SDSC

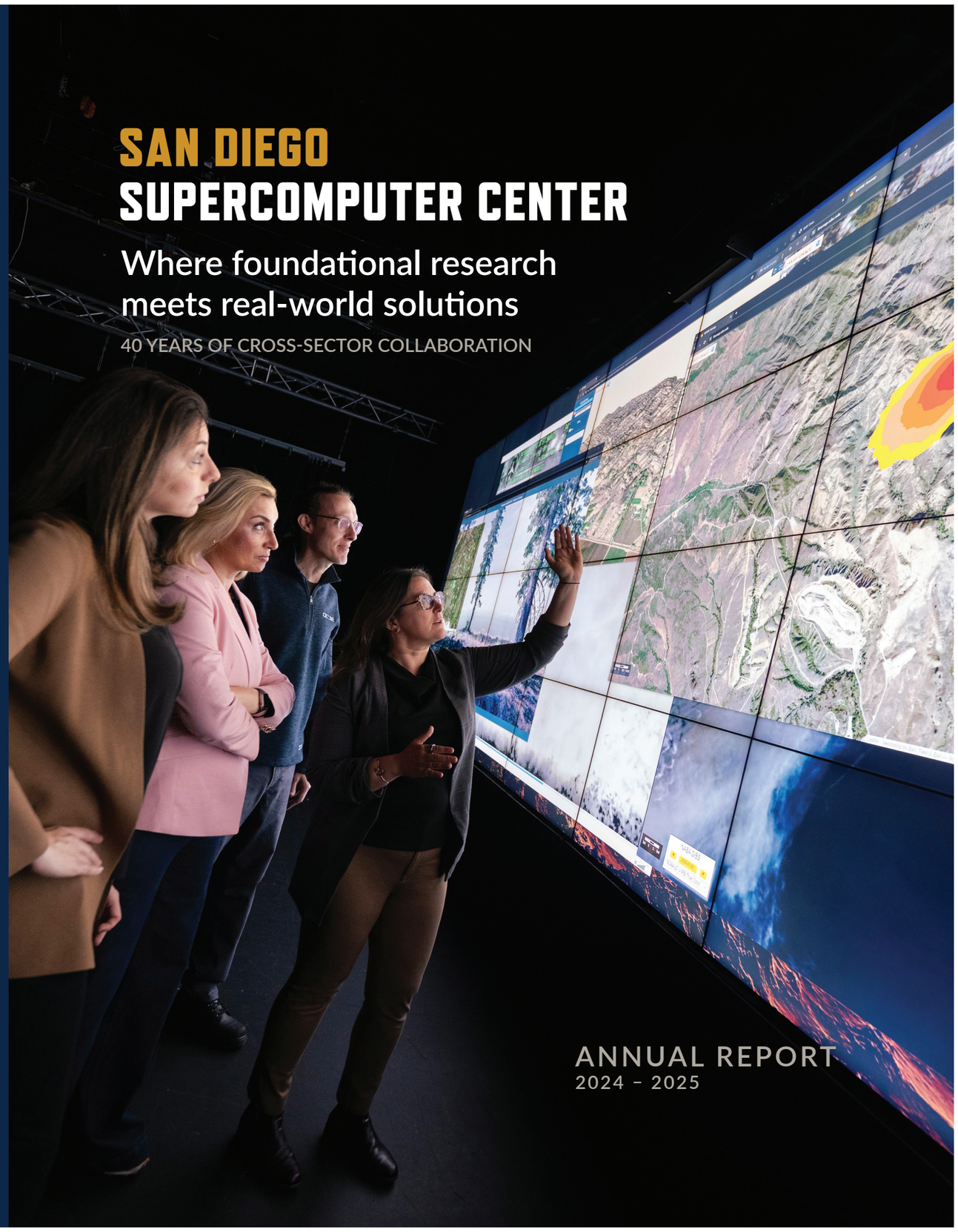
San Diego Supercomputer Center
University of California San Diego
9836 Hopkins Drive
La Jolla, CA 92093-0505

www.sdsc.edu
email/info@sdsc.edu
x.com/SDSC_UCSD
instagram.com/SDSC_UCSD
facebook.com/SanDiegoSupercomputerCenter
linkedin.com/company/san-diego-supercomputer-center
youtube.com/SanDiegoSupercomputerCenter

SAN DIEGO SUPERCOMPUTER CENTER

Where foundational research meets real-world solutions

40 YEARS OF CROSS-SECTOR COLLABORATION



ANNUAL REPORT
2024 – 2025

SAN DIEGO SUPERCOMPUTER CENTER
The San Diego Supercomputer Center (SDSC) is a leader and pioneer in high-performance and data-intensive computing, providing cyberinfrastructure resources, services and expertise to the national research community, academia and industry. Located on the UC San Diego campus, SDSC supports hundreds of multidisciplinary programs spanning a wide variety of domains, from astrophysics and earth sciences to disease research and drug discovery. Applying themes of "growing a versatile computing ecosystem" and "translating innovation into practice," SDSC continues its leadership with explorations in artificial intelligence, machine learning, cloud and edge computing, distributed high-throughput computing and more.

SDSC INFORMATION
Frank Würthwein, Director

San Diego Supercomputer Center
University of California, San Diego
9836 Hopkins Drive
La Jolla, CA 92093-0505
Phone: 858-534-5000

info@sdsc.edu
www.sdsc.edu

Cynthia Dillon
Division Director, Communications
communications@sdsc.edu

WHERE FOUNDATIONAL RESEARCH MEETS REAL-WORLD SOLUTIONS
SDSC Annual Report
2024 – 2025
Available online at www.sdsc.edu/news/annual_report.html

EDITOR:
Cynthia Dillon

CO-EDITOR:
Kimberly Mann Bruch

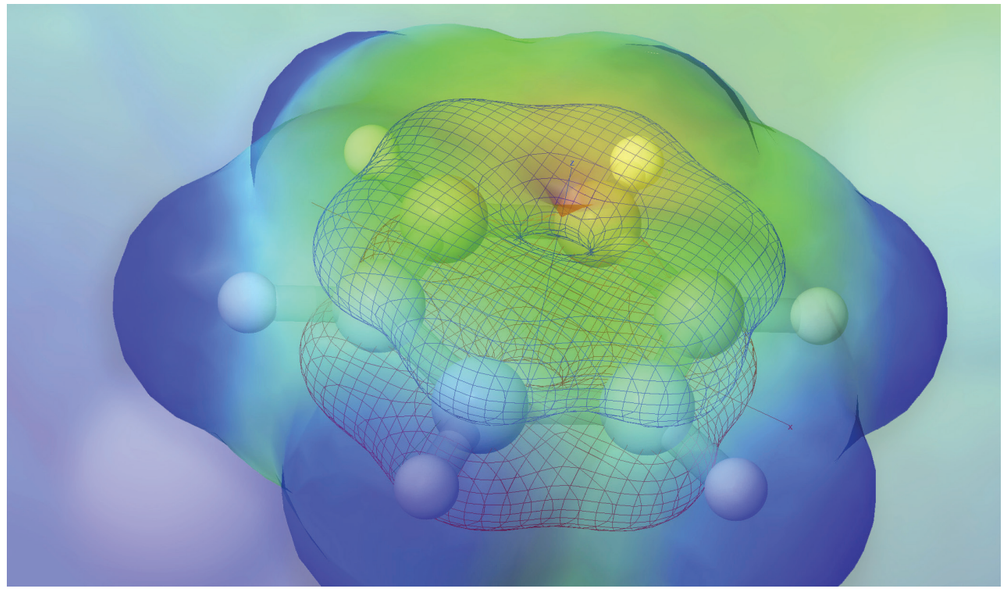
CONTRIBUTORS:
Ashley Atkins, Kim Bruch, Cynthia Dillon, Jake Drake, Fritz Leader, Amit Majumdar, Ben Tolo, Frank Würthwein

CREATIVE DIRECTOR:
Ben Tolo

PHOTOGRAPHER:
Jake Drake

All financial information is for the fiscal year that ended June 30, 2024. Any opinions, conclusions or recommendations in this publication are those of the author(s) and do not necessarily reflect the views of U.S. National Science Foundation, other funding organizations, SDSC or UC San Diego. All brand names and product names are trademarks or registered trademarks of their respective holders.

About SDSC



SDSC Helps to Expand Training for Chemistry Students and Faculty Nationwide

Sonoma State University Chemistry Professor and Chair Mark Perri received nearly $1M from the NSF to enhance ChemCompute, a free online platform he created for undergraduate students to perform advanced computational chemistry experiments. This funding also supports training faculty to integrate the platform into their courses using resources at SDSC.

"We are elated that the Expanse supercomputer is involved in the expansion of ChemCompute as this type of project aligns with our commitment to facilitate the use of data and computational resources for the development of next-gen scientific leaders," said SDSC Director Frank Würthwein. "With ChemCompute, Professor Perri has already created a way for students to easily access supercomputers and learn how to use them for computational chemistry projects – this new award allows him to take the project a step farther and expand class offerings as well as faculty trainings."

ChemCompute, which is web-based, allows students to use computer simulations for solving complex problems in topics like quantum mechanics, molecular dynamics and data science analysis. The platform, which has been in use since 2014, has seen widespread adoption at hundreds of universities, offering students the ability to perform computational chemistry calculations without the burden of licensing fees or additional costs for textbooks and software. Calculations submitted in ChemCompute are powered on Expanse as well as two additional ACCESS clusters: Bridges-2 at Pittsburgh Supercomputing Center and Jetstream2 at Indiana University.

Since its launch, ChemCompute has enabled more than 110,000 students to perform over 650,000 chemistry calculations – making a significant impact on science, technology, engineering and mathematics (STEM) education. Perri's new NSF grant will be used to further expand the platform's research capabilities, develop new curriculum, improve accessibility and foster a learning community of faculty using ChemCompute.

"I feel like I am reaching people, and the site will help students in developing new science and technology content for undergraduate research," Perri said.

Computing Systems and Infrastructure

Turning Bytes into Knowledge

Accelerating Science through Training, Application & User Support

Accelerating Science through Training, Application & User Support

Translating Research into Impacts

Virtual Environments for Science

DIRECTOR'S MESSAGE

Where foundational research meets real-world solutions

40 YEARS OF CROSS-SECTOR COLLABORATION

Dear SDSC Colleagues, Friends and Partners:

As we look back on the past year, the pace of change has been nothing short of extraordinary. Artificial intelligence is rapidly transforming research across disciplines, bringing both new possibilities and complex challenges. I am proud of the leadership role our team has played in helping the national research community navigate these shifts in ways that enable scientific advancements, translate discovery into real-world impact and make critical contributions to the nation's multi-sector science innovation ecosystem.

We strategically focus our work around five core areas of strength. This year's Annual Report is organized around these themes — Computing Systems and Infrastructure; Turning Bytes into Knowledge; Accelerating Science through Training, Application & User Support; Virtual Environments for Science; and Translating Research into Impacts. By advancing these core areas, we continue our tradition of pushing the frontiers of science, technology, education and society through innovations in data and computing that help to solve our world's most complex challenges.

As SDSC approaches its 40th anniversary in November 2025, we reflect on our legacy of leading and enabling cutting-edge scientific discoveries while setting our sights on advancing research that will shape the future. Whether through our contributions to the National AI Research Resource Pilot that empower the national research community to pursue groundbreaking AI innovations — or through translational research that saves lives and mitigates economic devastation caused by wildfires — we remain dedicated to driving transformative solutions.

