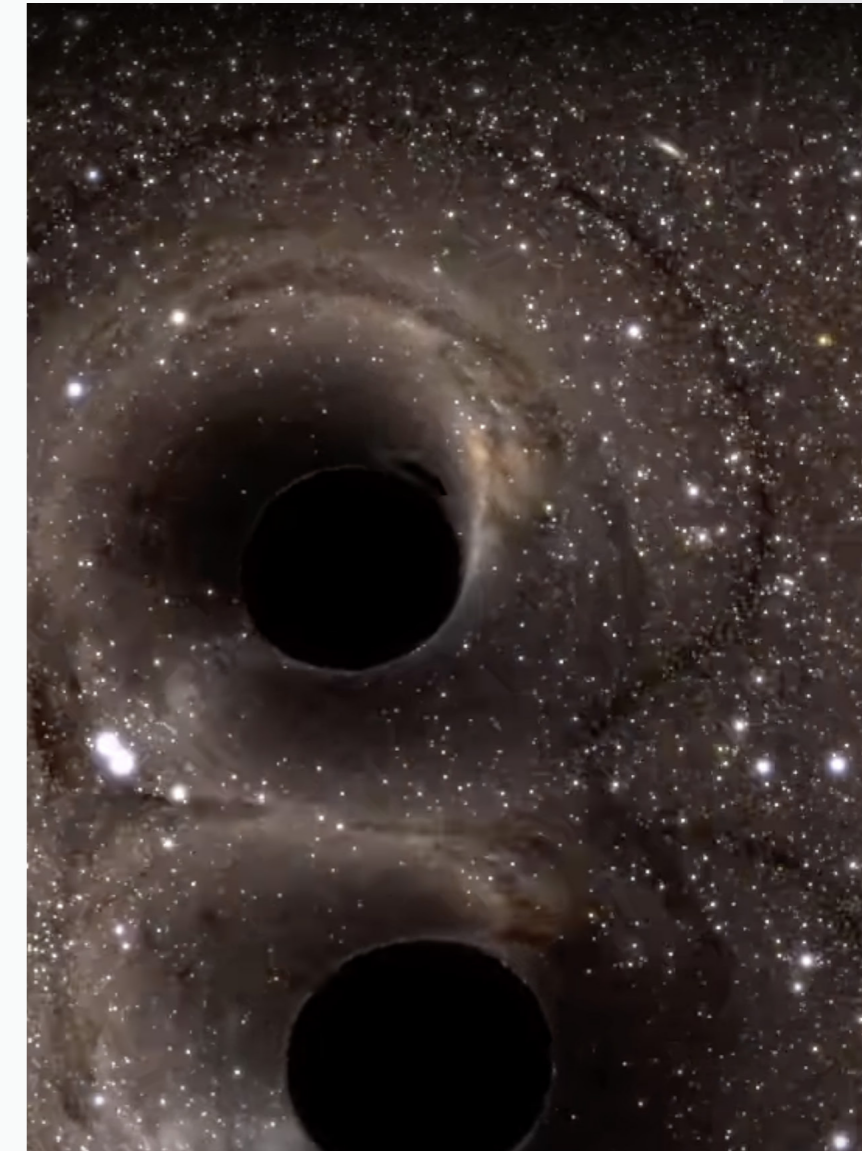
1) Scientists from the collaborative Simulating eXtreme Spacetimes (SXS) research group used (and still use) Purdue's Anvil supercomputer to explore the physics of cataclysmic space-time events and help shed light on the nature of one of the Universe's fundamental forces: gravity. Vijay Varma, Assistant Professor in the Department of Mathematics at the University of Massachusetts, Dartmouth, and Nils Deppe, Assistant Professor of Physics at Cornell University, are computational astrophysicists. The two use Anvil to help develop, test, and run state-of-the-art numerical relativity codes that make high-accuracy gravitational wave predictions for Laser Interferometer Gravitational-Wave Observatory (LIGO) signals. The Anvil supercomputer is indispensable to them for their research, and the group is pleased not only with the performance of the computer, but also with the support they received from the Anvil staff:

"The team at Anvil has been absolutely terrific in working with us to resolve any issues we had when we started," says Deppe. "At this point, everything is just smooth sailing. So a huge shoutout to the team for working really hard and being flexible with our needs. We run huge simulations that take several months, so we need to keep the data around for a long time. Most supercomputers have a frequent purge policy, so I'm extremely grateful for the team figuring out how to let us store data for longer."

Varma added, "Indeed, we definitely couldn't have done so many simulations without the special allocation of

scratch storage for the group; that was extremely important."

2) Omar Faruq Hamim is a Graduate Research Assistant at the Lyles School of Civil Engineering at Purdue University. He recently used Anvil to conduct research involving advanced computer vision techniques and machine learning algorithms, a subset of artificial intelligence (AI). Hamim worked alongside Surendra Reddy Kancharla, a postdoctoral researcher, and under the supervision of Professor Satish



"The team at Anvil has been absolutely terrific in working with us to resolve any issues we had when we started," Nils Deppe, Assistant Professor of Physics at Cornell University

Ukkusuri. Together, the team successfully created a deep learning architecture-based semantic segmentation model for mapping roads and sidewalks in developing countries from remotely sensed, open-sourced data. They were then able to automate the process of inventorying sidewalks by creating sidewalk maps of the study area on a neighborhood scale, based on the output from the developed model. All of this helps develop pedestrian infrastructure by enabling a relatively quick and efficient method of creating and maintaining a sidewalk inventory that local authorities can use for project planning. Hamim and his team were thrilled with what they could accomplish on Anvil:

"So we used HR net plus OCR," says Hamim, "that's the semantic segmentation model. It was state-of-the-art at the time of doing the research, but the problem with that model was that it had millions of parameters and we needed a large number of GPUs to use it. So Professor Ukkusuri contacted Rajesh [Kalyanam, Senior Research Scientist at RCAC] and he helped us get set up on Anvil. We trained our model there and were able to reduce our computational time while also increasing the pixel size of our images. In our case, the images were 640 x 640 pixels, but on our lab computer we were only able to use a random pixel count of 128 x 128. On Anvil we were able to capture the whole image, which increased the performance of our model. So without the GPUs on Anvil, I would not have been able to train the model completely."

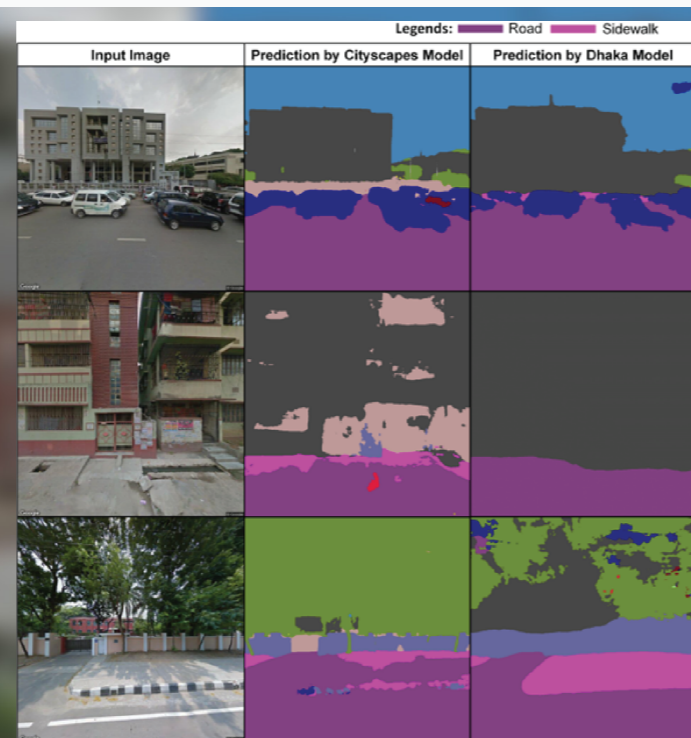
3) Dr. Marco Giometto is an Assistant Professor in the Civil Engineering and Engineering Mechanics Department at Columbia University, as well as the head of Columbia University's Environmental Flow Physics Lab (EFPL). He and the rest of the EFPL group use supercomputers to conduct research that focuses on the study of flow phenomena involving turbulence, heat transfer, and evaporation, specifically within the atmospheric boundary layer (bottom layer of the atmosphere which is in contact with the surface of the earth). Though the specifics of each research project can be quite different, the group's overall goal is to advance the understanding and ability to model turbulent transport in the atmosphere. The group knew they would need a very powerful supercomputer to model turbulence, as it is a multiscale phenomenon, so they turned to Anvil. The group initially requested 7 million core hours on Anvil, which they quickly burned through. They have since extended their allocation request and received approval for nearly 35 million CPU core hours, all on the Anvil system. According to Dr. Giometto, two unexpected benefits to using Anvil were the queue times and max allowable simulation time. On other systems the group has used in the past, the max time allowed per simulation was two days, after which they would be placed back in a long queue to wait until their next turn. This was not the case with Anvil, as the maximum simulation time was four days, and the queue was very short.

"With Anvil," says Dr. Giometto, "one nice thing was that it was not oversubscribed, and it had a four-day simulation duration. And these, I think, enabled us to close on a project that we wouldn't have been able to close on otherwise."

4) Pranav Sharma is an Associate Researcher in the Biological and Soft Materials Modeling Lab at Duke University.

He is using Anvil to develop and run a coarse-grained molecular model of nanoparticles in order to reverse-engineer a phenomenon called self-assembly, wherein nanoparticles spontaneously organize into higher-level structures due to intermolecular forces. Specifically, Sharma focused his project on DNA Origami, an emerging nanotechnology that takes advantage of the phenomenon of self-assembly as well as the properties inherent in DNA structures, with the ultimate goal of deepening the understanding of these structures and eventually leading to the development of a new nanoscale manufacturing paradigm. Sharma is no stranger to HPC, and he found that while Anvil had tremendous power and speed to offer for his simulations, the more exciting aspect of Purdue's supercomputer was its reliability:

"Anvil is great," says Sharma. "I've used many supercomputers, and Anvil has performed more reliably than many of these other systems. Often, I would leave a calculation running and come back to find something had gone wrong, only to run it the exact same way and have it work perfectly, but on Anvil, this was never an issue. I was able to just leave it running, and that's the best thing you can ask for."



Semantic segmentation model examples. From left to right: Input image; pre-trained model prediction; revised, locally trained model prediction.

"Anvil is great, I've used many supercomputers, and Anvil has performed more reliably than many of these other systems. Often, I would leave a calculation running and come back to find something had gone wrong, only to run it the exact same way and have it work perfectly, but on Anvil, this was never an issue. I was able to just leave it running, and that's the best thing you can ask for." Pranav Sharma, Associate Researcher in the Biological and Soft Materials Modeling Lab at Duke University

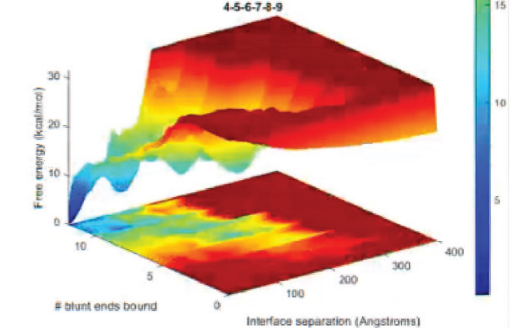
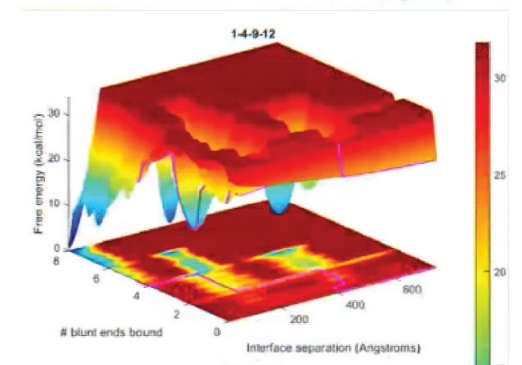
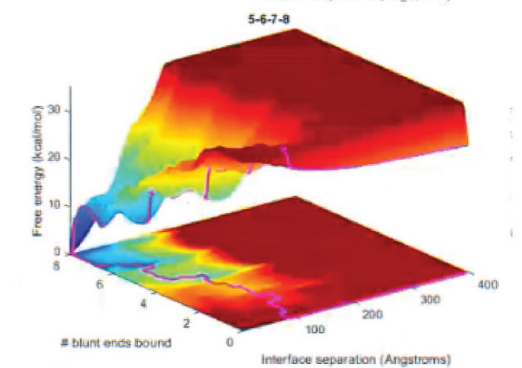
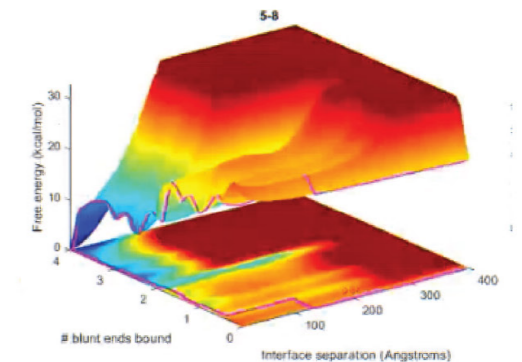
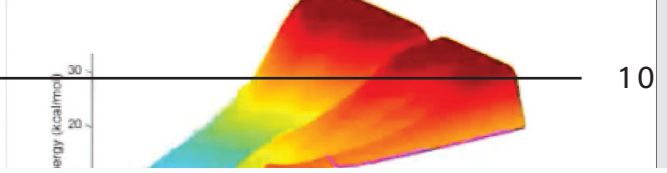
Aside from enabling groundbreaking research across multiple fields of science, Anvil is being used as a tool to develop the future workforce in computing. From professional training and workshops to hands-on learning experiences for students, Anvil is helping to forge the next generation of researchers and cyberinfrastructure professionals.

One major training and educational impact made by Anvil involved supporting the 2023 NIH BigCare Workshop. The BigCare Workshop was a National Cancer Institute-funded biomedical data analysis workshop designed to train cancer researchers on how to visualize, analyze, manage, and integrate large amounts of data in cancer studies. Dr. Min Zhang, the principal investigator on the NCI-funded project, taught the workshop participants the skills needed to analyze their research data, while Anvil provided an HPC environment that had a very low barrier to entry, ensuring that non-HPC professionals could quickly and easily complete their research without having to become an expert in computing.

"We don't want to turn everyone into a computer scientist, because they have more important things to do," says Zhang. "Previously, we had to teach users the front end, back end, command lines, all this kind of stuff, and now it's all gone! Life is so much easier. And everyone was so excited that they wanted to take Anvil to their own institution. Some of them would even say, 'We do have HPC, we do have cloud, but it's not as user-friendly as Anvil.'"

Anvil was so helpful for the workshop that Zhang intends to renew it as the resource for supporting BigCare for the foreseeable future. "It definitely made our workshop run much better and much smoother, and attracts many more researchers. So I think we will carry on this collaboration for not only next year, but the next five years. The cancer researchers will benefit a lot from the Anvil computing environment."