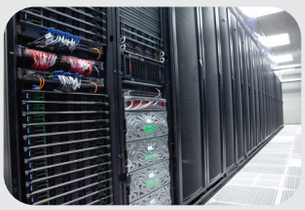
Internally, SDSC has undergone changes this year, including stepping into the role as a foundational pillar of the new School of Computing, Information and Data Sciences at UC San Diego. We also merged two Centers of Excellence into one new division called Stack Science, bringing new capabilities to the communities SDSC serves. And we launched the Societal Computing and Innovation Lab, which pioneers innovation pathways that move use-inspired problems to scalable, societally impactful solutions.

I hope you enjoy catching up with us through this year's report. We appreciate your interest. As members of our community of colleagues, funders, friends, researchers and partners, thank you for helping to make us who we are today — 40 years in and still going strong.

My best regards,
Frank Würthwein,
SDSC Director

SDSC's Five Key Themes

Operating across three scopes — national, state and local — SDSC is guided by the following five key themes:



COMPUTING SYSTEMS AND INFRASTRUCTURE

We enable discovery by architecting and operating innovative supercomputers; delivering research computing services and support; and providing infrastructure services including cloud, network, storage and data center colocation.

TURNING BYTES INTO KNOWLEDGE

We facilitate the extraction of meaningful insights from data through our expertise navigating complex, data-intensive problems. Our solutions range from addressing data management challenges to supporting AI-enabled science.

ACCELERATING SCIENCE THROUGH TRAINING, APPLICATION & USER SUPPORT

We bridge gaps by providing support to facilitate the impactful use of data and computational resources, and we remain committed to the development of next-generation scientific leaders through our mentorship and experiential learning opportunities.

VIRTUAL ENVIRONMENTS FOR SCIENCE

We make data and computationally intensive capabilities accessible to millions of researchers, educators and students across disciplines through science gateways, secure cloud enclaves, research software development and immersive visualizations.

TRANSLATING RESEARCH INTO IMPACTS

We advance scientific research in partnership with stakeholders across academia, government and industry to impact society for the better.

MISSION

Advance the frontiers of science, technology, education, and society through innovations in data and computing.

VISION

SDSC is a global leader in delivering integrated data and computing solutions that enable translational research, diverse partnerships and talent development.

Office of the Director

Frank Würthwein, Ph.D.
Director

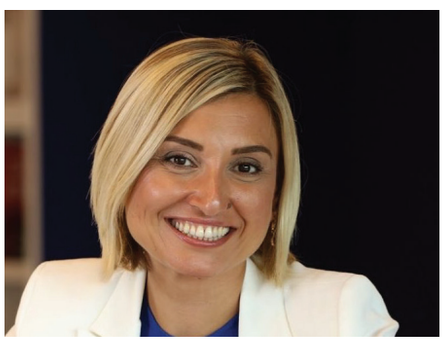
Frank Würthwein is the director of the San Diego Supercomputer Center. He holds faculty appointments at UC San Diego in the Department of Physics and the Halicioğlu Data Science Institute. After receiving his Ph.D. from Cornell in 1995, he held appointments at Caltech, MIT and Fermi National Laboratory, before joining the UC San Diego faculty in 2003. His research focuses on globally distributed compute and data systems (e.g., OSG, NRP, OSDF), experimental particle physics and distributed high-throughput computing.

Fritz Leader
Chief Administrative Officer

Fritz Leader is the chief administrative officer of SDSC and a member of the executive team. He leads SDSC's Business Office in its support of SDSC's finances, human resources and facility space. Leader is a 20-year employee of UC San Diego, and has held his role at SDSC since 2017.

Ashley Atkins, M.S.
Deputy Director

Ashley Atkins is the deputy director of SDSC. She most recently served as the executive director of the U.S. National Science Foundation-funded West Big Data Innovation Hub at UC Berkeley. Through her research, she has focused as a principal investigator (PI) and Co-PI on water resource modeling and societally beneficial applications of data and is also a Fulbright alumna.



Ilkay Altintas, Ph.D.
Chief Data Science Officer

Ilkay Altıntas is the chief data science officer of SDSC and director of the Cyberinfrastructure and Convergence Research Division. She is also a founding faculty fellow of the Halicioğlu Data Science Institute. Since joining SDSC in 2001, she has been a PI and a technical leader in a wide range of cross-disciplinary projects. With a specialty in scientific workflows, she leads collaborative teams to deliver impactful results through making computational data science work more reusable, programmable, scalable and reproducible. She holds a Ph.D. from the University of Amsterdam.

SDSC Executive Team

Frank Würthwein
Director, SDSC
Director, National Research Platform
Director, Open Science Grid
Professor, UC San Diego Department of Physics

Ilkay Altintas
Chief Data Science Officer
Division Director, Cyberinfrastructure and Convergence Research and Education

Ashley Atkins
Deputy Director, SDSC

Brian Balderston
Director of Infrastructure, Research Data Services

Sandeep Chandra
Division Director, Stack Science

Christine Kirkpatrick
Division Director, Research Data Services

Samuel 'Fritz' Leader
Chief Administrative Officer
Division Director, Business Office

Amit Majumdar
Division Director, Data-Enabled Scientific Computing

SDSC Experts

Ilkay Altintas, Ph.D.
Chief Data Science Officer
Division Director, Cyberinfrastructure Research, Education, and Development (CICORE)
Director, Societal Computing and Innovation Lab (SCIL)
Lecturer, Computer Science and Engineering, UC San Diego

Ashley Atkins, M.S.
Deputy Director
Director of Water Systems, SCIL

James Bordner, Ph.D.
Senior Computational Scientist

Stephen Burley, Ph.D.
Director, RCSB Protein Data Bank, Rutgers

Hans-Werner Braun
Research Scientist
Adjunct Professor, College of Sciences, SDSU
Director/PI, High-Performance Wireless Research and Education Network (HPWREN)
Internet Hall of Fame Inductee

Sandeep Chandra, M.S.
Executive Director, Sherlock Cloud
Director, Stack Science Division

Dong Ju Choi, Ph.D.
Senior Computational Scientist
Assistant Clinical Professor, Department of Radiation Medicine and Applied Sciences, UC San Diego

Kimberly Claffy, Ph.D.
Director/PI, Center for Applied Internet Data Analysis (CAIDA)
Research Scientist
Adjunct Professor, Computer Science and Engineering, UC San Diego
Internet Hall of Fame Inductee

Daniel Crawl, Ph.D.
Associate Director, Workflows for Data Science

Yifeng Cui, Ph.D.
Director, High-Performance GeoComputing Laboratory
Director, Intel Parallel Computing Center
PI, Southern California Earthquake Center
Adjunct Professor, SDSU

Diego Davila, M.S.
Computational Data Science and Research Specialist

Jose M. Duarte, Ph.D.
Assistant Project Scientist, RCSB Protein Data Bank

Melissa Floca, MBA
Director, Strategic Partnerships, CICORE Division
Chief Innovation Officer, SCIL

Anthony Gamst, Ph.D.
Director, Computational and Applied Statistics Laboratory

Sandra Gesing, Ph.D.
Senior Researcher
Executive Director, U.S. Research Software Engineer Association

Andreas Goetz, Ph.D.
Director, Computational Chemistry Laboratory
Co-PI, Intel Parallel Computing Center
Senior Investigator, Center for Aerosol Impacts on Chemistry of the Environment (CAICE), UC San Diego

Madhusudan Gujral, Ph.D.
Bioinformatics Programmer Analyst

Amarnath Gupta, Ph.D.
Director, Advanced Query Processing Lab of SDSC
Co-PI, Neuroscience Information Framework (NIF) project, Calit2

Bradley Huffaker, M.S.
Senior Research Programmer, CAIDA
Specialist, Computer Networks

Thomas Hutton
Chief Network Architect

Martin Kandes, Ph.D.
Research Specialist, Computational and Data Science

Christine Kirkpatrick, M.A.S.
Division Director, Research Data Services
Head, GO FAIR US
Secretary General, CODATA
PI, West Big Data Innovation Hub
Ex Officio Member, U.S. National Committee for CODATA for the National Academics of Sciences, Engineering, and Medicine
Co-Chair, FAIR Digital Object Forum

Valentina Kouznetsov, Ph.D.
Associate Project Scientist
Research Professor

Rodman Linn, Ph.D.
Associate Director, Fire Science of the WIFIRE Lab

Timothy Mackey, Ph.D.
SDSC Affiliate
Professor, Global Health Program, UC San Diego
Director, Global Health Policy and Data Institute

Amit Majumdar, Ph.D.
Division Director, Data Enabled Scientific Computing
PI, Voyager
Associate Professor, Department of Radiation Medicine and Applied Sciences, UC San Diego

Dmitry Mishin, Ph.D.
Applications Developer

Ka Pui Mok, Ph.D.
Research Scientist, CAIDA

Viswanath Nandigam, M.S.
Director, Advanced Cyberinfrastructure Development Lab
PI, OpenTopography
Co-I OpenAltimetry

Mai H. Nguyen, Ph.D.
Lead. Data Analytics

Michael Norman, Ph.D.
Distinguished Professor, Physics, UC San Diego
Director, Laboratory for Computational Astrophysics, UC San Diego

Wayne Pfeiffer, Ph.D.
Distinguished Scientist

John Porten, Ph.D.
Senior Research Manager

Zaira Razu, M.A.
Director, Convergence Research (CORE) Institute

Paul Rodriguez, Ph.D.
Research Analyst

Peter Rose, Ph.D.
Director, Structural Bioinformatics Laboratory
Lead, Bioinformatics and Biomedical Applications, Data Science Hub

Joan Segura, Ph.D.
Scientific Software Developer, RCSB Protein Data Bank

Igor Sfiligoi, M.S.
Senior Research Scientist, Distributed High-Throughput Computing
Lead Scientific Software Developer and Researcher

James Short, Ph.D.
Lead Scientist
Director, Spark AI
Director, BlockLAB

Robert Sinkovits, Ph.D.
Director, Scientific Computing Applications
Director, Education and Training

Subhashini Sivagnanam, M.S.
Lead, CyberInfrastructure Solutions and Services
Lead, Triton Shared Computing Cluster
PI, Open Science Chain
Co-PI, Neuroscience Gateway

Shava Smallen, M.S.
Manager, Cloud and Cluster Development
Lead Software Architect and Co-PI, CloudBank
Steering Committee Co-Chair, Pacific Rim Application and Grid Middleware Assembly (PRAGMA)

Claire Stirm, M.S.
Project Manager, Science Gateways Center of Excellence (SGX3)

Mahidhar Tatineni, Ph.D.
Lead, User Support
Research Programmer Analyst

Mary Thomas, Ph.D.
Computational Data Scientist
Lead, HPC Training
Co-PI, CC Compute: Triton Stratus*

Igor Tsigelny, Ph.D.
Research Scientist
Research Scientist, Department of Neurosciences, UC San Diego

David Valentine, Ph.D.
Research Programmer, Spatial Information Systems Laboratory

Rick Wagner, Ph.D.
Chief Technology Officer
Director, Industry Partnerships

Tanya Wolfson, M.A.
Senior Staff Member, Computational and Applied Statistics Laboratory

Nicole Wolter
Computational and Data Science Research Specialist

Frank Würthwein, Ph.D.
Director
Lead, Distributed High-Throughput Computing
Professor, UC San Diego Department of Physics
Professor, Halıcıoğlu Data Science Institute

Kenneth Yoshimoto, Ph.D.
Researcher, Computational and Data Science

Choonhan Youn, Ph.D.
Scientific Researcher

Ilya Zaslavsky, Ph.D.
Director, Spatial Information Systems Laboratory

Andrea Zonca, Ph.D.
Specialist, HPC Applications

Amarnath Gupta Named SDSC's Pi Person of the Year

Each year, SDSC recognizes an individual whose research contributions over the years span both science and cyberinfrastructure (CI) technology. The Pi Person of the Year Award, represented by the Pi symbol (∏), was first bestowed to an SDSC researcher in 2013. This year's recipient is Amarnath Gupta, a senior research scientist within the CICORE Division.

Gupta has made significant contributions to various domains, from neuroscience and oceanography, to public health and food security. His work has led to the creation of a 2008 large neuroscience ontology, which is still in use today; the 2012 Neuroscience Information Framework platform, recognized by the White House as a major "big data" CI and the Polystore-based AWESOME system, 2016-2018, used in multiple fields such as quantum materials science, political science and healthcare projects like Tempredict (Wearable Device for Physiological Data Monitoring) and, most recently, NOURISH. These solutions have been practically used by domain scientists and have earned Gupta recognition, including the ACM Distinguished Scientist Award.

A computer scientist by training, Gupta specializes in information systems, including heterogeneous information integration, query processing techniques, knowledge engineering, graph-based access control models and, most recently, conversational engines and novel use of large language models (LLMs) for different information systems tasks. He has more than 100 publications and four issued patents over his 27 years at SDSC.

CONTRIBUTIONS TO RESEARCH

Gupta has a long history of research in computer science, particularly in addressing the "Variety" problem of big data, which involves integrating heterogeneous information sources. Over the past 10 years, he has developed a novel framework that enables users to query information across multiple data sources and create virtual knowledge graphs. This work has resulted in several publications, research funding and two new U.S. patents related to optimized data ingestion and joint query processing.

One notable application of this architecture was a project funded by the U.S. Navy, which used knowledge graph analytics to identify knowledge gaps between the U.S. and other countries in areas of interest to the Navy. The project analyzed data from U.S. Patents, Publications, and News.

Gupta's current research focuses on using large language models (LLMs) to automate the construction and interrogation of knowledge graphs. He has published work on constructing domain-specific knowledge graphs from text using LLMs and deep learning methods like graph neural networks. His research also explores using AI techniques to plan and validate cross-model queries, perform semantic schema mapping, and generate polystore query plans from natural language questions.

USE OF CI IN RESEARCH

According to Ilkay Altintas, SDSC's chief data science officer, all of Gupta's work relies on scalable data systems. Many of the projects mentioned above offer Representational State Transfer (REST) services that internally use the National Research Platform (NRP), a notable utilization of SDSC's cyberinfrastructure.

An example is the Tempredict project that ingested physiological data daily from rings worn by 60,000 subjects in a UC San Francisco (UCSF)-led study. The data was analyzed to determine subjects that should take a COVID-19 test the following day because the wearable ring data suggested likely infection. Since the data had personal information, it was anonymized using the secure facilities of Sherlock, a service provided within the Stack Science Division at SDSC, then tunneled into the data services of AWESOME and sent to the National Research Platform (NRP) for statistical analysis. Results of the analysis were routed back to UCSF via the Sherlock secure system for de-anonymization, and ultimately shared with the at-risk subjects to inform them of their likely COVID infection with a suggestion to take a COVID test.

Further, Gupta's lab offers a set of CI services including the semantic search service used in NDP and several ontology and knowledge graph services offered for the NOURISH project. The AWESOME platform services are also offered to students as part of different classroom and mentorship activities.

"Amarnath's talent for translating research across domains into problem-solving systems typifies our mission at SDSC to advance the frontiers of science, technology, education and society through innovations in data and computing," SDSC Director Frank Würthwein said. "His many years of dedicated service to SDSC and the research community is noted and very much appreciated."

Divisions of SDSC

SDSC is organized into six unique divisions. Each operates under the guidance of a director, and each is a vital component of SDSC's collective whole. The divisions represent the five key themes of SDSC through their activities and projects, and they work collaboratively to meet the center's mission and vision.

BUSINESS OFFICE

The Business Office delivers experienced administrative support to all other divisions of SDSC and partial administrative support to the Halicioglu Data Science Institute. The Business Office provides effective financial management, human resources, business applications, and space management to enable high productivity for our customers. The team partners with UC San Diego's central administrative units to execute contracts and awards, provide comprehensive HR support at all stages of employment, oversee accounts payable and receivable, meet the space needs of employees, and more.

CYBERINFRASTRUCTURE AND CONVERGENCE RESEARCH AND EDUCATION

The Cyberinfrastructure and Convergence Research (CICORE) Division provides expertise in data science, AI, and advanced computing to partnerships and projects focused on achieving scientific and societal impact. The CICORE team has a passion for solving complex, data intensive problems and a commitment to developing the next generation of technological leaders. The CICORE Division is known for its broad expertise in cyberinfrastructure; domain-specific expertise in AI-enabled science; successful partnerships with policymakers, practitioners and industry; and experiential education initiatives.

COMMUNICATIONS

The Communications Division serves as a comprehensive public relations unit for SDSC. The team of communications experts works across the center with leaders, researchers, staff and students to stay informed about news and information to share with both internal and external audiences. The division helps to establish, maintain and elevate the SDSC brand, while promoting the center's systems and services through science and technology-focused storytelling via news, media relations, multimedia and marketing activities and resources. The division plays a unique role in SDSC's strategic success as part of the school of Computing, Information and Data Sciences at UC San Diego.

DATA-ENABLED SCIENTIFIC COMPUTING

The Data-Enabled Scientific Computing (DESC) Division is organized into multiple groups that have specific expertise to lead HPC and computational sciences innovation. The division serves and collaborates with thousands of SDSC's national, UC-wide and UC San Diego researchers, as well as with industry partners to provide full support and training to SDSC's user community.

RESEARCH DATA SERVICES

The Research Data Services (RDS) division provides research computing services and support for data-driven and computationally based research. Highlighted services are custom GPU compute solutions, large-scale high-performance storage, data management tool development, and piloting CI services to meet emerging research computing needs. Infrastructure services include cloud, networking, storage, and a data center colocation facility. RDS research and grants are supported by a data initiatives group that includes world-class leaders in research data management, storage at scale, cloud computing and science coordination.

STACK SCIENCE

The Stack Science Division is a leading academic organization specializing in cyberinfrastructure development, managed cloud services, and community building. The team collaborates with researchers and educators to facilitate breakthrough discoveries by building and operating virtual environments for science and education. Stack Science Division services drive research and scientific outcomes through a comprehensive suite of offerings, including science gateways as a service, research software development, secure cloud enclaves, and regulatory data management. Adopting a customer-centric and solution-focused approach, the Stack Science team is dedicated to advancing scientific community building, enhancing cyberinfrastructure, accelerating research goals and achieving societal impact through collaboration and excellence.

SDSC BY THE NUMBERS

FISCAL YEAR 2023/24

ORGANIZATION

$66M ANNUAL REVENUE

$34M GRANT FUNDING

$8M INDUSTRY REVENUE

319 EMPLOYEES & VOLUNTEERS

83 SPONSORED RESEARCH AWARDS

TRAINING PROGRAMS

32,663 STUDENTS ENROLLED WORLDWIDE IN SDSC-LED ONLINE COURSES

3,273 TRAINING & EVENT PARTICIPANTS

427 STUDENTS ENGAGED IN SDSC PROGRAMS

197 HIGH SCHOOL STUDENTS MENTORED

67 SDSC HOSTED PROGRAMS

SUPPORT FOR UC

4,015 UC USERS: EXPANSE, VOYAGER & NRP

1,150 UC USERS: TRITON SHARED COMPUTING CLUSTER

970 UC ACTIVE ALLOCATIONS: EXPANSE, VOYAGER & NRP

84+ PETABYTES OF STORED DATA

HPC SYSTEMS

222K x86 CORES ON SDSC HPC SYSTEMS

3,817 ACCELERATORS

332 PUBLICATIONS CITING SDSC/ACCESS RESOURCES

7 COMPUTE CLUSTERS

History, Services, Support

The San Diego Supercomputer Center has been a leader in high-performance and data-intensive computing and cyberinfrastructure for nearly 40 years. As a unique supercomputer center in the University of California system, SDSC provides systems, services and research expertise to local, regional and national partners in academia and industry.

Since July 2024, SDSC has been a part of UC San Diego's new School of Computing, Information and Data Sciences, which aims to transform data into actionable knowledge. SDSC continues to explore innovative areas of expertise such as artificial intelligence, machine learning and cloud computing, with a focus on growing a versatile computing ecosystem to translate research into applications that help address societal challenges.

SDSC offers a range of services and expertise to its user community — from high-performance computing and data science solutions to cyberinfrastructure and business services.

SDSC also provides in-depth technical support to its service users, including assistance with developing efficient HPC applications, accounts and allocations to the UC and UC San Diego academic research community, national HPC users and industry partners; user guides and documentation for its compute and data systems, and SDSC's experienced consultants are available to assist users as well.

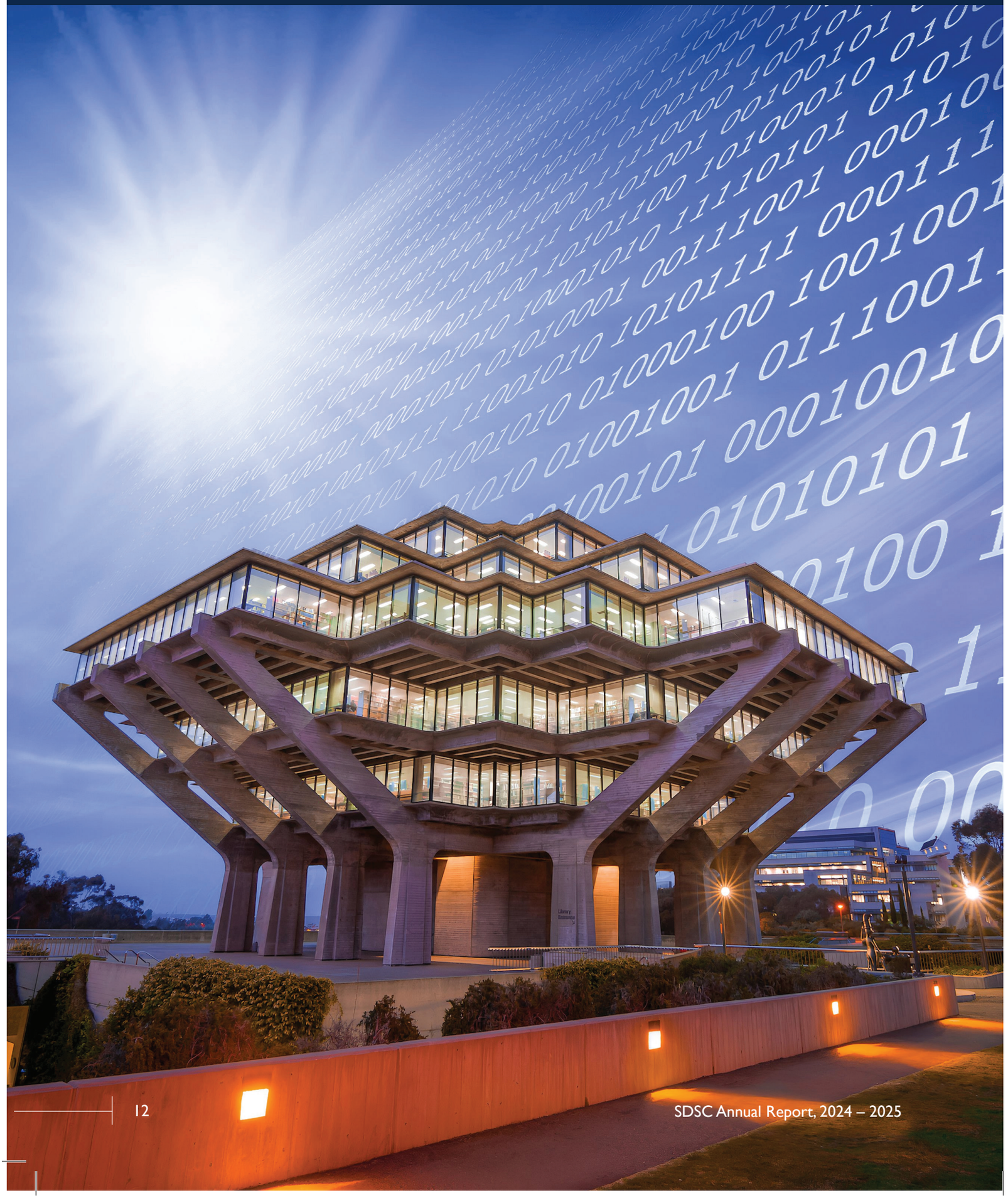
Additionally, SDSC offers education and training programs, including conferences, workshops and summer institutes, to support its advanced computing system users.