Anvil also supported roughly 1,800 students in a national

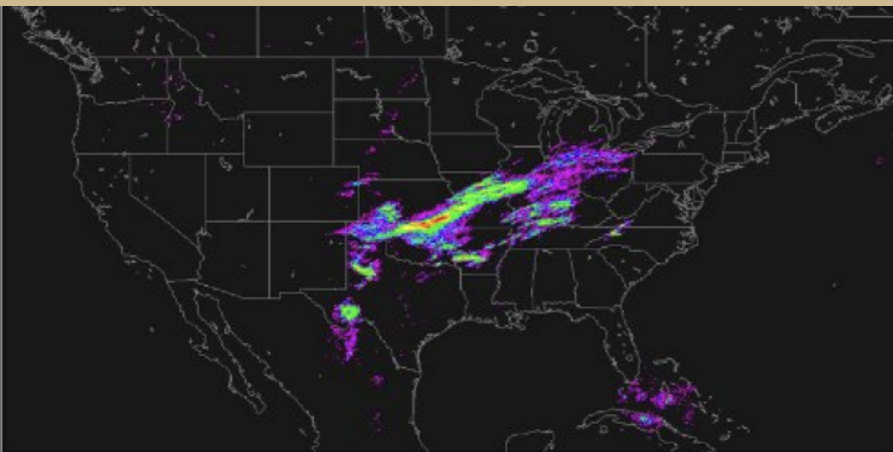


data science experiential learning and research program known as The Data Mine. The goal of The Data Mine is to foster faculty-industry partnerships and enable the adoption of cutting-edge technologies. The course introduces students of all levels and majors to concepts of data science and coding skills for research. The students then partner with outside companies for a year to work on real-world analytic problems. Anvil provided 1 million CPU hours for the program and allowed the students to manage extensive research datasets, thanks to the supercomputer's large capacity.

Anvil has been busy enabling research experience programs for undergraduates as well. One such program involved five students from the University of Illinois at Urbana-Champaign (U of I) who took part in a semester-long research experience program to gain practical knowledge on data analytics and statistical research using common computing HPC resources. Daniel Ries, a Principal Data Scientist at Sandia National Laboratories, led the project, teaching undergraduates in statistics and computer science the skills they will need once they enter the workforce. According to Ries, the intent behind this research experience was multifaceted—he wanted to utilize a real-world research problem to teach students how to scope, execute and refine, and draw conclusions, apply knowledge gained from coursework to actual research, and—importantly—create reproducible research on a common computing platform. By obtaining access to the Anvil supercomputer, the students were able to accomplish all of this and more.



From left to right: Austin Shwatal, Georges Durand, Spencer Bauer, Mallory Klostermann, Sarah Yang



Representative image (NASA) source: https://ghrc.nsstc.nasa.gov/home/lightning/index/data_nldn

"Getting the students on Anvil was not only a benefit, you know, in terms of reproducibility, but in terms of what these students will be doing when they either go to grad school or get a full-time job in the data science world," says Ries. "Most of the work that's done at a company, at a research lab, in academia—the computing is done on servers. You don't do computing on your own laptop or your own computer anymore. Just given the scale of models, the scale of data, it's very common to have to get used to working in a server environment, a Linux environment, things like that, and I don't think actually any of the students had experience with that. So it actually turned out to be a very good, I think, experience for them."

Not only did the students get HPC experience in an actual research application, but the research itself had practical implications. The group focused on a mode of predictive modeling known as "nowcasting." With nowcasting, a research team is looking to predict weather conditions in the near future based on conditions in the very recent past. As part of this project, the undergraduate students set out to build three different predictive models that could determine where lightning would strike in the next 15 to 60 minutes—a type of nowcasting that is immensely useful across multiple sectors. Using data collected from multiple sources, the students were able to develop two traditional statistical models and a third, U-Net deep learning model. The two traditional models, while typically not memory or computationally expensive, benefitted from the use of Anvil due to the sheer size of the data sets. And the U-Net model was trained on the Anvil GPUs, saving the team an enormous amount of time (30-60 minutes per training run versus a day or more without). By the end of the semester, the students successfully developed all three models.

Another research experience that Anvil supported was RCAC's very own Anvil Research Experience for Undergraduates (REU) Summer 2023 program. The 2023 Anvil REU program saw five students from across the nation gather at Purdue's campus in West Lafayette, Indiana, for 11 weeks to learn about HPC and work on projects related to the operations of the Anvil supercomputer. Eight members of RCAC's staff provided mentorship to the five students throughout the summer, helping them to complete four separate Anvil-enhancing projects. The student participants of the program were:

- **Ved Arora**, Data Science & Analytics major from Case Western Reserve University in Ohio
- **Nayeli Gurrola**, Computer Science major from the University of Texas Rio Grande Valley
- **Oluwatuminiu Oguntola**, Computer Science major from the University of North Carolina at Chapel Hill
- **Henry Olson**, Computer Science and Cybersecurity double major from Purdue University Northwest
- **Aneesh Chakravarthula**, Computer Science major from Purdue University



By summer's end, the students had made fantastic progress: they completed their projects, learned technical and people skills they will need when in the workforce, and gained an in-depth understanding of the world of HPC. Each of the four projects is still being used by the Anvil team in some capacity to help advance the system, and two of the students were even able to produce a conference publication from their work, which they presented at the 2023 International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC23).

Industry Partnerships

Anvil's second year of production saw the introduction of its Industry Partnership program. This program allows industry users to utilize the Anvil supercomputer for their business needs, but at a fraction of the cost of private HPC companies. Examples of some of the current Industry Partnership users, as well as projects under discussion, include:

- Smart building technology company (Kubernetes GPU workloads)
- Non-profit engaged in clean energy research (I/O intensive workloads)
- High-resolution weather prediction company (large geospatial datasets)
- AI-driven platform for airport power infrastructure management for electric aircraft
- Electromagnetic propulsion systems
- Generative AI for personalized content
- Company working on cancer detection technology using blood tests
- Life sciences diagnostics company
- Technology company aimed at early detection of TBI and cognitive impairment

Looking to the Future

With artificial intelligence (AI) becoming more and more prominent in society, it is clear that HPC systems will need to assist not only with studying AI, but efficiently utilizing AI for the advancement of science and technology. The Anvil team has recognized that need and has responded accordingly by becoming an official resource for the newly launched National Artificial Intelligence Research Resource (NAIRR) Pilot. The NAIRR is an NSF project aimed at creating a national infrastructure that connects U.S. researchers to responsible and trustworthy Artificial Intelligence

resource. The NAIRR will also provide these researchers equitable access to the data, software, training, computational, and educational resources needed to advance research, discovery, and innovation within the field of AI. By being part of the NAIRR Pilot program, Anvil will directly support research projects that focus on testing and validating AI systems, improving model performance, increasing the interpretability and privacy of learned models, reducing vulnerability to attacks, and assuring that AI functionality aligns with societal values and obeys safety guarantees.

With artificial intelligence (AI) becoming more and more prominent in society, it is clear that HPC systems will need to assist not only with studying AI, but efficiently utilizing AI for the advancement of science and technology.

"We are very excited to take part in such an important effort and help provide the nation with advanced AI computing resources," says Rosen Center Chief Scientist Carol Song, principal investigator and project director for Anvil. "Anvil was intended to lower barriers to applications of high-performance computing, with AI being a key

workload. Joining the NAIRR, the Anvil team will bring together Purdue's long history of supporting advanced computing with the experts being assembled on our campus to work on strategic initiatives, including Purdue Computes and the new Institute for Physical AI."

The Anvil team also intends to continue its pursuit of excellence in HPC education throughout production year three. This summer's Anvil REU program will see an increase from five undergraduate students to eight. These eight students will gain hands-on HPC experience while working to complete four separate, real-world application projects. The Anvil team will also be hosting two summer camp courses for high-school students. The first course, titled "CyberSafe Heroes," will take place from June 16th through June 21st, 2024, and will focus on creating interest in and preparing students for careers in cybersecurity. The second, titled "Code Explorers," will occur the following week, and will teach coding languages and introduce data analysis and visualization techniques to the students. Alongside these direct student interactions, Anvil will continue to provide world-class training sessions surrounding all things HPC. Many of the upcoming training sessions will focus on AI and data analysis applications, and how best to use them in a shared computing environment.

- ▶ Continue to drive for adoption of all Anvil capabilities
- ▶ Operations and collaboration with ACCESS
- ▶ Expand software and tools
- ▶ Expand user support and training
- ▶ Hardware expansions (GPUs/storage)
- ▶ Sustain and broaden outreach
- ▶ Increase Industry engagement

Performance
5.1 PetaFLOPs peak
132,096 CPU cores
360TB of system RAM

Key Features
1000 CPU Compute Nodes
Two 3rd Generation AMD Epyc processors @2.45 GHz, 64 cores per socket, 256 GB system memory, 480 GB local SSD

32 Large Memory Nodes
Two 3rd Generation AMD Epyc processors @2.45 GHz, 64 cores per socket, 1 TB system memory, 480 GB local SSD

16 GPU nodes
Two 3rd Generation AMD Epyc processors @2.45 GHz, 64 cores per socket, 128 GB system memory, 480 GB local SSD, 100 Gbps HDR InfiniBand, 1 PB capacity per node, 100 TB flash

Interfaces
Network Infrastructure

ANVIL
Forging the Future of Computing

NSF
DELL Technologies
EPYC
AMD
ddn

"Anvil has established itself as a major HPC resource to the national research community," says Preston Smith, Executive Director for the Rosen Center for Advanced Computing and co-PI on the Anvil project. "After two years in production, we are pleased with everything Anvil has enabled thus far, whether it be the science conducted on the machine or the training and education opportunities it has provided. Looking ahead to year three, our goal is to continue to innovate, helping expand the boundaries of scientific discovery, while still providing world-class support and education for researchers nationwide. With our inclusion as a resource for the NAIRR Pilot, we are looking forward to the new challenges in the upcoming year."



ENVISION CENTER DEVELOPS VR INCIDENT RESPONSE AND RECOVERY TRAINING

The Envision Center collaborated with a team at Purdue and Texas A&M's Engineering Extension Service to develop virtual reality (VR)-based trainings to help public safety officials better respond to cyberattacks during disasters.

The project, known as Cyber Security Awareness and Strategies to Enhance Resilience of Recovery and Response Operations During Disasters, is funded with a \$1.5 million grant by the Department of Homeland Security and the Federal Emergency Management Agency (FEMA).

"This is the first virtual reality training course funded by FEMA," says Umit Karabiyik, an associate professor in computer and information technology and a co-PI on the project. "It's really one of a kind."

Since July 2023, the team has been accepting students for both VR and web-based learning modules, and a total of more than 100 students have

already registered for the course, with the majority having finished and passed.

Karabiyik reports that the VR component was a particular hit with participants, even though it was initially daunting to some who hadn't used VR before, and the immersive scenarios were especially useful for learning the course concepts.

The training consists of a series of scenario-based, immersive, experiential learning modules in which cyber incidents and attacks occur concurrently with natural and man-made disasters. Through these scenarios, participants observe cause-and-effect reactions to the ubiquitous connected Internet-of-Things (IoT) technologies and identify strategies and techniques to adapt and prevent IoT-based attacks. The goal of the program is to increase the community's ability to prevent an attack and continue to enhance the community's resilience during times of catastrophic disasters.



PURDUE PROFESSOR USES ENVISION CENTER MOTION CAPTURE TECHNOLOGY FOR AI MUSIC PROJECT

A Purdue professor is using the Envision Center's motion capture technology for a project focusing on AI applications with musical instruments.

Kristen Yeon-Ji Yun, a clinical associate professor of music, is using motion capture technology while playing cello to collect data and transfer it to a robot. The end goal is to effectively use the robot to play cello just as well as a human cellist.

"Our goal is to have the robot play music it's never encountered," says Yun. "Our first step was the motion capture from the Envision Center. We wanted to have the data about cello movement because it's quite complicated. All the bow angles for the strings are different and there is left hand muscle memory. This data will be helpful for us to mimic human movement. We will transfer that data into a robot stimulating program."

This project used the Optitrack motion capturing system to track body joints, equipment, and left-hand fingers to generate a 3D reconstruction of the cellist's motion. The optical tracking technology provides just under 1mm (about 0.04 in) accuracy in each individual point movement, which can be used to describe body and instrument motion. The performer wears a specialized suit with retroreflective markers, which then can be extrapolated to a human skeleton, accompanied by the tracking of the cello body and bow.

The data collected from the motion capture will be used to train the robot, with the end goal being for the robot to learn how to play a musical instrument. Students working with Yun are currently building the right arm of the robot and the controller necessary for it to play the cello.

Yun is also working on a second project using this data, focusing on creating an app, Evaluator, to help string instrument players with their solo and ensemble practice.

POLITICAL SCIENCE PROFESSOR PUSHES BOUNDARIES WITH RESEARCH IN VIRTUAL REALITY

In collaboration with the Rosen Center for Advanced Computing Envision Center, a Purdue political scientist is leveraging the immersive nature of virtual reality and incorporating it into his research in ways that are novel for his field.

Bryce Dietrich, an associate professor of political science, studies political and social questions about how people move through space and relate to each other. One of the challenges of that work is that if you're looking at a large gathering, you can see that people positioned themselves in a certain way, but you don't know very much about what that behavior means.

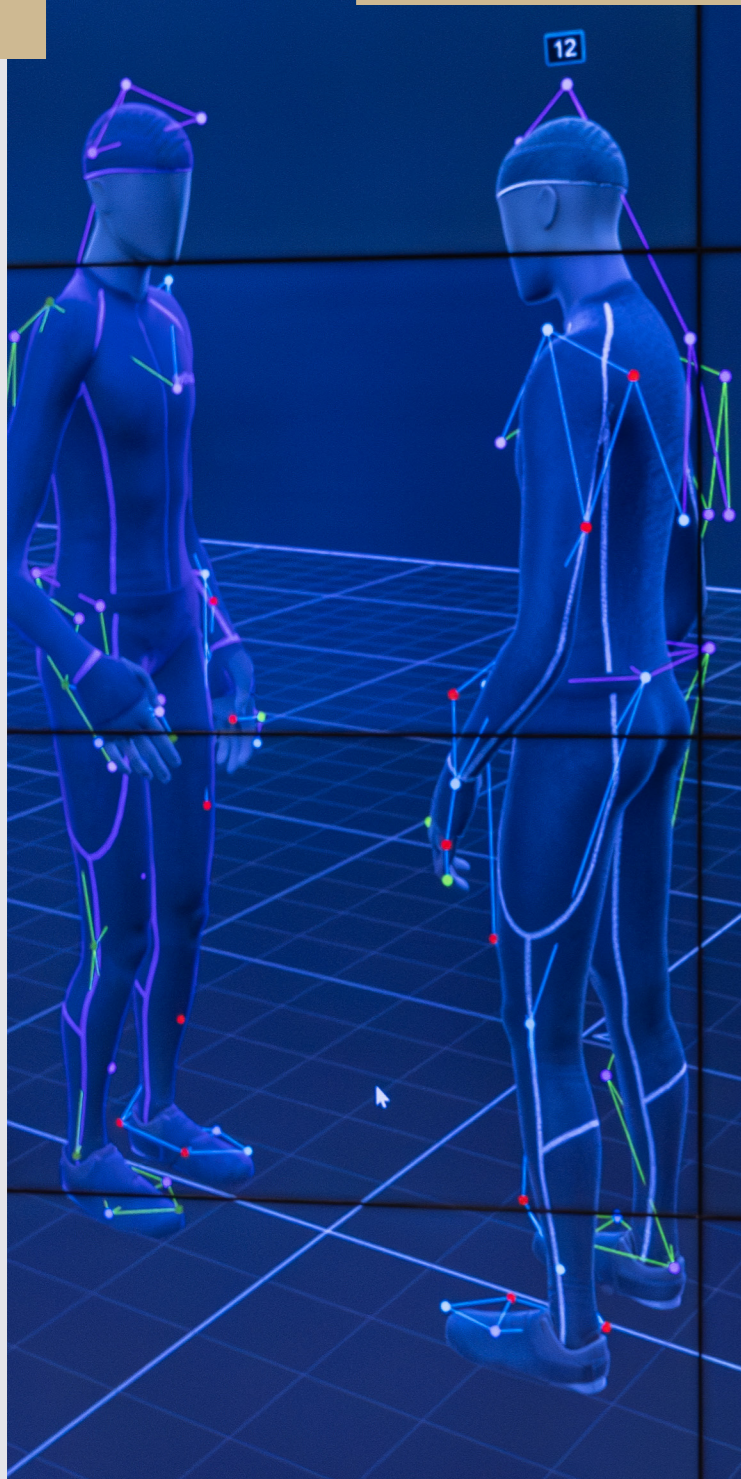
That's where virtual reality and the Envision Center come in – within a simulated environment, Dietrich can create the type of social interactions he's interested in studying and can do so in a controlled setting.