Please visit www.sdsc.edu/support/ for more information.

Celebrating 40 Years

Four decades ago this year, the San Diego Supercomputer Center was established as one of the nation's first supercomputer centers to provide computational resources to the national science community under a cooperative agreement by the U.S. National Science Foundation (NSF) in collaboration with UC San Diego and General Atomics (GA) Technologies. SDSC first opened its doors on Nov. 14, 1985, and today its relationships with the NSF, GA and the national research community remain strong. During 2025, the NSF, SDSC and its affiliate High Performance Wireless Research and Education Network (HPWREN) each will celebrate a key anniversary: 75th for NSF, 40th for SDSC and 25th for HPWREN. Congratulations on these milestones!

SCHOOL OF COMPUTING, INFORMATION AND DATA SCIENCES



SDSC Established as Pillar of New School of Computing, Information and Data Sciences

In July 2024, the University of California Board of Regents approved the creation of UC San Diego's new School of Computing, Information and Data Sciences (SCIDS), which aims to advance innovation and education in artificial intelligence (AI), computing and data science.

SCIDS is the fourth school to be added to UC San Diego in the 21st century and brings together faculty from various disciplines to improve the human condition by understanding how data shapes society, and by preparing the next generation of workers in AI. The school aims to make a positive impact on society by leveraging data science and AI to address real-world problems.

"The new school at UC San Diego will grow our impact on society via translational computing, information and data sciences, and bring AI education to community and teaching colleges across California and the nation via our AI infrastructure," said SDSC Director Frank Würthwein, a founding faculty member of the Halicioglu Data Science Institute (HDSI). "Combining our strengths with those of HDSI optimizes our leadership in innovation for science, technology, education and society."

Educating the next generation of machine learning engineers and data analysts, HDSI brings together an interdisciplinary team of faculty and researchers from areas ranging from computer science to communications to medicine. Working together, these researchers explore new computational methods, new mathematical models and guide societal and ethical impacts of data science.

"HDSI and SDSC share the unique challenge of building transdisciplinary academics and research. Their coming together under SCIDS will involve new synergies and realize tremendous new possibilities in creating talent in emerging areas, including artificial intelligence," said HDSI Founding Director and SCIDS Interim Dean Rajesh Gupta.

By translating data science from the classroom to research and the broader workplace, the school prepares students for their careers by providing opportunities for them to engage directly with industry and government partners, including emergency responders; municipal, state and national resource management organizations; and nonprofits. Students learn first-hand how data science can allow organizations to better address societal problems ranging from climate change mitigation and social justice issues, to technical challenges and healthcare.

Computing Systems & Infrastructure

SDSC enables discovery by architecting and operating innovative supercomputers, delivering research computing services and support; and providing infrastructure services including cloud, network, storage and data center colocation facilities. Following are examples of our impacts under this theme.

Expanse: A Computing Resource for ACCESS and AI

Last year, the U.S. National Science Foundation (NSF) awarded a $5.3 million supplement to SDSC to expand the artificial intelligence (AI) capabilities of Expanse, SDSC's flagship supercomputer. The award was made as part of the National AI Research Resource (NAIRR) Pilot program, which aims to connect U.S. researchers and educators to computational, data and training resources needed to advance AI research and research that employs AI.

Expanse is a computing resource for the NSF's Advanced Cyberinfrastructure of Coordination Ecosystem: Services & Support (ACCESS) core software stack, which includes remote login, remote computation, data movement, science workflow support and science gateway support toolkits. Expanse is accessible to ACCESS users, who are given time on the system by submitting a proposal through the ACCESS Allocation Request System. Interested parties may contact the ACCESS Help Desk for assistance with an Expanse proposal. Full instructions can be found by visiting the Expanse User Guide page on the SDSC website.

Technically, the Expanse cluster was designed by Dell and SDSC, delivering 5.16 peak petaflops and offering Composable Systems and Cloud Bursting. It has standard compute nodes that are each powered by two 64-core AMD EPYC 7742 processors and contain 256 GB of DDR4 memory, while each GPU node contains four NVIDIA V100s (32 GB SMX2) connected via NVLINK and dual 20-core Intel Xeon 6248 CPUs. Expanse also has four 2 TB large memory nodes.

The original resource is organized into 13 SDSC Scalable Compute Units (SSCUs), comprising 728 standard nodes, 54 GPU nodes and four large-memory nodes. Every Expanse node has access to a 12 PB Lustre parallel file system (provided by Aeon Computing) and a 3.5 PB Ceph Object Store system. Since its deployment in 2020, six SSCUs have been added to support the Partnership to Advance Throughput Computing (PATh), SDSC's industry program, the Center for Western Weather and Water Extremes at Scripps Institution of Oceanography, and the NAIRR Pilot program. Expanse uses the Bright Computing HPC Cluster management system and the SLURM workload manager for job scheduling.

The recent NAIRR expansion was funded by a $5.3M NSF supplement to the original award and includes two SSCUs, each consisting of 17 GPU nodes, one NDR NVIDIA InfiniBand switch and one 25 Gbps Ethernet switch, plus a dedicated 3 PB Ceph filesystem. Each GPU node contains two 36-core Intel Sapphire Rapids CPUs, four NVIDIA H100 GPUs and 6.4 TB of local NVMe storage. This resource is aimed exclusively for AI workloads of researchers from all domains.

Enabling AI Research on NVIDIA DGX Cloud Resources

SDSC received an Early-concept Grants for Exploratory Research (EAGER) award from the U.S. National Science Foundation (NSF) to support research groups using the NVIDIA DGX Cloud platform. The focus of this award is to provide IT support for optimization, performance monitoring and determining the best ways to run National Artificial Intelligence Research Resource (NAIRR) Pilot projects on the resources.

According to Principal Investigator Mahidhar Tatineni, the goal is to support dedicated systems for each research group, allowing uninterrupted access to perform complex modeling that can take weeks or months to complete. This is particularly important for tools that require special configurations not possible in shared system environments, and NVIDIA DGX Cloud enables the creation of custom environments for researchers. By using both cloud servers and physical servers at supercomputing centers, researchers can increase their potential for new discoveries.

The EAGER project aims to provide guidelines, training and tools to help NAIRR Pilot researchers use cloud platforms like NVIDIA DGX Cloud — a high-performance, fully managed AI platform ideal for NAIRR Pilot projects that require large-scale AI resources for focused research campaigns — more effectively, optimizing their work and speeding up research processes. The platform offers dedicated access to capacity that supports the entire AI lifecycle, from building and customizing foundation models to serverless inference, making it an essential tool for researchers working on AI projects, enabling them to work more efficiently and effectively.

Voyager Enters the Allocations Phase

Voyager is an Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) innovative AI system designed specifically for science and engineering research at scale. Voyager is focused on supporting research in science and engineering that is increasingly dependent upon artificial intelligence and deep learning as a critical element in experimental and/or computational work. Featuring the Habana Gaudi1 and Gaudi2 training and first-generation Habana inference processors, Voyager encompasses a high-performance, low latency 400 gigabit-per-second interconnect from Arista.

Voyager provides researchers with the ability to work on extremely large data sets using standard AI tools, like PyTorch, or develop their own deep learning models using developer tools and libraries from Intel Habana Labs.

Voyager is an NSF-funded system, developed in collaboration with Supermicro, Arista and Intel's Habana Lab. After completing a three-year testbed phase, Voyager is entering a two-year Allocations Phase to the broader NSF community starting on June 1, 2025. It is currently also a NAIRR Pilot resource. High energy physics, earth observation data models, biomedical text analytics, cancer therapy, EEG applications, image classification, deep learning models for biomolecular simulations, and LLM for wildfire smoke detection are some of the applications being used on Voyager. To request access to Voyager, please follow the ACCESS allocation process.

Innovative Cosmos Supercomputer to Democratize the Accelerator Ecosystem

SDSC received a U.S. National Science Foundation (NSF) award to develop the Cosmos supercomputer. The Cosmos system is built by Hewlett Packard Enterprise (HPE) and features innovative AMD Instinct MI300A accelerated processing units (APUs). These APUs provide both central processing unit (CPU) and graphics processing unit (GPU) capabilities with unified memory, making them highly performant and energy efficient.

The Cosmos system is designed to democratize access to accelerated computing, allowing researchers to exploit the power of accelerators via an incremental programming approach. This approach enables many communities to adopt GPUs and ease the process of porting and optimizing a range of applications.

According to Mahidhar Tatineni, principal investigator on the project, the benefits of accelerating applications will aid discoveries in various domains, including AI, astrophysics, genomics and materials science.

The AMD Instinct MI300A APU features an in-chip memory layout, which is integrated and shared between CPU and GPU resources. This type of memory architecture facilitates an incremental programming approach, making it easier for researchers to use accelerators. The directive-based programming approach has been demonstrated in a case study on the widely used OpenFOAM computational fluid dynamics code.

The Cosmos system is expected to significantly increase the range of applications that can effectively use accelerators, making it a powerful tool for researchers.

As Brent Gorda, senior director at AMD, noted, the APU advantage provides AMD Instinct MI300A APUs with unified memory and cache resources to deliver an easily programmable platform.

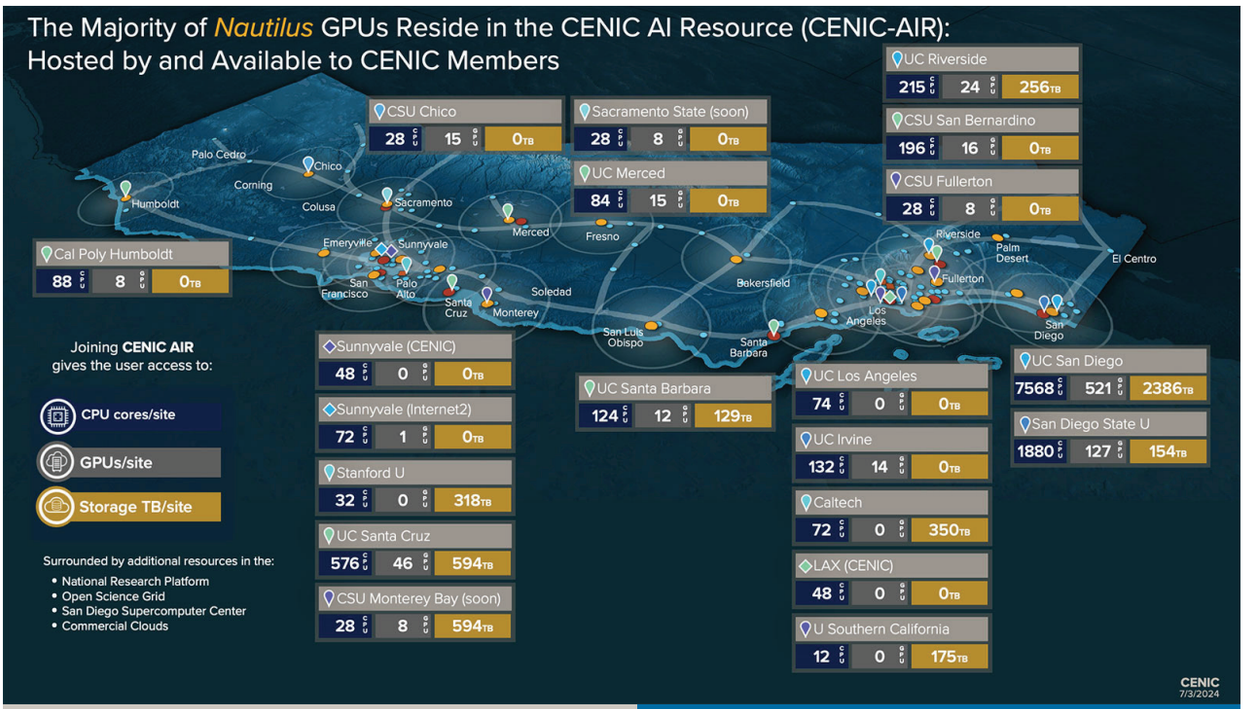
SDSC and CENIC Work to Expand AI Education Infrastructure in California

SDSC and the Corporation for Education Network Initiatives in California (CENIC) are working together to expand AI education infrastructure in California and beyond. Together, they have developed an AI education infrastructure called CENIC AI Resource (CENIC AIR) for K-12 education, public libraries, community colleges, the California State University (CSU) and the University of California (UC) campus systems.

According to Frank Würthwein, SDSC director, CENIC AIR was created to support the next generation of academics and innovators who will be working with artificial intelligence and machine learning.

Würthwein, along with UC San Diego's retired Distinguished Professor and CENIC Board Member Larry Smarr, and representatives from San Diego State University (SDSU) and the San Diego Community College District (SDCCD), presented details about the CENIC AIR program to CENIC's Board of Directors. The goal of this presentation was to gain wider support for the program within the state and beyond. Currently, there are nine CSU campuses, plus Miramar College, that are using CENIC AIR.

Tom DeFanti, a retired research scientist, noted that UC San Diego is the primary institution behind the National Research Platform (NRP). He stated that UC San Diego grants have purchased most of the computers and storage that make up CENIC AIR and NRP. DeFanti also mentioned that other campuses are adding compute nodes and storage. NRP provides operations support for users.



Key Attributes of UC San Diego's and CENIC's CENIC AIR Collaboration

- CENIC provides the Science DMZ networks that NRP depends on to integrate campus-owned hardware across colleges in California.
- NRP allows colleges to buy their own AI hardware without the need to have system administrators. The NRP team manages all hardware remotely for free, thanks to funding by the U.S. National Science Foundation (NSF).
- Educators from Community Colleges, CSUs and UCs use NRP to bring compute, data and AI resources into their classrooms. Students use LLMs to develop software to visualize data using Jupyter notebooks while working on in-class assignments.
- Curated content is made available according to FAIR (Findable, Accessible, Interoperable, and Reusable) principles via the National Data Platform.
- CENIC AIR bridges and smooths the transition from AI education to the application of AI research by providing the capacity to bring research-scale activities to the classroom.

National Research Platform

NRP is an open, nationally-distributed data and compute platform. Any college in the USA may join to have their hardware operated by the NSF-funded NRP operations team. As of July 2025, NRP manages hardware at more than 80 locations, including more than 50 colleges and more than 20 internet points of presence. Multiple DOE- and NSF-funded projects build on NRP. It is led by SDSC Director Frank Würthwein, in collaboration with experts at the University of Nebraska-Lincoln and Massachusetts Green High Performance Computing Center. NRP offers an array of resources, including NAIRR classroom resources. NRP is also part of SDSC's contributions to the National Artificial Intelligence Research Resource (NAIRR) Pilot program.

SDSC Leads $20M Grant to Provide Commercial Cloud Resources to National Research Community

SDSC will lead a $20 million grant from the U.S. National Science Foundation (NSF) for the acquisition and deployment of commercial cloud resources to support approximately 500 research projects annually across the nation over the next five years.

CloudBank 2 is a joint project between SDSC, the University of Washington (UW) eScience Institute and the UC Berkeley College of Computing, Data Science, and Society. It aims to provide commercial cloud resources to support various science and engineering research use cases, including access to Jupyter Notebooks, secure data, GPUs and quantum computers. This will enable researchers to work on projects that require resources beyond what is available in a single physical location.

“We will build upon our NSF-funded CloudBank pilot project, a flexible multi-cloud infrastructure now in its sixth year of operation, and scale to the broader science and engineering community,” explained CloudBank 2 Principal Investigator (PI) Shava Smallen, also the lead developer of the CloudBank Portal. Joining Smallen on the CloudBank 2 project as co-PIs are: UC San Diego’s Vince Kellen and Michael Norman, who serves as PI for the CloudBank pilot; UW’s Sarah Stone and UC Berkeley’s Katherine Yelick.

For national academic researchers, CloudBank 2 will integrate multi-cloud capabilities into the ACCESS ecosystem — a program established and funded by the NSF to help researchers and educators with or without supporting grants to use the nation’s advanced computing systems and services. Researchers will access commercial cloud resources free of charge via the ACCESS allocation process beginning in December (2025), just as they do any other NSF-funded computational resource,” Smallen said.

The CloudBank 2 project will utilize several commercial clouds, including Amazon Web Services, Google Compute Platform, Microsoft Azure and IBM Cloud, which are already available through CloudBank. These clouds provide a wide range of resources, platforms and services that support various research areas, including computation, data analysis and AI research.

Additionally, the commercial clouds offer diverse Infrastructure as a Service (IAAS) options for compute, storage and networking, as well as elastic and scalable platforms for big data management, machine learning and quantum computing.

The advantages of using commercial clouds include their widespread availability, zero upfront capital costs and ability to quickly scale up or down as needed. Additionally, commercial clouds provide continuous hardware and software updates, are highly reliable and offer a growing range of marketplace services, such as NVIDIA DGX Cloud.

The federal government has recognized the importance of commercial clouds in creating a federated advanced computing ecosystem, and CloudBank 2 will complement the NSF portfolio by expanding the resources available to researchers.

Liquid Cooling Upgrade Keeps Supercomputers Chill at UC San Diego

SDSC completed a major upgrade to its liquid cooling system, which is essential for keeping today's powerful supercomputers cool, efficient and online. The upgrade has quadrupled the liquid cooling capacity at SDSC's West Data Center from 600 kilowatts (kW) to 2.4 megawatts (MW). For context, 1kW (or 1,000 watts) is equivalent to running a refrigerator for one day; 1MW (or 1 million watts) can power roughly 750 to 1,000 households for an hour. The upgrade also laid the groundwork for future capacity improvement.

The need for the upgrade was clear, as supercomputers like the Expanse system – and systems designed for AI – generate a lot of heat while running complex simulations for research. If not cooled properly, these machines can overheat, damaging components and reducing their lifespan. Liquid cooling efficiently removes heat, allowing supercomputers to operate at optimal temperatures and helping to reduce energy consumption.

According to Christine Kirkpatrick, director of SDSC's Research Data Services Division, which completed the upgrade, SDSC's first supercomputers nearly 40 years ago were liquid cooled using now defunct methods. For two decades, supercomputers were air cooled. Now liquid cooling is the only option for many of the systems designed for AI and other modern supercomputers.

The recent upgrade at SDSC expanded the facility's liquid distribution system, making liquid cooling accessible anywhere in the data center, significantly increasing capacity. This infrastructure means SDSC can now fully utilize the West Data Center for upcoming computing challenges, particularly as high-powered systems demand precision cooling solutions.

Several advanced systems are already taking advantage of the expanded cooling network, including Aware, Cosmos and RDS' AI-ready racks equipped with Rear-Door Heat Exchangers. The next phase of improvements is already being planned, focusing on the basement level, which will add pumps and heat exchangers to the facility's process loop.

"Liquid distribution was our limiting factor — now, it's our strength," said RDS Director of Infrastructure Brian Balderston, who led the upgrade project with SDSC Data Center Manager Thomas Tate. "Our biggest constraint for deploying new, large systems has been solved with this upgrade, clearing the way for a new generation of high-performance infrastructure."

As SDSC Director Frank Würthwein noted, "As supercomputers grow more powerful — and hotter — liquid cooling has become the standard cooling method. With this upgrade led by Brian and his team, SDSC is now better positioned than ever to support the next generation of computational research — from AI and astrophysics modeling to genomics and beyond."

Turning Bytes into Knowledge

SDSC facilitates the extraction of meaningful insights from data through its expertise navigating complex, data-intensive problems. The solutions it offers range from addressing data management challenges to supporting AI-enabled science.

Using AI and Data Science to NOURISH Food Deserts

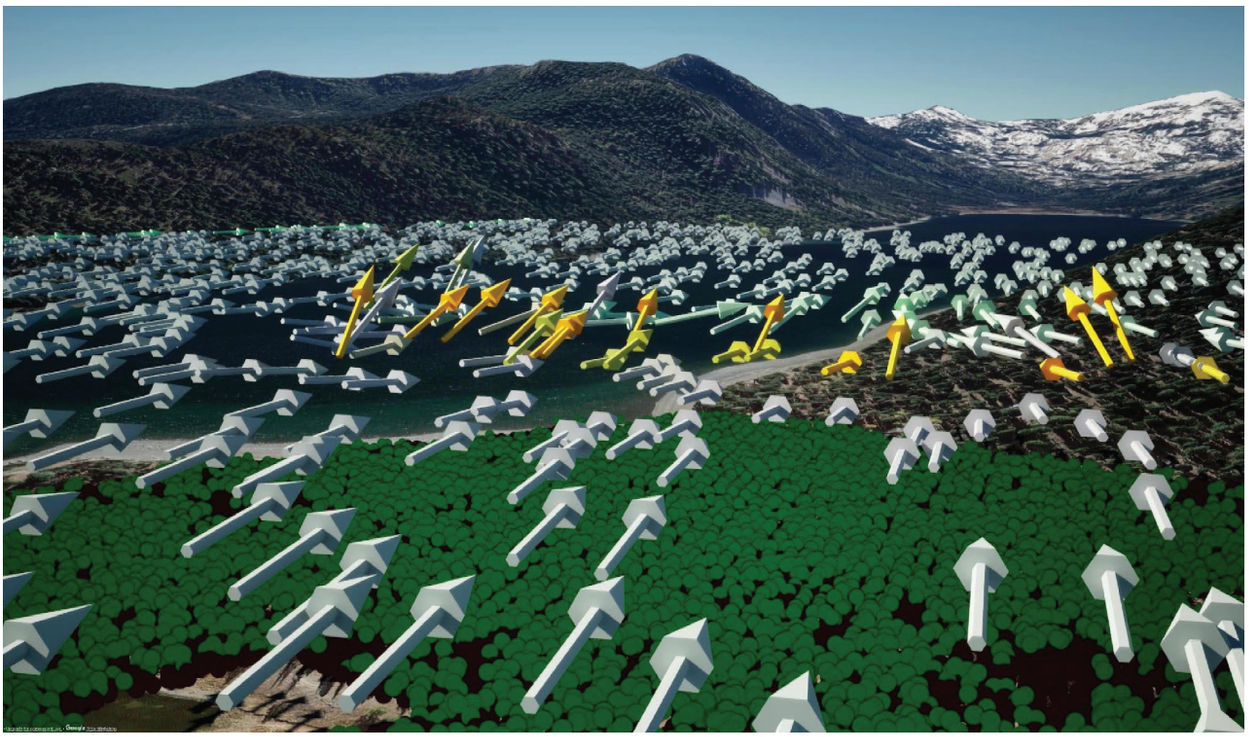
Twenty-four million Americans live in areas with limited access to fresh food, known as food deserts, where unhealthy, ultra-processed foods are abundant. To address this issue, researchers from UC San Francisco and UC San Diego created the NOURISH platform, which aims to provide small business owners in these communities with resources to sell fresh food. The platform offers access to loans and grants, online maps to optimize fresh food outlet placement, help with business permits and AI-guided sourcing of fresh ingredients.