Initially Dietrich's collaboration with the Envision Center was just an exploratory exercise to see what a virtual reality environment could look like.

"But when we saw the simulation they had created for us at the Envision Center, I was just blown away by the immersive nature of it," says Dietrich.

"What struck me the most was that you know that you're in a simulated world, but all your natural impulses based on the life experiences you've had are naturally reflected in it, and you move through it just as you would the real world."

The simulation Dietrich and the Envision team developed takes advantage of the Envision Center's virtual reality theater, The Forge, which allows multiple people to share the same virtual reality experience and to physically move through real space as they move through the simulated environment.



ENVISION CENTER EVENT LETS RCAC STUDENTS SHOWCASE PROJECTS

The Rosen Center for Advanced Computing (RCAC) recently hosted a student presentation event at the Envision Center, which gave RCAC's student employees a chance to demonstrate their hard work and achievements over the last year.

More than 20 student workers from RCAC presented their work at the event. Departments highlighted included the Envision Center, HPC hardware engineering, data science and operations and outreach. The event also provided a platform for the students to network with each other and learn more about the various kinds of research opportunities within RCAC.

"It was a good opportunity for the students and staff to learn about all the different student projects throughout RCAC, said Amanda Warren-Glowe, program manager at the Envision Center.

"The students did a great job presenting and we were especially excited to celebrate our graduating seniors." "I really enjoyed the lightning talks," says Uma Claire Teelock, a junior double majoring in game development and design, animation, and visual effects, whose project focuses on measuring how far away a person veers away from two people arguing depending on their race.

"I got a chance to show off my motion capture work and similar exciting work."

Teelock adds that it was also very interesting to see what

RCAC students outside the Envision Center are working on.

"The biggest thing I took away is that RCAC has a wide variety of projects serving different parts of Purdue, and we have lots of talented students on our team!"

RCAC's student program CI-XP (Cyber Infrastructure-eXPerience) includes student workers in RCAC, the Envision Center, the Scientific Solutions Group and Anvil REU programs, which provide paid summer internships on campus. The purpose of the CI-XP program is to provide work opportunities and real workplace experience that enhance the student's education through the development of professional skills, responsibilities, habits, attitudes, self-confidence and self-development. RCAC staff serve as mentors to the students.

RCAC provides advanced computational resources and services to support Purdue faculty and researchers. They also conduct their own research and development to enhance the capabilities of these resources. RCAC includes the Scientific Solutions Group, a team of research software engineers and cyberinfrastructure research scientists who develop cyberinfrastructure solutions to help advance scientific discovery and education, and the Envision Center, which is dedicated to assisting, supporting and collaborating with faculty, students and industry in scientific visualization, virtual and augmented reality and media creation.

20-YEAR BIRTHDAY CELEBRATION FOR THE ENVISION CENTER

Purdue's Rosen Center for Advanced Computing recently celebrated the 20th anniversary of its Envision Center, a group dedicated to assisting, supporting, and collaborating with faculty, students, and industry in scientific visualization, virtual and augmented reality, and media creation.

On April 17th, from morning until evening, the Envision Center (EC) opened its doors for all to celebrate 20 years of creativity and innovation. The day began with an Open House, running from 10 a.m. until 12 p.m., where EC staff and students helped to showcase many different projects taken on by the center throughout the years. Attendees of the open house were able to experience cutting-edge collaborative virtual experiences firsthand and engage with the development team dedicated to their advancement at Purdue University. One major project put on display was the newly developed Collab XR platform. This platform is a shared environment that allows anyone in a headset to view and interact with the same virtual content pieces together, viewed in fully virtual (the entire field of vision is virtual, obscuring all of the real environment, and other participants appear as virtual avatars) or passthrough augmented reality (cameras feed the real room and people to the headset view, overlaying virtual content to appear to exist in the same space). For the birthday celebration, the Envision Center used the Collab XR platform to showcase a variety of past projects, including a radiation detection simulation, supernova data visualizations, weather system visualizations, insect models, and more.

During the open house, the Envision Center also broke out the MoCap (motion capture) suits to highlight its motion capture systems and abilities. Attendees were invited to engage in a VR "sword fighting" game, wherein they dueled (using foam swords) in virtual reality against an opponent who donned a MoCap suit. The motion capture system tracked the movement of each participant's sword, as well as the body movements of the opponent, and the first-person perspective from the participant was projected onto a massive screen lovingly known as the Forge Wall. In this way, the cheering bystanders were able to see exactly what the participant was seeing—a platform surrounded by lava, and a realistic moving robot attacking with a lightsaber. While the VR duel is an exciting way to highlight the abilities of motion capture systems, rest assured that it can be—and is—used for serious research. Recent projects that have taken advantage of Envision's MoCap system include capturing tandem dance movements in order to develop virtual AI dance partners, recording body and instrument movements of a Cellist's performance to use as training data in robotics, and a character

animation performance capture to be used in a racial bias study.

Overall, the open house drew quite a crowd. If you weren't able to attend but would like to see an example of the Collab XR platform, please visit the Envision XR Lab webpage: <https://www.rcac.purdue.edu/envision/xr-lab>. You may also schedule a tour of the Envision Center and experience the center's work firsthand, here: <https://rcac.purdue.edu/about/tour>

After a one-hour lunch break, the 20th-anniversary festivities resumed. At this point in the day, the theme shifted from looking at past accomplishments to focusing on current and future XR research and capabilities. From 1 p.m. until 5 p.m., the Envision Center hosted an XR Symposium featuring multiple speakers from across campus, all highlighting the ways in which they are leveraging virtual and augmented reality technologies at Purdue. Each of these speakers discussed how they use XR in research and learning, obstacles they have found in their work, and potential ways to overcome these obstacles.



- **George Takahashi and Danny Milisavljevic** presented "XR Lab, Collaborative XR platform, and Purdue XR Consortium." The pair discussed the newly developed Collab XR platform and how Milisavljevic uses it to foster interactive learning in his astrophysics classes.
- **Voicu Popescu** presented "Immersive Visualization in Large Lectures." During the presentation, Popescu spoke about the fact that large lectures (100+ students in a single classroom) are here to stay, but they do not adequately engage students. He then discussed how he, along with Bedrich Benes, Alejandra Magana, and John Springer, are working to combat this by incorporating XR immersive visualizations.
- **Victoria Lowell and Stuart White** gave their presentation, titled "Emerging Technologies & Learning Design & Tech." The pair discussed the necessity of integrating emerging technology within learning environments to better prepare students for life after graduation and different ways in which this can be accomplished.
- **Matthew Ward** presented "VR Anatomy for Engineering in Medicine." During his presentation, Ward discussed his Spring 2024 Biomedical Engineering pilot course, which is built upon the use of a VR anatomy lab, and the benefits of having access to in-depth VR cadavers for the medical community.
- **Stylianios Chatzidakis** gave a presentation titled "Exploring the VR/AR/MR Options - The Nuclear Pharmacy & Nuclear Engineering Experience." In this presentation, Chatzidakis discussed the work that he, Kara Weatherman, and Wanju Huang undergo leveraging XR technology for environments where physics plays a significant role, such as mimicking radiation spills and exposure in a nuclear pharmacy or reactor.
- **Senay Simsek and Ben Paxson** presented "IV-STAAR: Integration of Virtual Reality to Support Transformative Active-Learning in Agriculture." Their presentation highlighted the usefulness of XR technology within the field of Food Science and showcased some of their work, particularly in the realm of Food Chemistry and Food Safety.
- **Tawfiq Salem** gave the final presentation of the day, titled "Integrating AI into VR Simulations - Intercultural Teamwork Skills." Salem discussed the work that he, Kris Acheson-Clair, Mesut Akdere, and George Takahashi have undertaken to train both students and the workforce in smart manufacturing on team/communication skills through the use of XR technology.

Once the presentations were complete, a panel discussion ensued to close out the festivities. The panel members were Voicu Popescu, Victoria Lowell, Shawn Jung, Wanju Huang, Hua Cai, Stuart White, Stylianios Chatzidakis, and George Takahashi. The panel discussion focused on using XR in the classroom, covering topics such as fleet management, consumer vs. enterprise, accessibility, best practices, assessment, and support. To view the full panel discussion, please visit [here](#).

The Envision Center's 20-year day of celebrations was a resounding success. The group was able to reminisce on all that has been accomplished throughout the years while also building excitement for all that is to come. The group would also like to make a special note:

"While we are proud of our work, we know that nothing can be accomplished in a void, so we would like to extend a huge thank you to all the researchers and users who have worked with us since our inception. Without you, we can truly say that none of this would be possible. Here's to the next 20 years!"

Overall, the open house drew quite a crowd. If you weren't able to attend but would like to see an example of the Collab XR platform, please visit the Envision XR Lab webpage.



THE FUTURE OF TEACHING AND LEARNING HAS ARRIVED

Purdue University's Envision Center, part of the Rosen Center for Advanced Computing, has developed a virtual reality platform that utilizes the Meta Quest 3 headsets to bring the future of teaching and learning to our doorstep.

The Envision Center (EC) has been in the field of virtual reality for quite some time. In fact, they recently celebrated their 20th anniversary. And with 20 years of learning and innovation under their belt, it's no surprise they are at the forefront of advancing the field of interactive education. Introducing "Collab XR"—the newly developed shared learning platform designed to leverage all of the Meta Quest 3 headset's newest features. The Collab XR (XR=extended reality) platform is a shared environment that allows anyone in a headset to view and interact with the same virtual content pieces together, viewed in fully virtual (the entire field of vision is virtual, obscuring all of the real environment, and other participants

appear as virtual avatars) or passthrough augmented reality (cameras feed the real room and people to the headset view, overlaying virtual content to appear to exist in the same space). Developed in conjunction with Professor Danny Milisavljevic, the Collab XR platform is the first of its kind, merging virtual reality (VR), augmented reality (AR), and mixed reality (MR) capabilities with training, research, and education. One of the best features of Collab XR is its ability to allow for co-located mixed-reality learning experiences. Instead of looking at images in a textbook, or even playing with a 3D model of an ancient artifact, imagine taking a classroom full of students and virtually placing them at the archaeological site. While immersed in the environment, they could interact with archaeological artifacts, making them bigger or smaller, viewing them from any angle they wish, all while having the instructor and other class members right beside them. With this level of data visualization, the

ability to reach the students and help them truly understand the data grows by leaps and bounds. In fact, Collab XR has been so successful that it is already being put to use in the classroom.

Dr. Danny Milisavljevic is an Associate Professor of Physics and Astronomy at Purdue University. He uses the Collab XR platform to help explore and share complex 3D data with his students—namely, the remnants of stellar explosions. Milisavljevic likens himself to a "CSI Bomb Tech" but for astrophysics. By looking at what is left behind after a star explodes, he learns about the properties of that star from before its death. Traditionally, the biggest hurdle in conveying this type of data (to other researchers and students alike) stems from the difficulty of mentally visualizing a 3D object based on 2D information. To combat this in the classroom, Milisavljevic first turned to 3D animations but found that this method was also lacking.

"I could create animations that provided some perspective," says Milisavljevic, "but ultimately, it was ME that decided which angle I wanted to rotate about, and therefore it was me that got the most information out of the animation. When I show it to a student, they aren't going to have the same spatial comprehension, and that's largely because they lack the agency to be able to rotate it themselves. I've found that if you have the ability to change the angle with your hand or by moving your head around, it allows you to unpack the 3-dimensional properties much more efficiently."

Milisavljevic took this problem to the Envision Center, and Collab XR was born. Now, Milisavljevic can use the Collab XR platform to bring a room full of students into a co-located AR environment, where they can all explore supernova remnants together in the same virtual classroom. This gives the students the ability to examine the 3D reconstructions on their own terms, while still being able to interact with their professor or classmates. Milisavljevic has used Collab XR to conduct immersive classroom lectures for some of his courses, and the students all love it.

"The response from students has been incredible," says Milisavljevic. "The excitement and reward of the experience is heightened by exploring content collaboratively, and they report far better comprehension of the multidimensional data as compared to traditional screen or blackboard presentations. This is a game-changing instructional platform."

Another Purdue professor who is using Collab XR in the classroom is Dr. Robin Tanamachi from the Department of Earth, Atmospheric, and Planetary Sciences. Tanamachi has conducted small-scale lectures for her atmospheric science students using the Collab XR platform. During these lectures, Tanamachi and her students interact with radar observations of storm clouds. Doing so in a co-located XR environment gives her the ability to provide more in-depth

insight into meteorological phenomena while allowing the students to better understand its spatial geometry.

"Much of what goes on inside severe thunderstorms is dynamic, three-dimensional, and invisible to the naked eye. Weather radar gives us the ability to probe the interior structure of the storm using a microwave beam. But, looking at those data on a flat screen and trying to reconstruct the full three-dimensional structure of the storm is cognitively challenging. With Collab XR, I've been able to let the students interact with the data in a way that isn't possible on a screen. For example, they can rotate the data, walk through it, look at it from different angles, and enlarge features of interest. Most importantly, they can switch between different radar variables and see the spatial relationships between features. These features and their behaviors may precede hazardous weather near the surface, like large hail and tornadoes."

Collab XR is able to accomplish all that it can by taking advantage of the Meta Quest 3 headset's latest features. It utilizes mixed reality, spatial anchors, shared environments, and multi-user connectivity. Even more impressive, Collab XR can be used for both local and remote instruction. As long as the user has a Meta Quest 3 headset, they can participate in the class from anywhere in the world, giving them the same level of instruction and experience that they would receive if they were in the room with the instructor. The best part of Collab XR? There's no need for individually built applications for each use case.

"Our platform is more of an ecosystem, not a hyper-specialized one-off," says George Takahashi, Lead Visualization Scientist at the Envision Center. "As long as we have the content, any class can see it."

This means that any field or discipline can utilize the Collab XR platform to bring advanced visualization capabilities to the classroom. As long as the instructor has an idea of what they want and the data needed to create the visuals, the Envision Center can get them up and running on the platform in no time.

"We are very excited to have developed our collaborative visualization platform and to have it grow into a functional co-located mixed reality tool," says Takahashi. "Having the ability to intuitively communicate complex three-dimensional information was previously limited to physical models. Virtual Reality opened the door to bring people together virtually to observe physics-defying large models. By moving to mixed reality and generalizing the models to accommodate data from any domain, this platform brings people together physically and is more scalable and accessible than ever before."