According to Laura Schmidt, the project's principal investigator, NOURISH complements government efforts to increase fresh food availability in food deserts by leveraging the entrepreneurial talent and unique food heritages of these communities.

A team of computer scientists, software developers and students led by Amarnath Gupta in the SDSC's CICORE Division is developing the platform. It will include interactive maps of local food systems and patented recommendation algorithms to customize business plans based on local consumer preferences.

Ilkay Altintas, SDSC's chief data science officer and director of the Societal Computing and Innovation Lab, stated that NOURISH embodies the center's vision of using data science and knowledge management to create impactful solutions to societal challenges.

The platform will be accessible from a mobile phone in multiple languages, aiming to democratize data access and create a more level playing field between large food companies and small businesses.



SDSC Launches the Wildfire Science & Technology Commons

Researchers from SDSC launched a new initiative to advance science and technology that addresses wildland fire challenges in an era of more frequent and devastating megafires — the Wildfire Science & Technology Commons.

A central hub for data, models, computing resources and expertise to enable wildland fire researchers to collaborate with each other and practitioners to move quickly from theoretical ideas and experimental workflows to impactful, scalable real-world solutions, the Wildfire Commons will advance fire tech by quickly moving research workflows into production through a community of practice, data standards and innovation pathways centered around open data, cutting-edge science and AI. The Wildfire Commons was initiated by the Proactive Wildfire & Environmental Sustainability Solutions (ProWESS) Center — a collaboration between UC San Diego and Los Alamos National Laboratory.

"Our goal is to accelerate cross-domain and cross-sector collaborations that advance solutions for a wildfire-resilient future. We are creating a community platform and marketplace for data and knowledge sharing," said Chief Data Science Officer at SDSC and Co-Director of ProWESS Ilkay Altintas.

The Wildfire Commons enables researchers, innovators and practitioners to collaborate on next-generation solutions by providing easy access to data, models, computing resources and expertise. Specifically, users have the opportunity to contribute to a comprehensive catalog of valuable data and models, share innovative solutions and connect with experts across many disciplines and domains.


Credit: David Esquivel, UCLA

SDSC Leads Effort to Enhance National Research Data Management Ecosystem Infrastructure

SDSC was awarded a U.S. National Science Foundation EArly-concept Grants for Exploratory Research (EAGER) grant to lead an effort to enhance the national research data management ecosystem infrastructure. The project involves using a system called Research Activity Identifier (RAiD), which was first developed by the Australian Research Data Commons.

SDSC's Director of Research Data Services, Christine Kirkpatrick, serves as the project lead and the center is the inaugural registration authority for the U.S. RAiD service, which provides researchers, funders, institutions and other stakeholders with a way to assign globally unique, persistent, and resolvable identifiers (GUPRIs) with corresponding metadata for specific projects or portfolios.

According to Kirkpatrick, RAiDs provide a transparent and reliable way to centralize and preserve vital project information, even after a project ends. This helps alleviate the administrative burden for researchers by simplifying tasks such as preparing funding applications, managing project collaborations, and generating reports.

The use of persistent identifiers (PIDs) is key to implementing some of the FAIR principles and will support new data architectures such as knowledge graphs to be more precise with context for AI readiness. Several countries have adopted national PID strategies, and the U.S. has an interagency working group currently working on a national strategy. Kirkpatrick noted that RAiD represents a missing piece needed for a successful PID strategy, giving stakeholders a way to find and collect information about projects.

The SDSC-led effort involves developing a hosted cloud service, a web application and an application programming interface (API) for issuing RAiDs, as well as creating a sustainable business model to support the service.

Biomedical Data Project Benefits Research Community

SDSC has been working with the GO FAIR Foundation, the National Center for Atmospheric Research and other partners on a project funded by the Frederick National Laboratory for Cancer Research. The project, led by SDSC's Research Data Services Director Christine Kirkpatrick, aims to improve the quality of metadata for biomedical research, making it findable, accessible, interoperable and reusable (FAIR).

The team has provided guidance on enhancing metadata quality within National Institute of Allergy and Infectious Diseases (NIAID) and National Institutes of Health (NIH) supported repositories and resources, and has offered structured trainings and guidance to support stakeholders. According to Kirkpatrick, the project will have a significant impact on the FAIR ecosystem and will be meaningful work for the team.

The project has also explored the relationship between FAIR resources and scientific impact through interviews with stakeholders and an assessment. The project will also work closely with another key project, the NIAID Data Ecosystem Discovery Portal, led by The Scripps Research Institute, to ensure repository improvements maximize the Discovery Portal's ability to search across NIAID-funded research data assets.

Accelerating Science through Training, Application & User Support

SDSC bridges gaps by providing support to facilitate the impactful use of data and computational resources, and its remains committed to the development of next-generation scientific leaders through its mentorship and experiential learning opportunities.

COMPLECS Award Funds SDSC to Conduct HPC Training for Domain Scientists

While researchers in various fields have the ability to find solutions to complex problems, often the implementation of their computational methods requires a machine that is bigger and more powerful than their desktop computers. Chemists, environmental scientists, geophysicists, biologists and others often find themselves in need of a supercomputer. These domain scientists and applied mathematicians typically work alongside a computational scientist to implement and run their workloads; however, thanks to an award from the U.S. National Science Foundation (NSF) to SDSC, a training program has been developed to better assist with learning supercomputing fundamentals.

"We are continuing to build on our earlier success with COMPrehensive Learning for end users to Effectively utilize Cyberinfrastructure, or COMPLECS, as we enter our second year," said Nicole Wolter, a member of the Data Enabled Scientific Computing Division (DESC) and principal investigator (PI) for the three-year $500,000 award.

In the first year, the COMPLECS training program, provided 17 webinars engaging with over 500 participants. The training material continues to evolve based on participant feedback to better address the fundamental three layers — beginning with foundational knowledge, such as parallel computing concepts and intermediate Linux, that serves as a base for learning other essential and specialized skills depending on the users' needs. The program hosts a multiday in-person workshop in the summer, an additional virtual two-day workshop to accommodate participants who might not otherwise be able to attend, in addition to the bi-weekly webinars as well as offering self-paced online study options.

Additional SDSC community members involved with the COMPLECS effort are Co-PI Martin Kandes, Advanced Computing Training Lead Mary Thomas, Senior Computational Scientist Bob Sinkovits, Director of User Services Mahidhar Tatineni and Director of Computational Chemistry Laboratory Andreas Goetz. The program is funded by the NSF Office of Advanced Cyberinfrastructure.

"One of our goals with the program is to recruit participants from domains that haven't traditionally used supercomputers in conjunction with our work on the ACCESS-CI project," Thomas said. "We are just thrilled to have this opportunity and grateful to the NSF for funding our work in this arena."

SDSC Offers ‘Real World’ Internships to Prepare Students for Big Tech Careers

SDSC offers undergraduate students “real world” internships to prepare them for big tech careers. The Research Data Services (RDS) Division at SDSC onboards students for positions in their agile application development program, where they work in a fast-paced environment to gain experience in project management, product management, front-end development and back-end development.

The program is led by Ryan Nakashima, Jenny Nguyen and Steven Yeu, who provide students with hands-on experience and exposure to different specialties.

One of the projects that student interns worked on was updating an app called NeuroRes, which was first created in 2017 by SDSC RDS for the UC San Diego Department of Neurosciences. The students, under the guidance of Nguyen and Yeu, added a feature that allows resident physicians to connect with attending physicians via the app.

In addition to NeuroRes, previous students worked on another health science project called CAPCHD, creating an app for heart disease researchers using AWS serverless microservices.

The RDS team actively seeks projects to be developed by the upcoming cohort of students. The program has been successful over the years, with many students going on to work in industry after developing skills during their time as interns.

HPC and Data Science Institute

Each year, SDSC hosts a week-long summer workshop focusing on introductory-to-intermediate topics in HPC, data science and AI. Designed for researchers and educators in academia and industry, the program equips participants to solve scientific challenges exceeding local computing capabilities.

CIML Summer Institute

The Cyberinfrastructure-Enabled Machine Learning (CIML) Summer Institute introduces researchers, developers and educators to the techniques and methods needed to migrate their ML applications from smaller, locally run resources, such as laptops and workstations, to large-scale HPC systems.

SDSC Participates in HPC Education Project to Produce New Cyber Training Catalog

SDSC Computational Data Scientist Mary Thomas has been participating in a U.S. multi-institutional initiative that streamlines discovery and collaboration through an innovative cyber training catalog. The High Performance Computing-Education (HPC-ED) project has integrated an intuitive interface, structured metadata and active community engagement.

"Our HPC-ED team is creating a system that allows training material owners to share their resources while maintaining ownership," Thomas said. "Organizations can also enrich their local portals with shared materials, ultimately broadening the reach of valuable educational tools."

The work has been published in the Journal of Computational Science Education.

In addition to SDSC's Thomas and Jiesen Zhang, co-authors on the paper are HPC-ED project team members Susan Mehringer, Richard Knepper and Zilu Wang of the Center for Advanced Computing at Cornell University; Katharine Cahill of the New Jersey Institute of Technology; Charlie Dey of the Texas Advanced Computing Center; Brian Guilfoos of the Ohio Supercomputer Center; David Joiner of Kean University; John-Paul Navarro of Argonne National Laboratory and Jeaime Powell of Omnibond Systems.

Thomas explained that a key strength of the HPC-ED project is leveraging the flexible and well-established Globus Search framework. She said that this architecture allows the team to offer a scalable solution for a wide range of users.

"With insights gained from our pilot phase, we are now evolving to implement new features designed to better meet community needs," Thomas said. "Enhanced tools and interfaces will cater to diverse preferences, such as enabling filtered search result downloads and integrating Jupyter Notebooks for streamlined resource sharing."

According to Thomas, by fostering collaboration and leveraging cutting-edge technologies, HPC-ED is well on its way to revolutionizing how cyber training materials are shared and discovered.

"Creating a federated repository for training materials supports the ability of computational scientists to learn and extends the offerings available through local research computing and data teams," said Knepper, director of the Cornell Center for Advanced Computing (CAC).

Mehringer, first author of the paper and an associate director at the CAC, concurred. "Building a well-defined catalog achieves two important goals: supporting researchers by enabling improved material discovery and reducing training development duplication by simplifying sharing," she said.

Open Storage Network Welcomes New Campus Computing Partners

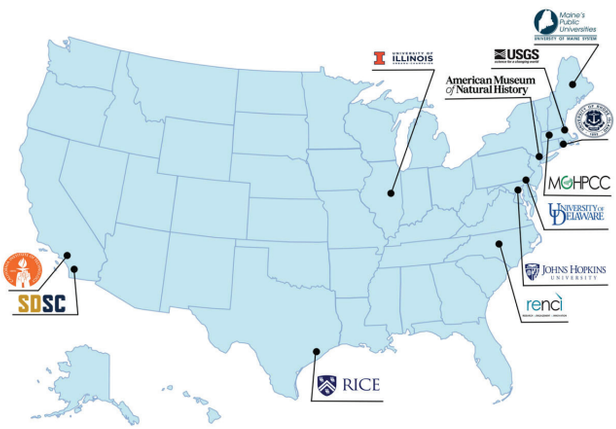
SDSC has been leading the Open Storage Network (OSN) program for years, and along with its collaborators – Massachusetts Green High Performance Computing Center (MGHPCC), National Center for Supercomputing Applications, Renaissance Computing Institute, Johns Hopkins University and Rice University – provides low-cost, quality, sustainable distributed storage cloud for the research community.

Most recently, SDSC Research Data Services (RDS) Director Christine Kirkpatrick worked with MGHPCC Executive Director John Goodhue on expanding the OSN to include more pods (i.e., storage nodes) at additional sites throughout the U.S. and welcomes new campus computing partners.

"The past decade has seen rapid growth of data sets from scientific instruments, simulations, internet postings and other sources – allowing new insights through big data analytics and more recently training of AI models," Goodhue said. "This torrent of data has created a need for a platform that supports simple and cost-efficient storage and sharing of large volumes of data."

In 2018, the OSN set out to address that need with a pilot that was sponsored by the U.S. National Science Foundation. Since then, the platform has evolved into a large-scale production storage system that supports allocations of up to 50 terabytes at no charge via the NSF ACCESS program, as well as paid participation for entities with larger needs.

"While the OSN started as a way to store and share data across geographically distributed sites, it has also been useful as a way to share data between different groups on individual campuses," Kirkpatrick said. "OSN continues to expand with several new pod sites for projects by the principal investigators writing campus computing storage into their proposals."



CORE Summit Focuses on Leveraging AI for Agriculture

The 2024 Convergence Research (CORE) Summit was held at the San Diego Supercomputer Center (SDSC) and brought together researchers and practitioners to explore innovative solutions for a more productive food system.

The summit featured a keynote address by Shefali Mehta, former Deputy Under Secretary of Research, Education, and Economics and Acting Chief Scientist at the U.S. Department of Agriculture (USDA), who discussed the importance of interdisciplinarity and cross-sector collaboration in using artificial intelligence (AI) for agriculture.

The summit also included a student Design Challenge, where students worked with the VINE, an initiative launched by the University of California Agriculture and Natural Resources (UC ANR), to imagine how to use data and AI to create a productive and resilient food system. Selected participants further developed their design concepts through internships with the CORE Institute.

The CORE Institute is an initiative of the Societal Computing and Innovation Lab at SDSC, launched in 2022 with support from the U.S. National Science Foundation Convergence Accelerator.

Congressional App Challenge Inspires San Diego's Next Generation of Innovators

Last fall, students from local middle and high schools received valuable insights from experts at SDSC during the 2024 Hackathon organized by U.S. Congressional Representative Scott Peters — a long-time champion for STEM education — as part of the annual Congressional App Challenge. The Congressional App Challenge is a nationwide competition, established in 2013, that encourages young people to explore STEM through app development. And, according to Peters, the annual event is a crucial platform for cultivating the talents of young innovators who will drive San Diego's burgeoning tech economy.

"The Congressional App Challenge offers students in middle and high school a chance to showcase their coding skills and creativity," said Peters. "I want to thank the San Diego Supercomputer Center at UC San Diego for once again hosting our annual Hackathon and helping to inspire the next generation of innovators and problem-solvers right here in California's 50th District."

Secondary school students from the district participated by learning tips on developing their own software applications, which can be designed for mobile, tablet or computer devices on any platform. This hands-on event served as an opportunity for students to not only improve their technical skills but also explore the limitless possibilities of coding, computer design and STEM in general.

As a strong advocate for preparing students for the tech-driven future, Peters emphasized the importance of the competition in empowering students to think creatively and embrace the wide range of applications via STEM — science, technology, engineering and math.

"The most important investment we can make is in the next generation," Peters added, "because these students will go on to power San Diego's innovation economy."

Winners of the challenge join others from across the country in receiving recognition in Washington, D.C., where their apps are featured as part of a Capitol Hill exhibit highlighting the achievements of young coders nationwide.

Translating Research into Impacts

SDSC advances scientific research in partnership with stakeholders across academia, government and industry to impact society for the better.

SCIL Director and SDSC Chief Data Science Officer, Ilkay Altintas. Credit: Erik Jepsen, UC San Diego Communications

SDSC Launches Societal Computing and Innovation Lab

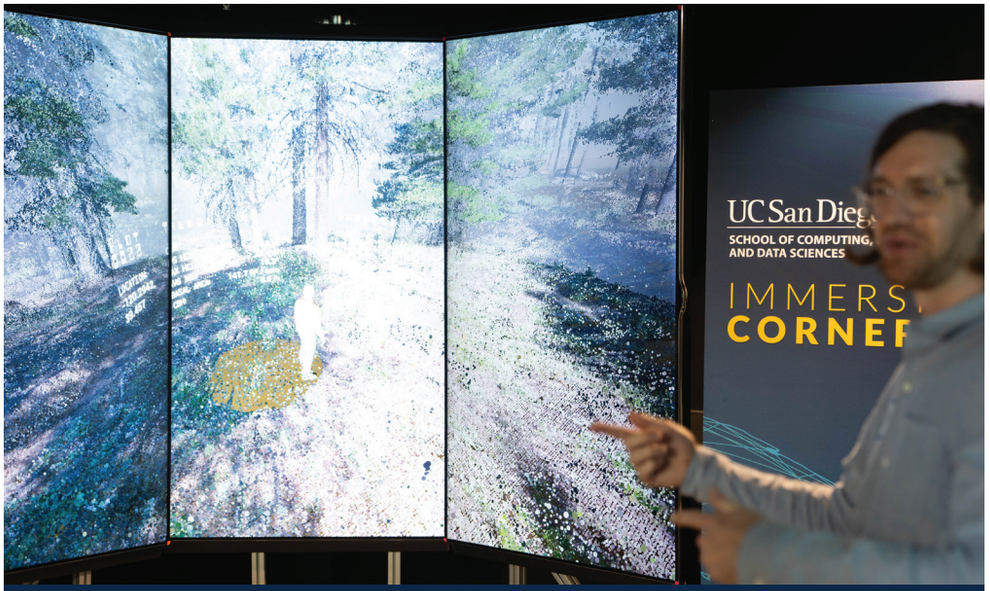
The Societal Computing and Innovation Lab (SCIL) held its official launch event in May 2025 at the SDSC. Created by SDSC's Chief Data Science Officer Ilkay Altintas, SCIL pioneers innovation pathways to address complex societal challenges. At the heart of SCIL's translational approach is a commitment to moving from use-inspired problems to scalable, real-world solutions.

The lab's work is informed by over two decades of Altintas' leadership in scientific workflows and computing and a decade of collaboration through the WIFIRE Program with the fire management community.

According to Altintas, "SCIL grew out of our deep commitment to identifying critical gaps in today's solution landscape – gaps where we can uniquely innovate at the intersection of science, technology, and data and AI to drive real-world, transformative impact on society's most pressing challenges."

SCIL offers data platforms and fire management solutions that are powered by dozens of research projects and partnerships aimed at translating science into practice. This work is grounded in five primary research areas:

1. integrated research workflows,
2. composable systems across the digital continuum,
3. data and knowledge systems,
4. digital twins for complex systems and
5. agentic AI for scientific applications.



Isaac Nealey gave a presentation using the Immersion Studio during the launch event.
Credit: Bobby Yu, UC San Diego

The renowned WIFIRE Program — which provided valuable predictive mapping data to responders during the recent fires in Los Angeles — develops technologies with the fire management community driven by cutting-edge science and data to increase resilience to wildfires. Meanwhile, additional SCIL programs develop and support a suite of open, interoperable data platforms that serve as critical infrastructure for collaborative research, decision-making and real-world impact. These platforms include the National Data Platform, which has earned national recognition as a "Project to Watch," as well as the NSF Quantum Foundry's Quantum Data Hub and the Wildfire Science & Technology Commons.

"SCIL's work will build a stronger, more connected innovation ecosystem by reimagining how university researchers and cross-sector community partners can work together to tackle pressing societal issues," said SDSC Director Frank Würthwein.

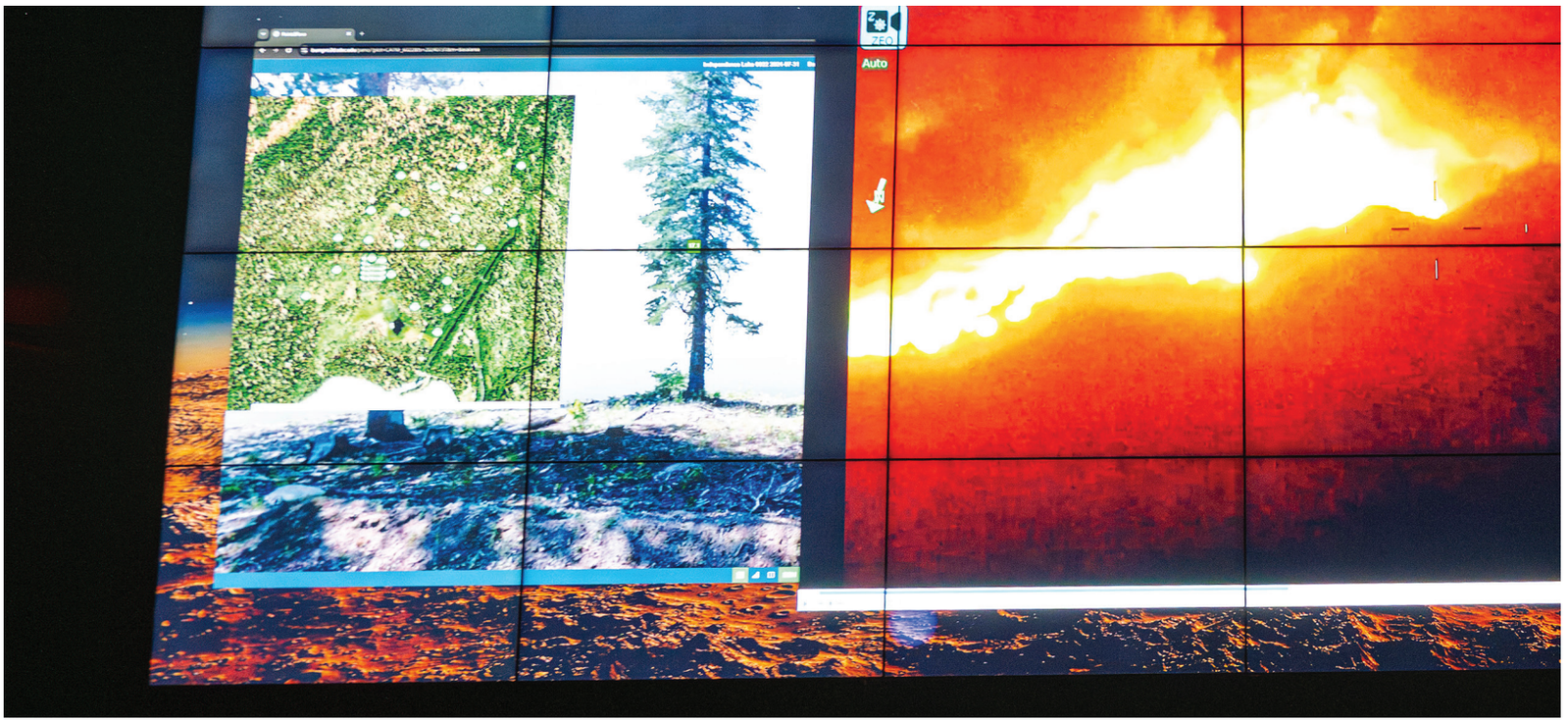
A defining feature of SCIL is its Immersion Studio, which leverages the AI-readiness of scientific data to power new modes of teaching, training, decision-making and public engagement. This dynamic research and visualization space offers large-scale multi-panel displays and virtual reality experiences.