PURDUE UNIVERSITY'S RESEARCH SOFTWARE ENGINEERING CENTER OPENS!

The Purdue Center for Research Software Engineering (the RSE center) recently opened its doors as a university-approved center within the Rosen Center for Advanced Computing. Its official establishment recognizes the increasingly vital role that software plays in all fields of scientific research, and formalizes RCAC's software engineering efforts at RCAC to better support research at Purdue.

The RSE center's mission: To help accelerate research and increase their impacts through the creation of innovative, robust and sustainable research software.

With streamlined processes for project intake and management, consistent software engineering practices and oversight, proposal development, and business management, the new RSE center will increase capacity to better support the university's large and strategic initiatives and center needs.

Purdue RSE Services:

- Proposal collaboration
- Software design
- Code development
- Software configuration and deployment
- Consulting
- Training

What is a Research Software Engineer (RSE)?

A Research Software Engineer (RSE) combines professional software engineering expertise with an intimate understanding of research. They design, write, and maintain software that is used to support, disseminate, or undertake research.

Notable Projects:

New GLASSNET Tools: Climate, land-use, and agricultural data tools developed by the Purdue RSE GLASSNET team members have recently been added to the online platform MyGeoHub, allowing researchers across domains to easily access the data they need to progress towards meeting the United Nations Sustainable Development Goals.

StreamCI: Hosted on Purdue's Geddes cluster, StreamCI is a cloud-based, real-time smart sensor and device data management and analysis platform designed to streamline the entire lifecycle of sensor data. It is associated with GABBs (Geospatial Analysis Building Blocks), an NSF-funded project aimed at creating a powerful web-based system that enables researchers worldwide to manage, curate, share, analyze, and visualize geospatial data for diverse applications, from predicting damaging floods to projecting climate change effects on vulnerable populations. The project integrates geospatial data hosting, processing, and sharing capabilities into Purdue's HUBzero platform. From registration and ingestion to preprocessing and storage, StreamCI offers data providers a robust solution for managing their sensor data in the cloud. For data consumers, StreamCI enables powerful data querying

with filters such as date, spatial extent, and value range, alongside intuitive data visualization tools. Whether you're a data provider or consumer, StreamCI provides a comprehensive and user-friendly platform to manage, analyze, and visualize sensor data in real-time

Cyberfaces: CyberFACES is a platform for delivering CyberTraining in the areas of water, climate, and environmental sustainability. It is funded by research grants from the US National Science Foundation. Currently, the platform hosts a suite of modules that involve the use of online data and tools to address related to water, climate, and environmental sustainability. Specifically, the users of CyberFACE modules will use Jupyter Notebooks to access data, process data and perform computational tasks. One or more modules are then packaged to offer learners badges, courses, and certificates. Purdue RSE has significantly contributed to the advancement of cyber-enabled education through its involvement in projects like CYBERFACES. Leveraging its cutting-edge HUBzero technology, Purdue developed innovative tools such as RWater, an internet-based educational platform that empowers students to analyze complex hydrologic processes in real-time.

RSE Team Highlights:

I Luk Kim presented "Japper: A comprehensive Framework for Streamlining Jupyter-Based Scientific Web Application Development" at PEARC24, where it won the award for "Best Short Paper." Japper is a comprehensive framework designed to simplify and accelerate the development of Jupyter-based scientific web applications. Japper addresses common challenges such as limited front-end capabilities, complex environment management, and lack of architectural guidance in Jupyter Notebook development. The framework provides a robust toolkit, best practices, and architectural patterns tailored specifically for the Jupyter ecosystem, enabling intuitive interfaces, advanced customization, and streamlined deployment.

smart buildings, AXIN addresses key challenges such as data heterogeneity, storage, and real-time processing demands. The platform leverages technologies like StreamCI, JupyterHub, and NVIDIA Triton Inference Server, offering a scalable and flexible solution that supports the entire lifecycle of data-driven modeling, from development and training to deployment and execution.

Jaewoo Shin presented "Integrating ML/AI Workflows in a streaming Data Management and Processing Platform for Building Energy Research" at PEARC24. The AXIN data lake, developed by the RSE team, is an open-source platform designed to streamline the integration of machine learning (ML) and artificial intelligence (AI) workflows in building energy research. By managing and processing large volumes of streaming IoT data from

Rajesh Kalyanam presented "The Science-i cyberinfrastructure platform" at the 2024 Science-i Bridging Worlds Workshop. Kalyanam discussed the Science-i cyberinfrastructure platform and how researchers can seamlessly request access to relevant datasets and perform analysis and computation on them, all from a web-based platform. He also gave an overview of the broader cyberinfrastructure resources available at Purdue University, including the Anvil supercomputer that can be used to perform resource intensive computations as well as machine learning workflows.

UPDATES TO GILBRETH GPU CLUSTER

To better serve the needs of researchers doing AI, machine learning 80 NVIDIA A100 GPUs recently added to Gilbreth community cluster

After expanding the Gilbreth community cluster's GPU nodes in 2022 and yet again last year, RCAC has once again added even more GPUs to the Gilbreth cluster to meet demand from the Purdue community.

With the recent addition of 80 new NVIDIA A100 GPUs inside 20 Dell PowerEdge xe8545 compute nodes, the Gilbreth cluster now has 411 GPUs, nearly four times its original capacity.

The new nodes include NVIDIA's NVLink technology, which allows for faster communication between the GPUs and will improve speed and access to memory for researchers who use multiple GPUs at once.

The infrastructure expansion supports RCAC's broader strategic plan to offer AI and machine learning expertise, offer new trainings, and partner with faculty on proposals.

"Recently, our group was able to achieve a 44% increase in thermal emitter efficiency, which in large part was due to the high-performance capabilities of Gilbreth. Without Gilbreth, tackling these research challenges would be infeasible," says Michael Bezick, a sophomore in computer science working in the lab of Alexandra Boltasseva, the Ron And Dotty Garvin Tonjes Distinguished Professor Of Electrical and Computer Engineering, who is working on developing machine learning techniques to optimize the topologies of nanophotonic structures.

"This training would take significant amounts of time on a single consumer grade GPU. Parallelizing the workload across many GPUs, the industry standard in training state-of-the-art models, allows us to achieve a many-fold speedup in training, and the greater memory capacity allows us to utilize larger, deeper models," adds Bezick.

The new A100 GPUs are offered under a similar pricing model as CPU-based community cluster systems, meaning researchers can choose between purchasing per-GPU units through either a one-time five-year charge or an annual subscription.

Gilbreth cluster storage capacity doubled to meet the needs of AI researchers

Recognizing the need for faster and larger storage capacity in emerging areas of science, RCAC has recently doubled the storage capacity in Gilbreth,



its community cluster that is optimized for communities running GPU intensive applications such as AI and machine learning.

Gilbreth's storage was not only upgraded to feature twice the capacity of the previous storage, but also features an improved design that results in faster storage transactions.

The new storage system uses DDN's Exascalr 400NVX2-S appliance with a total capacity of 4.56 PB (4.3 PB usable) and features a tiered approach just like before, but it offers a much larger persistent, fast nonvolatile memory express tier to improve the metadata handling and data caching. The new storage controller designer and its significantly improved hardware will bring the controller closer to the storage and speed up data storage processes.

These updates enhance the data pipeline to Gilbreth's GPUs and processors and

ultimately help host and access larger datasets and speed up applications, reducing researchers' time to science.

"This improvement to Gilbreth's storage is part of our many steps to support cutting-edge research in artificial intelligence and scientific domains," says Arman Pazouki, RCAC's director of scientific applications.

"As the number of GPUs and FLOPs (floating point operations per second) of each GPU grow, they will continue to work with larger quantities of data in each simulation. AI applications also rely on large amounts of data and files. The Purdue Computes initiative will have a defining role in the growth of GPU utilization on campus. All of this points to the significance of data, both the capacity and transfer speed, and that's what this upgrade aimed to address," adds Pazouki.

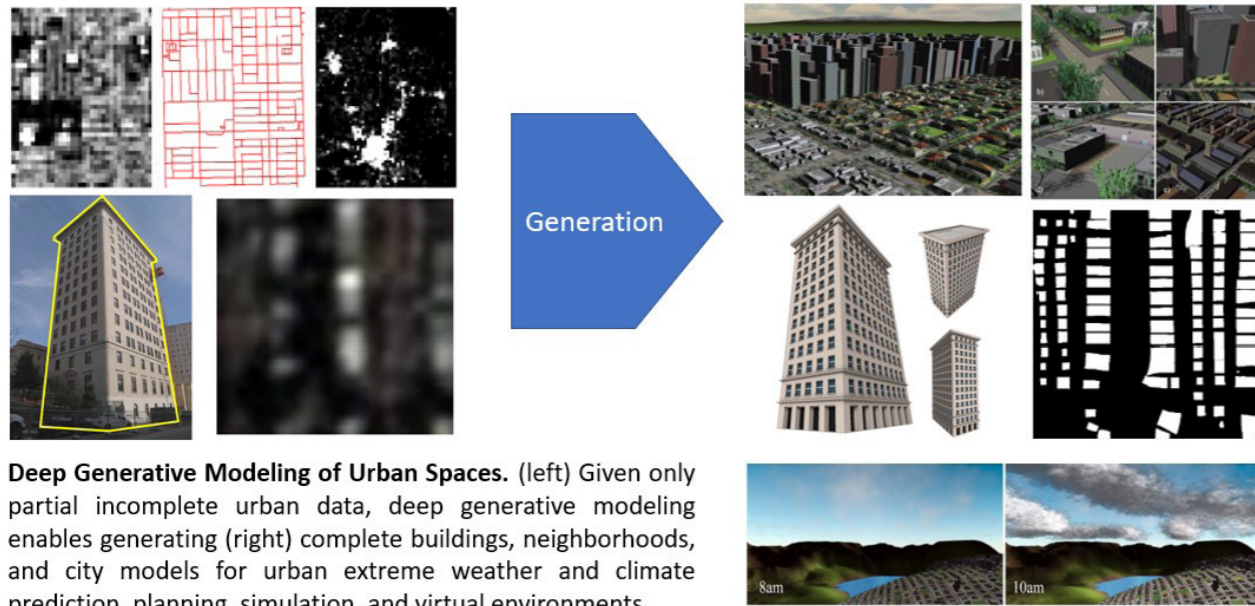


Purdue research team uses 100 RCAC GPUs to create urban representations of 330 US cities

In a multidisciplinary endeavor that has far-reaching implications for computer science, urban planning, digital urban forestry, ecological sciences, and related domains, a team led by Daniel Aliaga, associate

professor of computer science, has leveraged the power of 100 of RCAC's Gilbreth GPUs to advance the field of deep generation of urban-related content.

Urban Layouts



Deep Generative Modeling of Urban Spaces. (left) Given only partial incomplete urban data, deep generative modeling enables generating (right) complete buildings, neighborhoods, and city models for urban extreme weather and climate prediction, planning, simulation, and virtual environments.

Aliaga, along with Adnan Firoze, Liu He, Aocheng Li and Nikhil Makkar, all doctoral degree candidates from the same department, used the GPUs to create urban representations for 330 cities across the United States, with a significant emphasis on sustainability and the cities of the future.

"The availability of the GPUs enabled us to improve algorithm development as well as run the model to enable creation of the results," said Aliaga.

Aliaga and his team embarked on three interlinked projects using the computational power of the GPUs: urban forestry, urban layouts and historical urban layouts.

Aliaga expressed his excitement over the accomplishments of the team, stating, "Our research has pushed the boundaries of what is possible in generating urban-related content. The utilization of the RCAC GPUs has allowed us to achieve unprecedented milestones and significantly contribute to various fields, from computer science to ecological sciences."

Urban Forestry

The team used the RCAC GPUs to detect trees in all of the target cities, resulting in 273 million trees throughout 8.2 million acres of urban space. The GPUs helped to explore hyperparameters and train the underlying models. Once trained, the actual inference process from satellite imagery is relatively

The researchers also used the GPUs to generate urban layouts from sparse data. In particular, the GPUs enabled the team to explore hyperparameters including clustering and sampling strategies to understand the huge design space and arrive at a trained deep engine that using only 2-4% of the urban layout information of a city (such as building locations, sizes and heights) can generate the entire urban layout with only a few meters, or percent, of error. The GPUs were used to optimize more than one billion parameters, using about 100 hours wall time, resulting in position errors of less than two meters, area errors of less than 20 square meters, height errors of less than half a meter, and almost zero building count error for several million buildings nationwide.

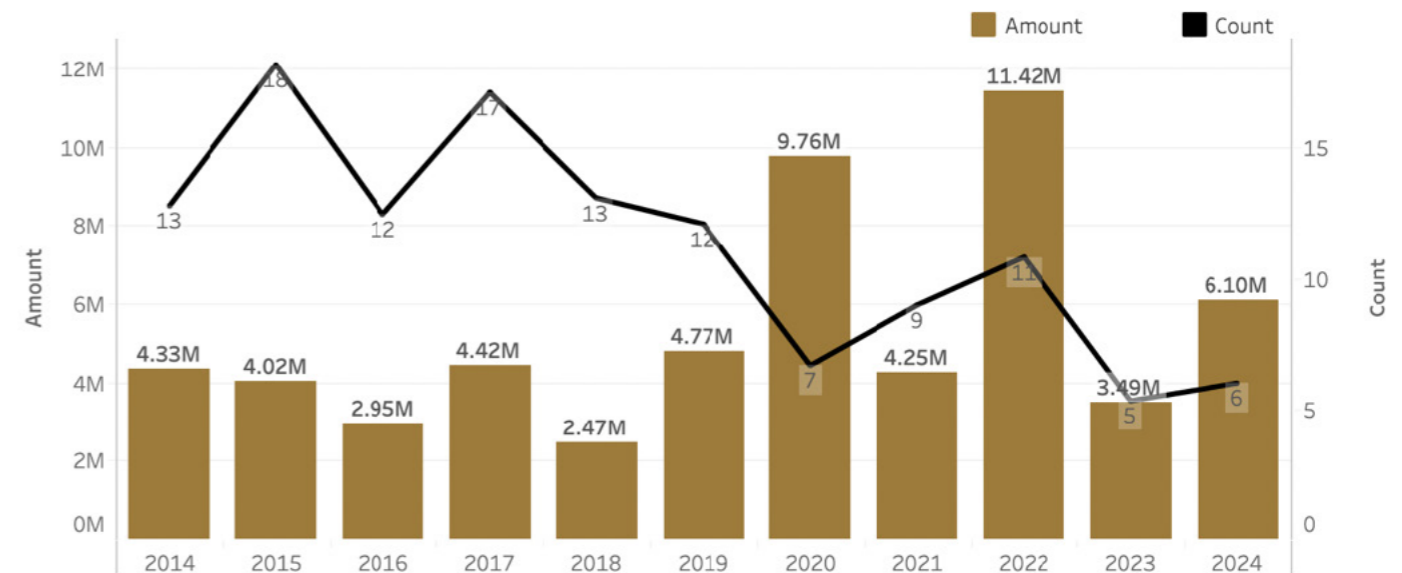
Historical Urban Layouts

The team is also using the GPUs to help in-progress research of pluralistic image completion and infilling which is useful for archaeological sites, among other applications. Unlike a current city, an ancient archaeological site only has a small percentage of the site remaining. However, this is a similar inference setup, so the same methodology can be applied. Initial results from sites in Peru, Greece, and Turkey are quite promising.



Aliaga expressed his excitement over the accomplishments of the team, stating, "Our research has pushed the boundaries of what is possible in generating urban-related content. The utilization of the RCAC GPUs has allowed us to achieve unprecedented milestones and significantly contribute to various fields, from computer science to ecological sciences."

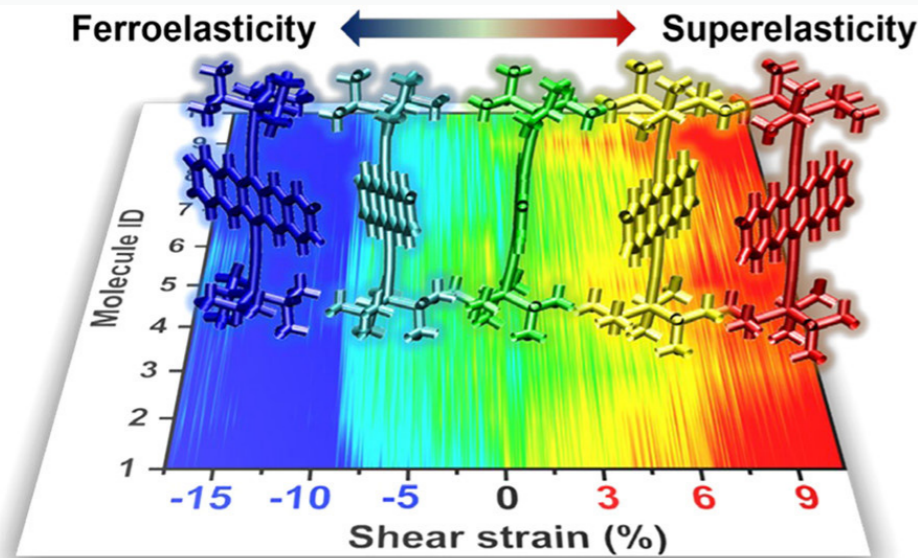
Research Scientists from RCAC were awarded \$6.1M of new sponsored research in FY 2024.



Filters: Purdue System-Wide; Purdue Only; Campus: All; Colleges: All; Department: 6802000000 - IT Research Computing

CLUSTER/STORAGE SYSTEM SCIENCE HIGHLIGHTS

RCAC supercomputers help Purdue professor unlock mysteries of mechanics and material chemistry



A Purdue researcher who studies material chemistry is using RCAC's Brown community cluster to learn more about how materials behave in products like organic electronics and lithium-ion batteries.

Kejie Zhao, professor of mechanical engineering, uses his Brown nodes to explore why a class of materials known as superelastic organic crystals are deformable without sacrificing electrical conductivity – properties that make them well-suited for use in making the

flexible and printed devices used in bio and optoelectronics. Using Brown, Zhao has run simulations to understand on the molecular level why this material is deformable and how the electrical conduction happens in the polymorphs.

This project forms the basis of an NSF CAREER grant, a prestigious award for early-career faculty, awarded to Zhao in 2020.

"Brown is a critical resource for us because these simulations are computationally-intense and expensive, and it's not something we could do on our

own workstations," says Zhao. "The cluster is a really important resource for us to run the projects."

RCAC, College of Agriculture partner to reshape agriculture research

RCAC has partnered with the College of Agriculture to develop cyberinfrastructure that better serves the computational and data management needs of agriculture researchers.

The College of Agriculture's Ag Alumni Seed Phenotyping Facility (AAPF) provides researchers with a controlled environment in which to run experiments on plants and simulate conditions like drought and different nutrient mixes. Experiments run in the AAPF can take months or even years to complete, and researchers can collect extensive data without damaging the plants via three different specialized cameras. These cameras collect information about the plants in multiple wavelengths, creating huge amounts of data in the form of RGB, hyperspectral, and x-ray images.

The College of Agriculture IT group initially partnered

with the AAPF facility, select researchers, and an external vendor to create a high-performance computing and flexible storage solution to address the significant compute and storage required to process these images.

"Once Geddes [RCAC's Community Composable Platform] was implemented, we immediately started working with RCAC to compare our solution with theirs, where we learned their enterprise offering would be our best path forward long term," says Biggs.

"It was clear their flexible system and dedicated support would allow us to focus more on research-specific needs and less on the underlying technology. By partnering with RCAC using Geddes, we can help our researchers save time and money, putting processed data in their hands faster and shortening the time to discovery."

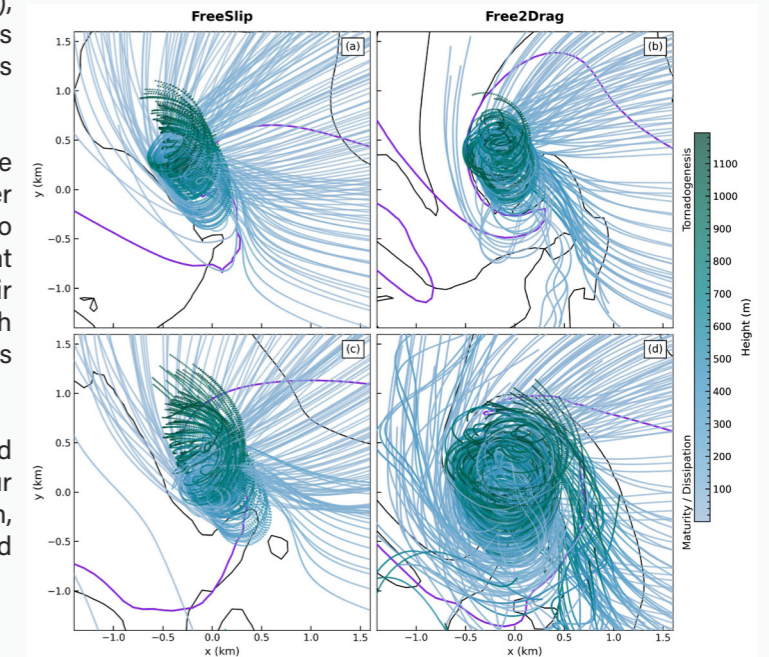
Purdue professor uses RCAC community clusters for weather simulation research

Daniel Dawson, an associate professor of earth, atmospheric, and planetary sciences, conducts high-resolution simulations of severe thunderstorms and tornadoes using the Bryan Cloud Model (CM1), which demands extensive computational resources that the Bell and Negishi community clusters adeptly provide.

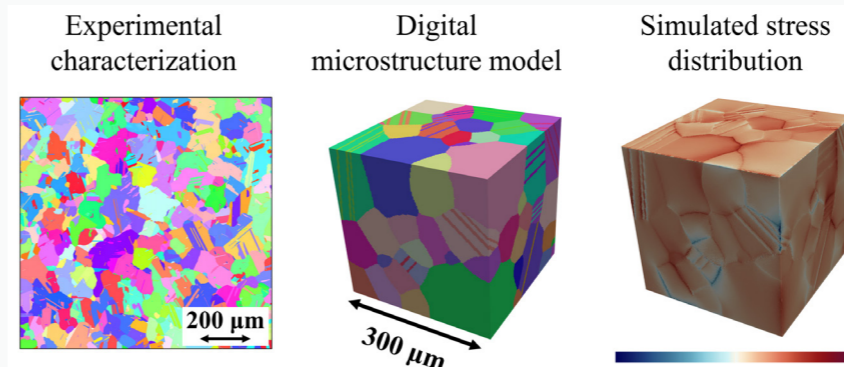
"The clusters are essential for performing the weather simulations because of both the sheer number of simulations we will eventually need to perform, as well as the high grid resolutions that some of them will require," says Dawson. "Their ability to handle thousands of simulations with intricate detail is why we can push the boundaries in severe weather modeling."

"The efficiency, power, and reliability of Bell and Negishi have been paramount. Their role in our success cannot be overstated," adds Dawson, who says he also appreciates the responsive and

supportive RCAC staff, emphasizing their contribution to his team's achievements.



RCAC cluster Negishi powers study of additively manufactured materials



Researchers at Purdue are studying the behavior of additively manufactured (otherwise known as 3D-printed) materials, thanks to the powerful capabilities offered by RCAC's Negishi cluster.

Krzysztof Szymon Stopka, a research engineer and postdoctoral research associate working under Michael Sangid, professor of aeronautics and astronautics and director of the Advanced Computational Materials and Experimental Evaluation (ACME) Lab, is involved in a project focused on advancing the qualification of additively manufactured materials, particularly for structural applications in the aerospace field.

Sangid's research group combines knowledge of materials science, solid mechanics, and advanced manufacturing to solve complex problems in materials behavior and

processing. They employ physics-based computational modeling and design tools, which are experimentally validated and verified. The goal of this work is to improve our understanding and our tools for designing, processing, and "lifing" (testing the durability and expected useful life of) materials through simulation-based modeling of the material structure at the microscale, known as microstructure.

The researchers use Negishi to perform computational modeling to supplement their experimental test campaigns, which are indispensable but costly and time-consuming. Negishi's remarkable speed and reliability make it the preferred choice for these simulations, ultimately enhancing the cost-effectiveness and efficiency of the research.

"The computational demands of our research are immense," Stopka states. "Some simulations require nearly 200 CPU cores and can take weeks to complete. Given the volume of simulations needed for our work, RCAC's clusters are indispensable."

RCAC, College of Agriculture partner to reshape agriculture research

A team led by Purdue researchers has used the RCAC's Bell and Brown community clusters for a groundbreaking catalyst research and design paper recently published in Nature.

Jeffrey Greeley, the Charles and Nancy Davidson Professor of Chemical Engineering, and Zhenhua Zeng, a research scientist at Purdue, worked with researchers at Hunan University in China, Leiden University in the Netherlands and Clarkson University, to demonstrate the existence of a fundamentally new class of "active sites," spots on the surface of a catalyst where molecules bind and undergo a chemical reaction.

The findings not only offer valuable perspectives on prior studies in catalytically active sites, but also forecast the development of new catalysts with notably enhanced performance for hydrogen fuel cells and other electrochemical devices.

The use of heterogeneous catalysts in chemical reactions is widespread, with the understanding that high catalytic activity occurs only on specific surface sites, known as active sites. Determining the atomic structure of active sites is a central and longstanding goal of catalysis science.

Until the Nature paper led by Greeley and Zeng, these active sites were identified and classified through distinct surface motifs such as steps and terraces. This categorization often oversimplifies the complexity of active site identification, which can in turn lead to uncertain classification of active sites and incorrect

predictions of catalytic activity, which hinders catalyst design.

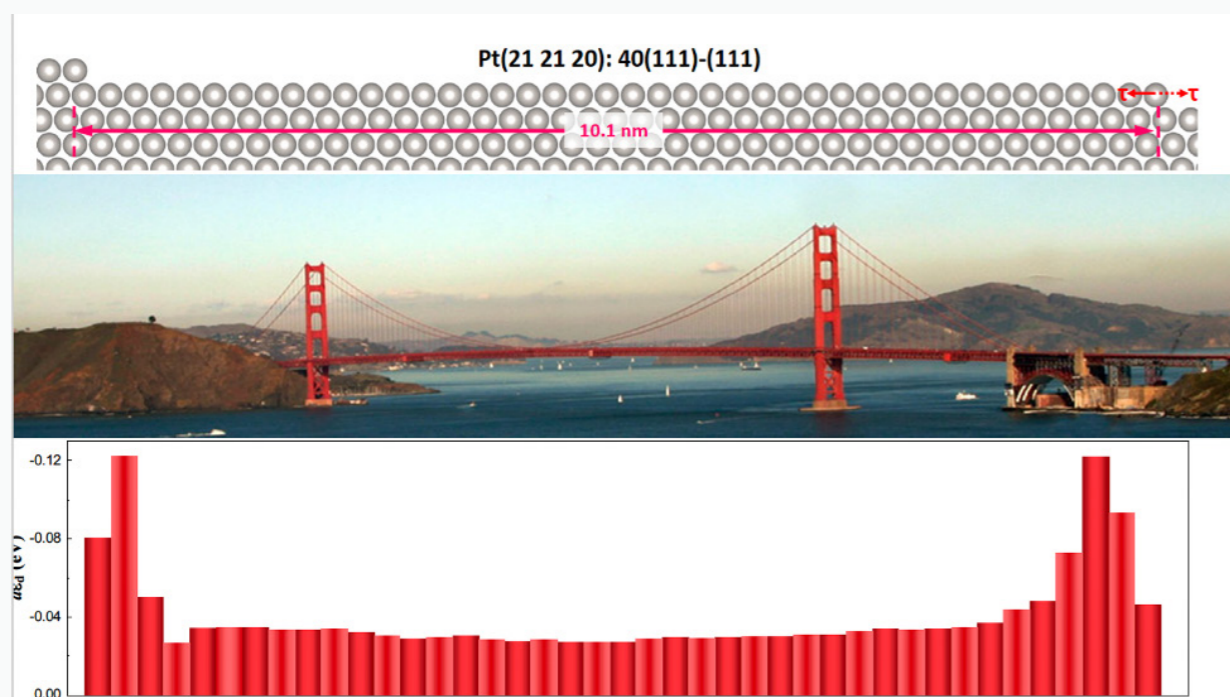
Using stepped Pt(111) [platinum] surfaces and the oxygen reduction reaction (ORR) in fuel cells as examples, the paper demonstrates that surface stress release produces inhomogeneous strain fields, leading to distinct electronic structures and reactivity for terrace atoms with identical local coordination. Furthermore, the terrace atoms surrounding the steps can experience up to 50 times enhancement higher than the atoms in the center of the terrace causing some areas of the terrace to experience higher or lower ORR activity.

The researchers concluded that the ability to control ORR reactivity through altering terrace widths or regulating external stress opens new doors for catalyst design.

The research team used RCAC's powerful Brown and Bell clusters for this work, and in the process simulated the largest systems ever published in this field. The calculations require an enormous amount of computational power, to the point that they wouldn't have been possible a decade ago even on a powerful supercomputer.

The design of clusters like Bell was crucial to accomplish the computational work efficiently.

"The new generation of the clusters with multiple CPUs per node, faster CPUs and better connection between the nodes to enable them to more efficiently work in parallel has been key to accomplishing the simulations in this paper," says Zeng.



857 Million CPU hours provided		1.9 Million GPU hours provided	
261 Community cluster PIs	3897 Community cluster users	66 Community cluster departments	
83,382 Community cluster OOD sessions	27% % of doctoral graduates using community clusters	97 RCAC Outreach events	
52 Students supported by CI-XP internships	73% % of Purdue research expenditures to RCAC users	33 RCAC Training sessions	
Sponsored research expenditures within RCAC			5 Million

PARTNER AWARDS AND SUCCESSES

Long time community cluster partner wins prestigious Purdue research award



Matthew Huber, the David E. Ross Director of the Purdue Institute for a Sustainable Future and a professor of earth, atmospheric and planetary sciences, was recently awarded the Herbert Newby McCoy Award for outstanding work in the natural sciences. Huber, who was a founding partner in the Rosen Center for Advanced Computing (RCAC)'s community cluster program when it launched in 2004, has owned nodes on every traditional RCAC cluster and has relied heavily on the computing power RCAC supplies over the course of his illustrious career.

Huber's research focuses on global modeling of past, present and future climate conditions and climate's impact on human settlements, managed landscapes and natural land, ocean and cryosphere ecosystems. His work is helping to explain the physical processes that generate tropical "thermostats" as well as the amplification of warming at the north and south poles, and the environmental, economic, ecological and evolutionary implications of these processes. He is specifically examining the human health and economic impacts under different future greenhouse gas emission scenarios.

In a paper recently published in the Proceedings of the

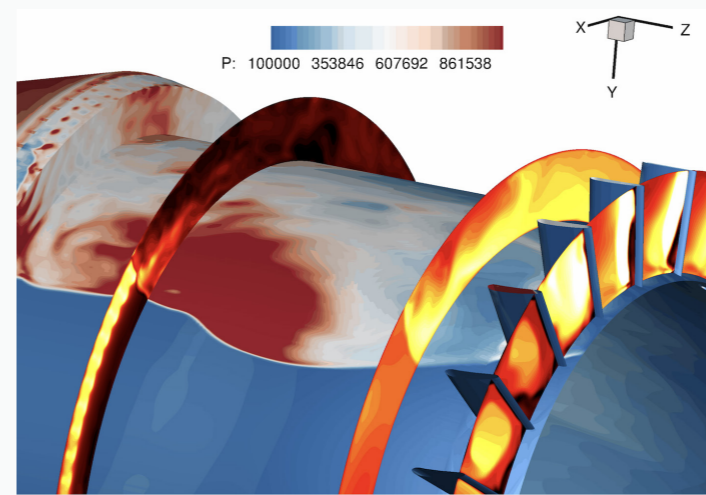
National Academy of Sciences, Huber and his collaborators used many years of past climate data to calculate future heat stress and its impact on the world. While the calculations used were not particularly novel and the climate data is publicly available, no one had been able to do anything like this previously because of the computing resources needed to download and process this massive dataset, which is estimated to be at least 300 terabytes.

"Due to the computational resources we have at RCAC, we're able to crunch the numbers at a scale that nobody else is," says Huber. "More accurate numbers lead to more accurate predictions for the future, which in turn lead to more accurate protections for the financial and social implications of climate change."

Another long term project of Huber's that uses RCAC's clusters involves paleoclimate research, studying how past warm climates can inform us about future warming. Each simulation takes many years to run and analyze, exceeding the duration of a typical grant. Huber says that having RCAC resources available has been critical for continuing these extended research projects and eventually turning them into publications and dissertations.

"We have great capabilities available through RCAC," says Huber. "In addition to the computing power, the fact that we can buy good storage on a fast network for a yearly fee is very useful. It's been very useful to me to have the tiered storage model with scratch file systems for work-area storage, Data Depot that you pay for annually as long as you need it, and then Fortress for storing longer term datasets."

Team led by Purdue professor receives industry award for work done on RCAC's Bell cluster



A team led by Purdue professor Guillermo Paniagua was honored with the "Best HPC Collaboration" award in HPCwire's 2023 Readers' and Editors' Choice Awards for its work developing novel engine components that will have implications for decarbonized power generation. The team used the RCAC's Bell supercomputer to run simulations.

"We are honored to receive this recognition from HPCwire, and it reflects the hard work and dedication of the many people involved in this multi-institution collaboration," says Paniagua, a professor of mechanical engineering, who moved to Purdue in 2014 primarily so he could pursue research on advanced turbines for clean power and propulsion at a faster pace. "This

award shines a light on Purdue's world-class experimental and computational resources, including our advanced turbine facilities tailored for pioneering small core clean aviation propulsion."

Paniagua, who is also the head of the Purdue Experimental Turbine Aerothermal Laboratory (PETAL), and his team are studying how to replace conventional combustors with hydrogen-fueled rotating detonation combustors, taking advantage of a combustor pressure increase to reduce the number of compressor stages. This could lead to smaller and lighter engines.

The team used Bell for various aspects of the project, including to:

- perform numerous computational fluid dynamic simulations of the combustor,
- optimize the design of the turbine, which involved running hundreds of direct evaluations of turbine geometries to maximize efficiency and reduce pressure distortion,
- and design the turbine's diffuser to ensure the correct temperature and pressure profile upstream of the turbine.

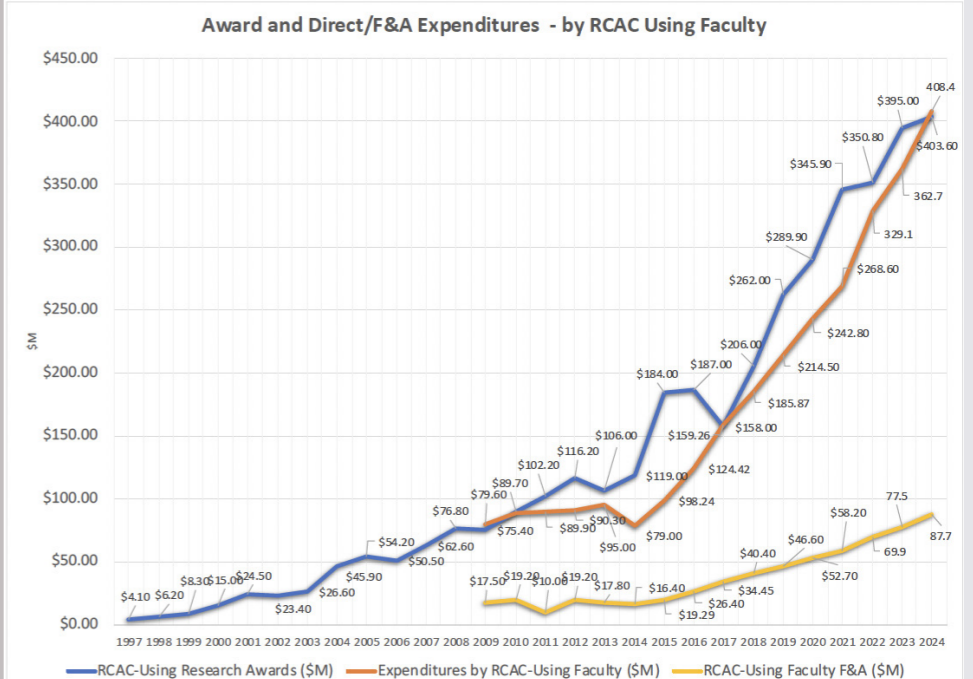
The Bell community cluster, which came online in 2020, debuted as number 431 on the list of the world's 500 most powerful supercomputers. Bell was the ninth Purdue community cluster to make the Top 500 list. Altogether, Purdue's community clusters provide nearly 10 petaFLOPs to drive research at Purdue.

Sponsored Research

Awards: RCAC-using faculty partners accounted for 67% of awards in FY24, for a total of \$403M. (From 60%/\$395 in FY23) This is a 7% increase in the total share to RCAC PIs.

Expenditures: RCAC-using faculty accounted for 73% of sponsor expenditures in FY24, a total of \$408.4M. (Up from 64%/\$352.7M in FY23) This is a \$55M increase in expenditures by RCAC PIs.

F&A: RCAC-using faculty paid \$88.5M of F&A to Purdue's general fund in FY24 – 78% of all F&A recovered. (Up from \$77.5M/88% in FY22)



EXTERNAL FUNDING

University collaboration receives \$6 million in funding for new CI workforce project

CI-PIVOT

The National Science Foundation (NSF) has awarded nearly \$6 million to Purdue, Indiana, and Mississippi State Universities to fund a collaborative program aimed at fortifying the cyberinfrastructure (CI) workforce within research communities. The new project, known as Cyberinfrastructure Professionals InnoVating and

broadening the adoption of advanced Technologies (CI PIVOT), is funded under the NSF's Strengthening the Cyberinfrastructure Professionals Ecosystem (SCIPE) program. SCIPE's goal is to advance the Cyberinfrastructure Professional (CIP) workforce throughout the nation. CI PIVOT will help to accomplish this by recruiting individuals with domain-specific expertise and training them to be CI facilitators, making advanced CI accessible and effective for all.

Purdue awarded two NSF grants to enhance networking for research



The NSF's Office of Advanced Cyberinfrastructure has recently awarded Purdue University two new Campus Cyberinfrastructure (CC*) grants. The two awards, totaling nearly \$1.5 million, will directly address and enhance the cyberinfrastructure at Purdue University. CC* awards are given to teams who are seeking to partake in coordinated campus-level cyberinfrastructure improvements, innovation, integration, and engineering for science applications and distributed research projects.

The first award is for a project titled "Enhancing network connectivity for data-intensive, multi-institution collaborative science," and Arman Pazouki, the Director of Scientific Application at RCAC, is the Principal Investigator (PI). This project will enhance Purdue's external connectivity to the world, delivering fast networking for data-intensive research within Purdue, and extending connectivity to under-served research facilities. It will accomplish this by connecting Purdue to Indiana GigaPOP and the ESnet network with a 400 Gbps wide-area network. The project will also expand Purdue's science DMZ to enable high-bandwidth, low-latency communication required for real-time video analytics and public-private partnerships within the

Purdue Discovery Park District. Several educational and workforce development activities, including formal training and mentorship, are included in the project to engage undergraduate students in the deployment and operation of network infrastructure and support experiential and residential learning programs on Purdue's campus.

The second award is for a project titled "Programmable Network Testbed for 400 Gbps Science DMZ." It is led by Professor Sanjay Rao of the Elmore Family School of Electrical and Computer Engineering (ECE). Erik Gough, a Senior Research Scientist at RCAC, is a co-PI for the project and will lead the testbed deployment and its integration into RCAC's network infrastructure. Professor Vishal Shrivastav of ECE and Professor Muhammad Shahbaz of the Department of Computer Science will also serve as co-PIs on the project. The network integration project will design, implement, and deploy PRONET, a testbed to support cutting-edge programmable networking research for managing Science DMZ traffic. This project will also aid in workforce development by involving students—both graduate and undergraduate—and the PIs will work to actively recruit students from underrepresented groups.

Anvil receives funding for national pilot project aimed at advancing AI



Purdue University's Anvil supercomputer received \$4.9 million in supplemental funding from the NSF in order to support the newly launched National Artificial Intelligence Research Resource (NAIRR) Pilot. The NAIRR is an NSF project aimed at creating a national infrastructure that connects U.S. researchers to responsible and trustworthy Artificial Intelligence (AI) resources. The NAIRR will provide these researchers equitable access to the data, software, training, computational, and educational resources needed to advance research, discovery, and innovation within the field of AI. The NAIRR Pilot is the first

step towards realizing the full NAIRR vision. Its purpose is to demonstrate the NAIRR concept and advance its primary goals of spurring innovation, increasing diversity of talent, improving capacity, and advancing safe, secure, and trustworthy AI in research and society. By being part of the NAIRR Pilot program, Anvil will directly support research projects that focus on testing and validating AI systems, improving model performance, increasing the interpretability and privacy of learned models, reducing vulnerability to attacks, and assuring that AI functionality aligns with societal values and obeys safety guarantees.

Anvil REU program receives additional funding



The Anvil Research Experience for undergraduates (REU) summer program was awarded \$64 thousand in supplemental funding from the NSF. After two years, the Anvil REU program has proven to be tremendously successful, providing undergraduate students with hands-on HPC experience and directly impacting

workforce development within the field of HPC. As such, the NSF awarded this supplemental funding to support the continuation and growth of the program, allowing for an increase in the total number of participants. This supplement enabled eight undergraduates to work full-time on-site at Purdue University for a 11-week period where they collectively accomplished four separate Anvil-related projects (two students per project.) During the program, the students were able to travel to the PEARC conference, where they not only shared about their work, but also all secured future internships at a national research laboratory. Six of the REU participants will also be continuing their work in student positions with RCAC during the fall semester.

EDUCATION AND OUTREACH

Summer camp fun: RCAC teaches high schoolers about cybersecurity and coding



The Rosen Center for Advanced Computing recently hosted two summer camps aimed at high schoolers, with the hopes of giving them an introduction into the college experience by providing them the ability to earn college credit, explore potential majors and experience campus life.

The first camp, CyberSafe Heroes: A Week of Cybersecurity Mastery, focused on cybersecurity best practices and career pathways. Students participated in encryption challenges, ethical hacking simulations, cybersecurity escape rooms, online safety workshops, and engaging career panels with cybersecurity professionals.

The second camp, Code Explorers: Coding and Environmental Discovery, focused on creating an

immersive introduction to coding, connecting it to environmental science. Students were able to code with microcontrollers, conduct data analysis with Python, create environmentally themed games and more.

Both camps were lauded as great successes by both students and staff members alike, bringing together a diverse and vibrant community of enthusiastic students and dedicated volunteers in order to create an atmosphere of excitement and learning. The meticulously planned activities and smooth execution ensured every participant had a memorable and enriching experience.

"I had an incredible time at Purdue University's CyberSafe camp," says camper Kanishk Thamman. "I explored cybersecurity topics like password cracking and also used Purdue CyberTap's cyber range to conduct a scavenger hunt! For my final project, I presented on AI's impact on cybersecurity, showcasing my passion to peers and mentors. I'm excited for more opportunities like this in the future."

Plans are already underway to make these camps annual events. The undeniable enthusiasm and joy it brought to both students and volunteers has energized the organizers to take it to the next level. Future iterations will build on this year's success, introducing new and exciting activities that will continue to inspire and engage participants.

RCAC hosts outreach event for local high schoolers

RCAC recently hosted an outreach activity for the Southport High School Engineering, Computer Science, and Robotics students and the school's Girls Who Code group. During the event, the high schoolers learned about the Anvil supercomputer as well as RCAC's Envision Center, where innovation and imagination meet research and education.

On Saturday, April 27th, Southport High School brought 24 students to Purdue University's main campus in West Lafayette to tour RCAC's facilities. The students, most of whom had never before seen a supercomputer, were able to experience a walkthrough of the data center to

see the Anvil supercomputer in person. Keegan Adams, a student employee within the CI-XP program, led the Data Center tour alongside Luke Monroe, an Associate Research Solutions Engineer for RCAC. After splitting everyone into two groups of 12, Adams took each group through the Data Center, showcasing the different supercomputing clusters housed on campus. The tour's main focus was the NSF-funded Anvil system, Purdue's most powerful supercomputer. The students were able to see the supercomputers in action—hearing the roar of the systems, feeling the heat radiate between the aisles—all while Adams discussed the specifications of Anvil and related personal stories from his time spent helping to

build the cluster during his freshman year at Purdue. The students were even able to open Anvil's doors and take a peek inside at the configuration and build.

Next up on the tour was RCAC's Envision Center. Adams first gave a presentation on Anvil and its capabilities and showcased some of the work completed on the supercomputer. He then turned the microphone over to Amanda Warren-Glowe, Program Manager for the Envision Center. She spoke with the Southport High School students about how Envision enhances research and education outcomes through scientific visualization, virtual and augmented reality, and media creation. Warren-Glowe shared insight into some of Envision's past projects, highlighting the various ways in which complex research concepts can be effectively communicated. The students were then able to experience some of Envision's work first-hand by demoing the newly developed Collab XR platform, part of the XR Lab at the center. This platform is a shared environment that allows anyone in a headset to view and interact with the same virtual content pieces together, viewed in fully virtual (the entire field of vision is virtual, obscuring all of the real environment, and other participants appear as virtual avatars) or passthrough augmented reality (cameras feed the real room and people to the headset view, overlaying virtual content to appear to exist in the same space). Upon wearing the XR headsets, the students could view supernovas

in intricate detail, pass objects amongst themselves, and, importantly, not collide with each other during the experience. Warren-Glowe closed out the students' trip by answering any questions they had about the Envision Center and its capabilities.



RCAC Hosts 2nd Annual Cyberinfrastructure Symposium on AI



The Rosen Center for Advanced Computing held its second annual "Cyberinfrastructure and AI" symposium on October 12. The event featured keynote speaker Paul Bloch, co-founder of DataDirect Networks (DDN), who discussed the future growth opportunities catalyzed by AI. Plenary speakers included NVIDIA's Emmett Kilgariff, who highlighted AI's impact on supercomputing, and Beth Plale from Indiana University, who emphasized AI ethics and democratization.

RCAC leaders Preston Smith and Carol Song presented on Purdue's campus cyberinfrastructure advancements. George Takahashi showcased how the Envision Center's extended reality (XR) technologies are enhancing learning across various disciplines.

The symposium concluded with a reception sponsored by DDN, AMD, and Piers Group, and Purdue faculty members, including Daniel Aliaga, Alex Psomas, and Ziran Wang, also presented their research.

The event was a significant platform for connecting researchers, students, and industry partners, reinforcing RCAC's vital role in Purdue's research efforts. With less than 9% of Purdue IT's budget, RCAC supported faculty responsible for 63% of FY23 research expenditures (\$363 million). The community cluster for machine learning and AI research, Gilbreth, now has 148 faculty partners, more than double from last year.

STUDENT NEWS

RCAC student employee successfully defends Masters Thesis



Yiqing Qu, a Graduate Research Assistant at RCAC, recently obtained her Master of Science (MS) degree in Computer Information and Technology. Her MS thesis was related to the work she conducted at RCAC, which ensured that there is a way to measure and improve adherence to the FAIR principles of scientific data management.

Qu began working on her research project in September of 2022, when she first joined RCAC. Her work is part of the GeoEDF (Extensible Geospatial Data Framework), an NSF-funded project with the goal of providing seamless connections among platforms, data, and tools and making large scientific and social geospatial datasets directly usable in scientific models and tools. Part of the GeoEDF project is to develop a resource data management portal that allows researchers to then publish their workflows and their results, allowing other colleagues to reproduce their work. A key strategy in ensuring such reproducibility is to adopt and adhere to the FAIR data principles. The FAIR data principles are guidelines researchers can (and should) use to enhance the reusability of their data—specifically, to enhance the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. Without the use of FAIR principles, it is incredibly difficult for scientists to find and utilize the most relevant data for their research.

The problem with FAIRness comes from its implementation. Research Software Engineers had no clear guidelines for adhering to the FAIR principles when creating a new data repository from scratch. So, in

theory, incorporating FAIRness into data portals and data sets is a wonderful idea that everyone should start doing immediately. In practice, it is more akin to conquering the Chimera—or at least it was.

After discussing the idea with Rajesh Kalyanam, a Senior Research Scientist for RCAC and Qu's mentor for her project, and her academic advisor, Dr. Baijian Yang, the Associate Dean for Research and a Professor at the Purdue Polytechnic Institute, Qu decided to tackle the problem of FAIRness for her Master's thesis using the GeoEDF project as the functional basis of her research. The ultimate goal of her project was to create a methodology for evaluating FAIRness and to develop a structured approach to implementing improvements that would lead to the creation of FAIR-compliant data portals. If successful, researchers could then use her work to build their own data repositories that adhere to the FAIR principles.

"For working on my Master's thesis," says Qu, "I wanted to solve two questions: How to evaluate the FAIRness of the project, and how to improve the FAIRness of the project? I decided to spend two years working on these research questions, and chose this particular project due to its real-world applications and its close alignment with my personal research experience."

Starting with a bare-bones Django portal framework built by the Globus team, Qu needed to:

- Evaluate the FAIRness of the data portal.
- Implement new features on the portal to improve the FAIRness score.
- Re-evaluate the FAIRness score of the improved data portal and iterate.

Once she decided to use the Globus portal framework as a starting point, Qu began by deploying the portal on the Anvil Composable Subsystem, a Kubernetes-based private cloud managed with Rancher that provides a platform for creating composable infrastructure on demand. Anvil is an NSF-funded shared computing resource and Purdue's most powerful supercomputer. By using Anvil, Qu ensured that her work would be immediately available to researchers nationwide. Qu then moved on to the first FAIRness evaluation. She tested three separate evaluation tools and determined that a tool known as F-UJI was best for the project. Qu used F-UJI to score the FAIRness of the barebones Globus framework, which received a score of 47%. For comparison, Qu chose a well-known, mature, FAIR-compliant platform, that had diverse data types and

well-designed metadata, to test against. The platform, known as HydroShare, scored a 64% in the FAIRness evaluation. Now that Qu had a target score to aim for, she began to implement new features to design her new data repository, named the GeoEDF Data Portal.

Aside from simply giving a FAIRness score, F-UJI also provided feedback on what improvements could be made. Qu looked at the feedback and prioritized the features that would lead to the greatest overall impact on the score. She then systematically added new features to the portal, increasing its FAIRness score from 47% to 60—a huge improvement that puts the portal on par with HydroShare in the FAIRness assessment. Needless to say, Qu and her mentors were thrilled with this result. Qu is now working on a paper based on her work and hopes to present this to other portal developers and researchers who are involved in building similar data portals.

"We are beyond excited by what Yiqing was able to do while working with us," says Kalyanam. "You would be hard-pressed to find anyone who could complete so much in such little time, yet she did it while going to graduate school. She is one of the hardest working students I've ever had the pleasure of working with, and it was a joy to have her as part of the team."

Now that Qu has successfully defended her thesis and graduated, she will be transitioning to her new job at Klaviyo, a company that provides intelligent marketing automation powered by customer data. In her new role, Qu will be working on Klaviyo's real-time data pipeline, facilitating the ingestion, processing, and movement of data points that power Klaviyo's core functionalities.

"For working on my Master's thesis, I wanted to solve two questions: How to evaluate the FAIRness of the project, and how to improve the FAIRness of the project? I decided to spend two years working on these research questions, and chose this particular project due to its real-world applications and its close alignment with my personal research experience."



Industrial engineering students analyze RCAC cluster data for senior capstone project

For their senior capstone course IE431, a group of recent industrial engineering graduates worked with RCAC staffers to analyze data from RCAC systems and build an interactive dashboard aimed at reducing system downtime.

Working with Kyle Purple, RCAC's director of integrated cyberinfrastructure, and Ashish, senior research data scientist, the students, Naphatsakorn Simasathapornkul, John Duffy, Mihir Mahale, Avery Mantyla, and Hao Zhuo, all recent graduates in industrial engineering, developed equipment monitoring techniques to reduce downtime for RCAC clusters and other systems.

System downtime disrupts the productivity of faculty and researchers, and leads to inefficient and costly maintenance efforts. RCAC was looking for a way to proactively identify and predict factors leading to node failures, thus preventing downtime before it happens.

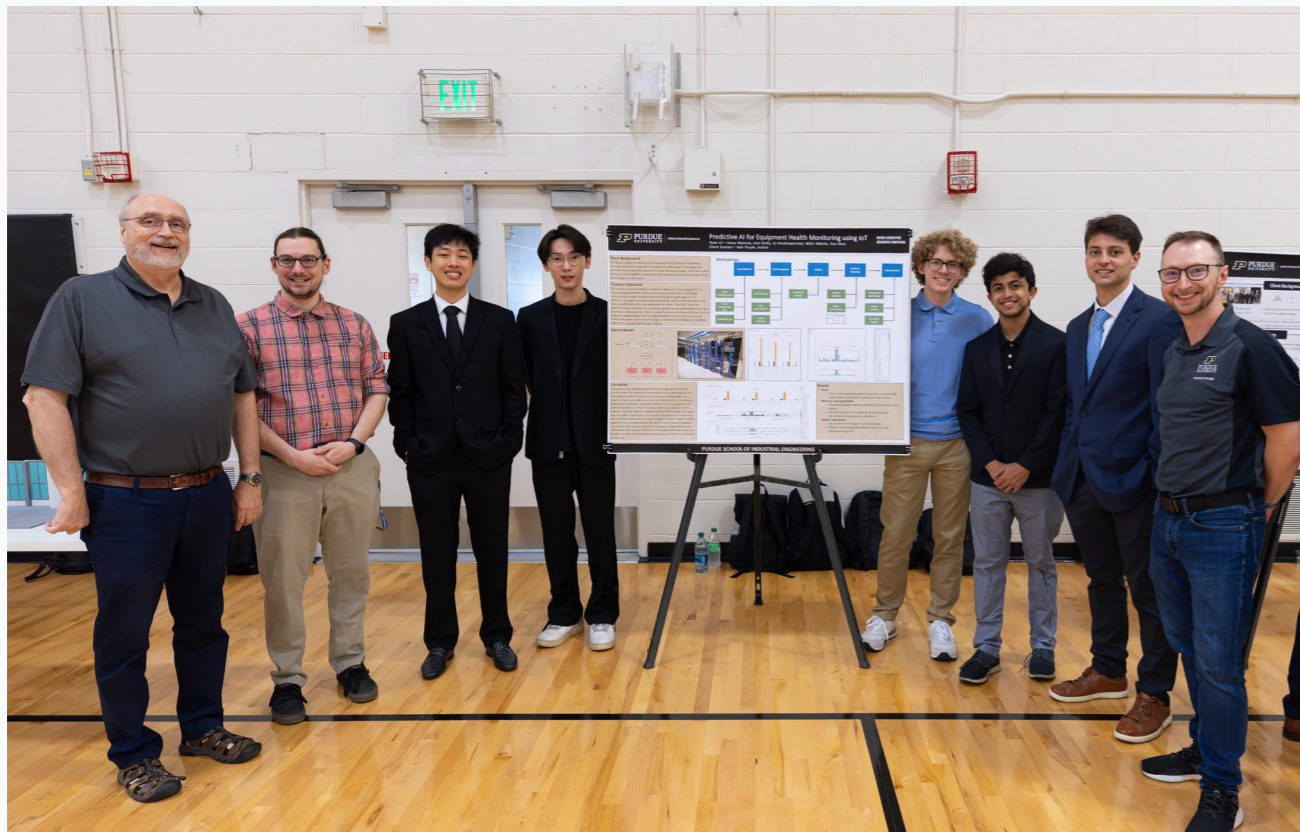
The students used several machine learning models to process and analyze large volumes of data and chose the model that best predicted the node's downtimes. They then created an interactive, visual dashboard that facilitates data-driven operational decision-making and provides customizable views for identifying trends and patterns in the data.

"The work they put in, leveraging machine learning models and dashboards, laid a foundation we can build upon in the future," says Purple. "I'm impressed with what they were able to do in such a short period of time."

The students presented their work to the RCAC team before sharing with a wider audience at the industrial engineering end-of-semester event, where more than 35 IE431 projects were showcased.

RCAC has a robust student employment program, CI-XP (Cyber Infrastructure-eXperience) with opportunities for student workers in the data center, the Envision Center, the Scientific Solutions Group and the Anvil REU programs. The purpose of the CI-XP program is to provide work opportunities and real workplace experience that enhance the student's education through the development of professional skills, responsibilities, habits, attitudes, self-confidence and self-development. RCAC staff serve as mentors to the students.

"We're always looking for ways to be more proactive at engaging students in our operations," says Purple. "Working with this group of IE students to aid in the efforts to reduce system downtime was a tremendous opportunity, and one I hope to continue in the future."



RCAC students shine in summer internships across the nation

Last summer, RCAC students ventured far and wide to gain invaluable professional experience at various companies throughout the nation. Their diverse journeys have showcased their talents and highlighted the significant role RCAC plays in preparing students for real-world challenges. Here are some compelling stories from our students:



Dhruv Sujatha: Hewlett Packard Enterprise, North Carolina: Dhruv Sujatha spent his summer interning with Hewlett Packard Enterprise (HPE), working on the primary storage team. His role was to develop a tool that would analyze the performance of NimbleOS and FleetOS, HPE's upcoming object storage platform. Sujatha's project involved creating a Signature Collection and Log Analysis tool to streamline bug triaging by parsing logs and extracting relevant information. This significantly reduced debugging time from days to minutes, akin to a smart compiler for an entire operating system.



Karl Oversteins: Raytheon Missiles & Defense, Arizona: Karl Oversteins ventured to Tucson, Arizona, for his internship with Raytheon Missiles & Defense. Working with the classified infrastructure HPC team, Oversteins made impressive strides. He pioneered the integration of Open OnDemand in separate environments, initiated the organization-wide adoption of RHEL 9, and successfully started and troubleshooted clusters for two new programs.

Aaron Ni: CACI International Inc., Indiana: Aaron Ni's internship took him to Crane, Indiana, where he worked

as a systems integration engineer intern for CACI International Inc., a defense contractor. Despite coming from a software engineering background, Ni dove into the world of systems integration for counter-unmanned aerial systems (CUAS). His tasks included integrating sensor suites like radar and EO/IR, developing system requirements, and conducting system operability verification testing.

Prachi Modi: Tesla Inc., Nevada: Prachi Modi began her internship as a data engineering intern with Tesla's quality data and systems team. She successfully resolved legacy server airflow issues to migrate directed acyclic graphs to a new server with zero downtime and facilitated smooth ownership transfer. Additionally, Modi played a pivotal role in preparing for ISO 9001 audits, improving documentation, and adhering to the high industry standards.



These stories exemplify the versatility and preparedness of RCAC students, showcasing their ability to excel in diverse roles and industries. Upon their return to campus, they brought back not only valuable experiences but also a broader perspective on their academic and professional journeys. We applaud their accomplishments and eagerly anticipate the achievements that lie ahead.

STAFF NEWS

The Rosen Center for Advanced Computing has experienced astounding growth across all the teams in our organization throughout the past year. In order to support your computing and data needs, we have made strategic hires to expand our depth of expertise in key fields, such as artificial intelligence, machine learning, and research software engineering.

Integrated Cyberinfrastructure Team:

This HPC Engineering Team designs, builds, and maintains Purdue University's integrated cyberinfrastructure (CI), which includes eight high-performance computing systems, each with scratch storage, secure, archival, and large-scale data storage solutions, as well as high-speed networking, advanced campus CI facilities, and our secure Weber cluster for controlled and unclassified information. Their engineers deliver cutting-edge technology solutions to support Purdue researchers, including the newly launched Gautschi cluster, and researchers nationwide through the NSF-funded Anvil cluster. The Integrated CI team has seen the addition of six new members:

- *Kyle Purple, Director of Integrated Cyberinfrastructure*
- *Frank Bakhit, Research Solutions Engineer*
- *Taylor Graham, Associate Solutions Hardware Engineer*
- *David LeFevre, Research Solutions Engineer*
- *Chris Orr, Senior Research Solutions Engineer*
- *Alex Younts, Principal Research Engineer*

Scientific Applications Group:

The Scientific Applications Group is responsible for a wide variety of tasks within RCAC, including user support, training, documentation, research consultation, and project partnership, all with the goal of helping researchers adopt and efficiently utilize HPC. In short, they aim to increase access to research computing resources and make high-end research easier to achieve. The Scientific Applications group has seen the addition of four new team members:

- *Michael Carlson, Senior Computational Scientist*
- *Charles Christoffer, Senior Computational Scientist*
- *Haniye Kashgarani, Senior Computational Scientist*
- *Tomas Ratkus, Senior Research Project Manager*

Research Software Engineering Center:

The Purdue Center for Research Software Engineering (RSE Center) is a newfound center dedicated to helping accelerate research and increase its impact through the creation of innovative, robust, and sustainable research software. The RSE center is composed of two teams, the Scientific Solutions Group (SSG) and the Envision Center (EC), with additional support provided by staff members from across RCAC. The SSG group develops innovative cyberinfrastructure solutions to help advance scientific discovery and education, and enable greater research reproducibility and impact at Purdue and beyond. The EC provides novel solutions to effectively communicate complex research concepts. Computer graphics, advanced visualization, auditory (sound), haptic (touch), and multimodal interaction integrate with state-of-the-art high-performance computation to assist researchers, instructors, and leaders in their quest for new knowledge and innovative products. The RSE Center has seen the addition of three new team members:

- *Daniel Madren, Senior Research Development Administrator, Center Operations and Visualization*
- *Noah Oller Smith, Research Software Engineer, Scientific Solutions Group*
- *Julian Triveri, Associate Research Software Engineer, Envision Center*

Communications Team:

The Communications team handles all internal and external marketing and communication tasks for RCAC. Through comprehensive articles, seminars and symposiums, and social media engagement, the Comms team showcases to Purdue and beyond the world-class research being conducted on Purdue's HPC systems, as well as highlights and promotes the benefits of HPC. The Communications Team has added one new team member:

- *Jonathan Poole, Technical Writer*

Fields of Expertise:

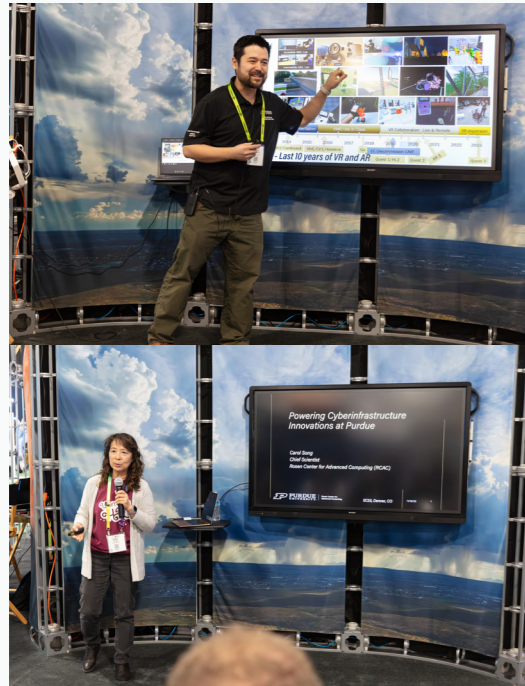
With the addition of these individuals, we have helped our organization to accumulate and cultivate even more expertise in a variety of fields. These fields include:

- *Artificial Intelligence*
- *Machine Learning*
- *Software Development*
- *Cybersecurity*
- *Systems Engineering*
- *Climate and Geosciences*
- *Life Sciences*
- *CI Project Management*
- *Marketing and Communications*
- *HPC in Cloud*
- *Geospatial Data*
- *LINUX*



RCAC STAFF PRESENT AT NATIONAL CONFERENCES

SC23



RCAC hosted a booth at the SC23 Supercomputing conference, which took place in Denver, Colorado, from Nov. 12-17. The Purdue booth highlighted the recently announced "Purdue Computes" initiative.

Purdue Computes is a university-wide, multi-pronged initiative to meet rising student interest in computing-related majors and the growing societal impact of AI and semiconductor chips. Over the next five years, Purdue plans to hire 50 new faculty in the department of computer science, hire another 50 faculty who will be associated with the new Institute for Physical AI, and build \$100 million worth of semiconductor facilities.

With nearly 10 petaFLOPS of HPC and 57 PF of AI compute resources available to Purdue researchers, RCAC supports computational faculty from over 60 departments across every academic college and all Purdue campuses.

The SC23 conference exhibit hall featured numerous presentations hosted at the Purdue booth (#1525), as well as presentations from RCAC staff in other booths.

As an international conference, SC23 attracted more than 400 exhibitors and more than 12,000 attendees from around the world who come to see the latest innovations in high-performance computing and related fields.

The goals of the SC conference revolve around fostering collaboration, sharing research and innovations in high-performance computing, promoting discussions on challenges and advancements in the field, and providing a platform for networking among professionals and researchers in the global HPC community.

PEARC 23

RCAC staff presented posters and papers and led workshops at the Practice and Experience in Advanced Research Computing (PEARC) conference held in Portland, Ore. in July 2023.



The PEARC conference is held to provide a forum for discussing challenges, opportunities and solutions among the broad range of participants in the research computing community. This community-driven effort builds on successes of the past, and aims to grow and be more

inclusive by involving additional local, regional, national, and international cyberinfrastructure and research computing partners spanning academia, government and industry.

Several RCAC staff members were also on the conference program committee.



RCAC participates in the Minority Serving - Cyberinfrastructure Consortium annual meeting



Suzanna Gardner, Senior Research Operations Administrator of Outreach and Engagement for the Anvil supercomputer at RCAC, and Laura Theademan, Director of Center Operations and Visualization at RCAC, recently presented at the 2024 Minority Serving - Cyberinfrastructure Consortium (MS-CC) Annual Meeting. The pair presented alongside Dameion Brown, the Director of Information Security & Chief Information Security Officer in the Division of Information Technology at Jackson State University, on the topic of "Embracing Diversity in Technology Leadership."

The Minority Serving - Cyberinfrastructure Consortium (MS-CC) is a community of information technology (IT) professionals, campus leaders, faculty members, researchers, and students from across the nation who have come together to improve cyberinfrastructure (CI) capabilities at historically Black colleges and universities, tribal colleges and universities, Hispanic-serving institutions, and minority-serving institutions. The goal of MS-CC is to promote advanced CI for teaching, learning, research, and enterprise technology capabilities at these historically under-resourced institutions. This year, the second annual meeting took place in Washington, D.C., from May 29th through the 31st. At the meeting, MS-CC members gathered to discuss sustainable campus-level IT capabilities for data-intensive education and research programs.

During their presentation, Brown, Gardner, and Theademan discussed the importance of embracing and recognizing diversity within technology leadership on campuses. In their view, diversity not only includes women and traditionally underrepresented minority groups, but also those with a non-technical background.

The trio highlighted the importance of diversity and inclusion within CI (innovation and creativity, better decision making, etc.) as well as the challenges faced by underrepresented leaders in the field and strategies for overcoming these challenges. Some of these strategies include hosting mentorship and sponsorship programs, actively addressing bias and discrimination, and creating a safe space for open dialogue. The group also noted that the MS-CC Annual Meeting is exactly the type of event that can help overcome equity challenges.

"I am grateful for the opportunity to present alongside Laura and Dameion," says Gardner. "This experience has been eye-opening, highlighting the absolute need for inclusion. I came back with the stark realization: diversity is inevitable. Due to immigration, migration, technology, and the blending of socio-economic, religious, cultural, and racial backgrounds in our workplaces and communities, diversity will continue to shape our society. We must look ahead and ask ourselves: How can we achieve genuine inclusivity? How can we bring individuals from various backgrounds into innovation and decision-making spaces? How can we cultivate cultural sensitivity in our interactions? The diversity train is unstoppable. We must harness and celebrate our differences, inviting change into our organizations. By doing so, we not only adapt but thrive in an ever-evolving world."

The 2024 MS-CC Annual Meeting was a widely attended event, allowing for excellent discourse on improving campus CI on underrepresented campuses. Outside of presenting, Gardner and Theademan were able to make new connections and discuss how Anvil can help support minority-serving institutions. Some of these discussions have already led to the potential for exciting new projects and collaborations.

"We are thrilled to discover new ways Anvil can broaden CI adoption at historically underserved institutions, including MS-CC member institutions," says Gardner. "Our innovative ideas are already taking shape, and we look forward to presenting them at next year's annual MS-CC meeting."



WHPC HOSTS GENDER BIAS WORKSHOP, HACKATHON, AMONG OTHER EVENTS

Purdue's WHPC group, which is part of a broader engagement initiative by RCAC and is led by RCAC staff members, has had a busy year, hosting several events for the Purdue community.

Workshop on gender bias, stereotypes

WHPC hosted a workshop on Feb. 9 about the "default male" phenomenon and the pervasive influence of gender stereotypes based on the book "Invisible Women: Data Bias in a World Designed for Men" by Caroline Criado Perez. The event featured presentations from four Purdue faculty members, followed by an audience question and answer session.

"I came to the workshop because I'd previously read the book 'Invisible Women,' so I thought it would be interesting to hear other people's interpretations of it. Each of the presenters sparked a new and instructive conversation about the challenges of a world built around one single default identity, and the struggles of people who don't have access to that default," says George Lyle, IT security risk analyst.

"Attending the Purdue WHPC workshop was an enlightening experience that broadened my perspective on the nuances of gender in both

professional and personal spheres," adds Sruthi Dasika, a doctoral degree student in environmental and ecological engineering.

Inaugural InnovateHer Hackathon

RCAC and WHPC recently participated in the inaugural InnovateHer Hackathon, a weekend-long event at Purdue University aimed at fostering inclusivity in the field of technology.

The hackathon began at 8 a.m. on Saturday, March 2nd, and continued until 4 p.m. on Sunday, seeing teams work on their projects throughout both days and overnight.



On Saturday, RCAC's Suzanna Gardner, Taylor Graham, and Keegan Adams led a Networking 101 workshop on behalf of Purdue's WHPC group, of which they are members. The workshop showed students ways to develop their professional network and included a hands-on resume-building segment. The response to the workshop was great, with the participants learning to make impactful connections, leverage their strengths, and capture opportunities.

On Sunday, RCAC hosted a high-performance computing workshop, led by Eric Adams and Geoffrey Lentner. The workshop was open for HPC newbies and pros alike, with discussions ranging from what HPC is and why you would want to use it to more advanced topics, such as how to construct specific lines of code on a shared-computing resource.

Importance of Women in STEM presentation

WHPC recently hosted a presentation surrounding Rebecca Sharples's Doctor of Education (Ed.D) dissertation, which was inspired by the remarkable life journey of Suzanna Gardner, the Senior Research Operations Administrator of Outreach and Engagement for Anvil and a member of WHPC.



On April 26th, Gardner and Sharples gave a dual presentation at RCAC's Envision Center. During the presentation, the two discussed Gardner's early encounters with the entrenched biases of a society steeped in misogyny, how her path led to the pair becoming friends, and how this culminated in Sharples' decision to use Gardner's life story as a platform for her dissertation.

The pair then spoke about some of the specific research that went into Sharples' dissertation, namely the gross inequalities women face, not only in this country, but throughout the world. The statistics surrounding this topic tied into the unifying theme of the presentation: The Importance of Women in STEM.

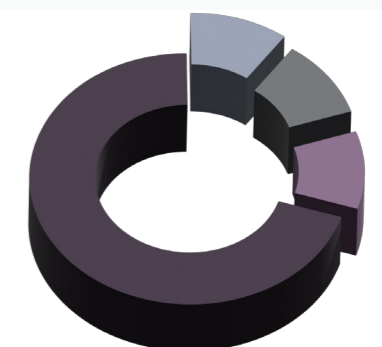


Later in the presentation, Gardner and Sharples discussed the barriers and challenges that women face, specifically when it comes to getting more women in STEM. The day then concluded with a positive and thought-provoking Q and A session. If you weren't able to make it in person, you may watch the full presentation here:



Scholarships

The Rosen Center for Computing (RCAC) at Purdue University has successfully provided partial travel assistance scholarships to undergraduate and graduate students identifying as women, fostering their participation in national conferences related to advanced computation.



- Grace Hopper
- SIGGRAPH
- SC23
- Society of Women Engineers

WHPC AWARDS CONFERENCE TRAVEL SCHOLARSHIPS TO WOMEN STUDENTS

Over the last year, Purdue's Women in High-Performance Computing (WHPC) group has awarded scholarships to enable women Purdue students to attend a variety of conferences related to computing, engineering and science.

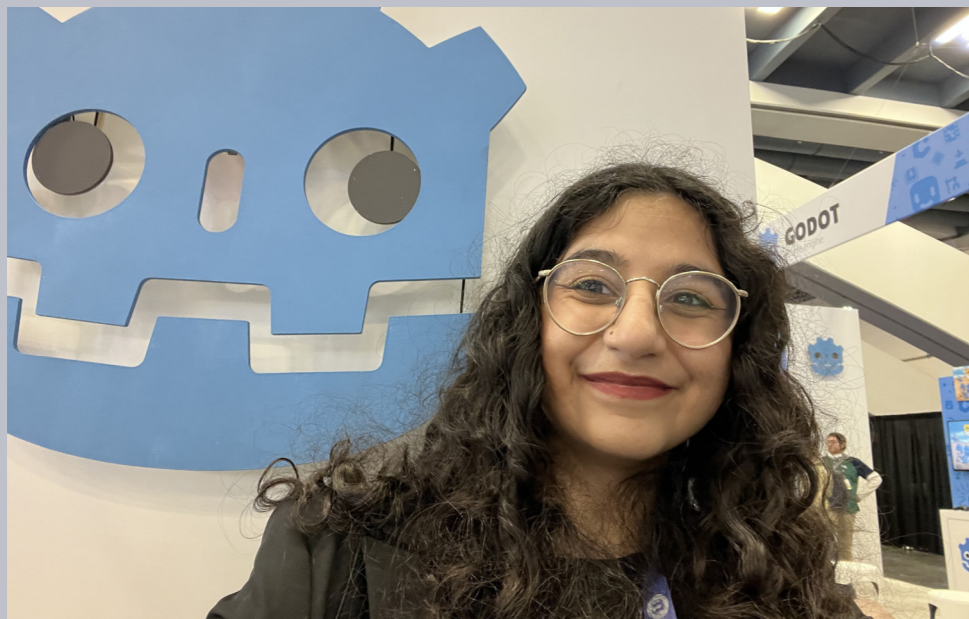


Jeanna Blake, a fourth year doctoral student in the College of Veterinary Medicine whose research focuses on canine genetics, attended the 12th International Conference on Canine and Feline Genetics and Genomics (ICCFGG) in Helsinki, Finland with WHPC support and was honored as having one of the three best posters at the conference by the Scientific Organizing Committee. Blake was also one of just fifteen poster authors invited to give a lightning talk about her poster.

"The WHPC scholarship is empowering and investing in the next generation of female scientists, like me, by allowing women to attend key conferences in their field that they might not be able to go to in order to advance their career opportunities," says Blake, who adds that she might not have been able to attend ICCFGG without the WHPC support, since travel to international conferences can be cost-prohibitive.

WHPC scholarships also supported student travel to the 2023 Grace Hopper Celebration of Computing, the 2023 Society of Women in Engineers conference, the 2023 Game

Developers Conference and Supercomputing (SC23), among others.



"Attending GDC was amazing and inspiring" says Uma Claire Teelock, a junior in game development and design and animation and visual effects who also works as a student artist at the Envision Center, of her conference experience at GDC, the game industry's premier event that included nearly 30,000 attendees.

"I had the opportunity to meet like-minded people and hear the lessons others have learned."

"The diverse range of tutorials and workshops on HPC, deep learning and cloud computing have expanded my understanding of current trends and best practices in my field, providing me with a deeper insight into the intricacies of the industry," says Sathvika Kotha, a master's degree student in computer and information technology who was sponsored by WHPC to attend SC23.



The WHPC program is part of the Purdue Information Technology broader engagement initiative to expose and encourage women in the Purdue community to pursue research and careers in HPC and technology fields.

To learn more about WHPC travel scholarships, please visit:



To learn more about Purdue Women in HPC or be added to the mailing list, contact



20
24

High-performance computing at
the highest proven value