"From 3D vegetation models and fire science simulations to real-world data from high-resolution sensors, the immersion studio brings complex information to life through cutting-edge visualization," said SCIL's Chief Innovation Officer Melissa Floca.

SCIL's commitment to solving complex societal challenges requires more than technical knowledge — it entails hands-on engagement with real-world problems — providing experiential education opportunities. These activities train the next generation of innovators through a variety of programs such as data challenges, the CORE Institute and the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship.

"The lab's experience, applications and partnerships will inform a specialization that provides students with a foundation to imagine, build and scale solutions to society's most pressing problems," said Rajesh K. Gupta, SCIDS interim dean and founding director of the Halıcıoğlu Data Science Institute.

SCIL's projects are supported by funding from SDG&E, California agencies, NSF, NIST, DHS, USDA, DHS, NOAA, the Moore Foundation, UCOP, ESTCP, NASA and others.



WIFIRE Program: A Useful Resource for Responders

On January 7, 2025, when a fire broke out in the Pacific Palisades neighborhood in Los Angeles, the WIFIRE team at SDSC was ready to respond. They used their predictive software, Firemap, to create maps of the fire's spread and provided this information to responders within minutes of the fire's ignition. The team worked with the California Governor's Office of Emergency Services to provide data on the fire's location and predicted path, aiding in evacuation efforts and damage assessment. The team also used UC San Diego's camera and sensor network, ALERTCalifornia, to triangulate the fire's position and run the predictive models.

The WIFIRE program, is part of UC San Diego's new Societal Computing and Innovation Lab (SCIL) at SDSC. The program's goal is to provide accurate information to responders so they can deploy resources effectively and prevent catastrophic wildfires. The program has been successful in providing real-time insights, which has been transformational in responding to wildfires.

"Initial attack is the first moments of responding to a fire, and accurate information is crucial," said Ilkay Altintas, SDSC chief data science officer and founding director of WIFIRE. "Providing data on where the fire is and where it will be is our priority. When responders have the right information to deploy resources they are better equipped to prevent catastrophic wildfires."

A HISTORY OF COLLABORATIVE INNOVATION

The WIFIRE program at UC San Diego has evolved over the years through innovative ideas and unique partnerships. It originated from the High Performance Wireless Research and Education Network, which recognized the potential for public safety applications after the 2003 Cedar Fire and 2007 Witch Fire in San Diego. Altintas began exploring how to provide responders with real-time, data-driven insights, and the program started with federal funding and support from the U.S. National Science Foundation.

WIFIRE takes a use-inspired approach to making data and fire science actionable for decision support, working closely with partners to develop sustainable and scalable solutions. The program has partnered with various fire departments, including the Los Angeles Fire Department, and has expanded to support similar fire resource networks in other states.

WIFIRE's goal is to bridge the gap between research and development and the frontlines, ensuring that the tools they create are useful and actually used. The program has also partnered with utility companies and government agencies to advance its capabilities and improve preparedness and response strategies.

Credit: Erik Jepsen, UC San Diego Communications

PROACTIVE PARTNERSHIPS GOING FORWARD

"We can't respond our way out of this wildfire crisis," said Chris Anthony, the former chief deputy director of CAL FIRE. Anthony was an advocate for the work being done at UC San Diego during his time in the fire service; in his retirement, he became a fellow of the ProWESS Center, a collaboration between UC San Diego and the Los Alamos National Lab focused on proactive measures to prevent and reduce wildfire risk.

"I saw the impact UC San Diego had on the fire service community — not only in aiding response but in applying data to improve fire resilience across the state," he said. "New technology is necessary to significantly scale prescribed fire programs and make wildlands more resilient."

In terrain that has not seen a significant fire for decades, built-up fuels increase the risk of catastrophic fire spread. But this can be mitigated by reintroducing fire through controlled, prescribed burns — controlled being the key word. And for greater control, one needs better information.

Drawing from their expertise in understanding and modeling fire behavior with WIFIRE, the SCIL team built the modeling program BurnPro3D to help prescribed burn crews plan and predict what a controlled fire will do under different weather and fuel conditions. This is a powerful tool for ensuring these risk-reducing burns are conducted efficiently and safely, and detailed three-dimensional models bring prediction analysis to another level.

"BurnPro3D is designed to support the preparation of burn plans as well as the implementation of burns," Anthony said. "The interface allows burn bosses to create and visualize high-resolution 3D fire simulations and compare fuel consumption and risk under different weather and ignition scenarios."

FIRE BURNS IN BRUSH ON THE SIDE OF THE ROAD

WIFIRE tools like BurnPro3D and Immersive Forest produce digital models of actual wildland terrain to inform prescribed burns. The lab is also advancing modeling with the Immersive Forest Project, which produces digital models of actual forests to virtually simulate the complexity of fire to train researchers, educate policymakers and attract more AI technologists into firetech — all without the need for live burns. And even beyond the fire, the lab is now incorporating smoke management into predictive models, a critical factor for public health and safety.

"We're doing all these things, but we're not doing them alone," said Melissa Floca, Strategic Partnerships director for CICORE and chief innovation officer for SCIL. "Our partnerships and collaborations have helped grow a bigger community of researchers, fire practitioners and public and private organizations that all want to contribute to solutions."

HPWREN's Quarter-Century of Wireless Innovation

The High Performance Wireless Research and Education Network (HPWREN), a University of California San Diego partnership project led by SDSC's Hans-Werner Braun and the Scripps Institution of Oceanography's Institute of Geophysics and Planetary Physics' Frank Vernon, supports Internet-data applications in the research, education and public safety realms.

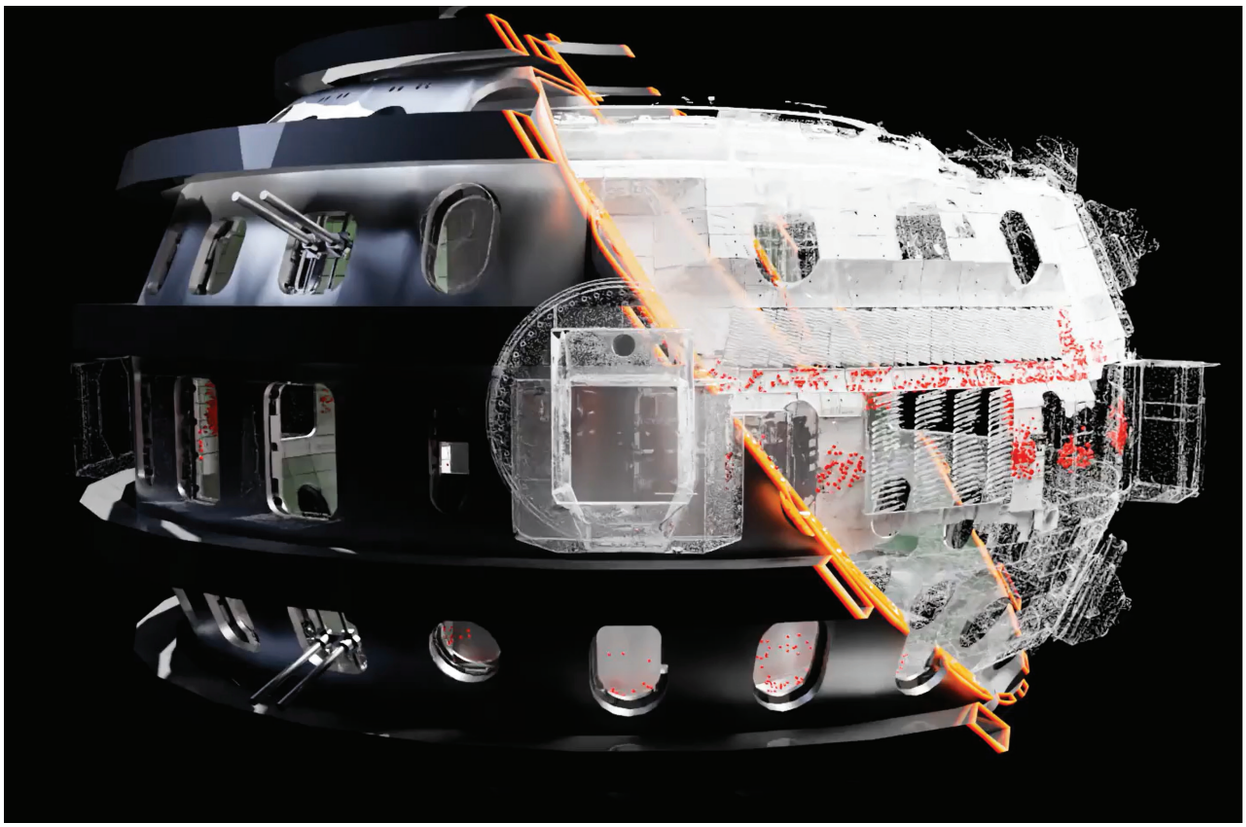
HPWREN functions as a collaborative, Internet-connected cyberinfrastructure. The project supports a high-bandwidth wireless backbone and access data network in San Diego, Riverside and Imperial counties in areas that are typically not well-served by other technologies to reach the Internet. This includes backbone locations, typically sited on mountain tops, to connect often hard-to-reach areas in the remote Southern California back country.

HPWREN marks a milestone this year as it recognizes the program's 25th anniversary of groundbreaking research. In 2000, with funding from the U.S. National Science Foundation, HPWREN established the first wide-area wireless communications research network. The success of HPWREN's groundbreaking network created the foundation for ALERTCalifornia, an award-winning UC San Diego research and public safety program that incorporates the HPWREN cameras and technology into a growing statewide network of more than 1,150 natural hazard monitoring cameras used by emergency managers and the public.

Today, HPWREN continues pioneering research and providing connectivity across remote backcountry. The HPWREN cameras are an essential tool for emergency response during natural disasters. In September 2024, the Airport Fire threatened communities in Orange and Riverside counties and critical mountaintop infrastructure on Signal Peak. The HPWREN cameras provided situational awareness and incident monitoring for CAL FIRE and local emergency managers. Despite damage, the HPWREN cameras aided response through the incident and the field team was on site as soon as it was deemed safe to assess and make repairs. The Airport Fire is just one example of how HPWREN's 25-year legacy of research and to-the-minute data collection and distribution impacts the lives of Californians, and will continue into the future.

CAIDA

Founded in 1997, the Center for Applied Internet Data Analysis (CAIDA) conducts network research and builds research infrastructure to support large-scale data collection, curation and data distribution to the scientific research community. CAIDA is based at SDSC, part of the CICORE Division, at UC San Diego. For more information about CAIDA's work, please visit: www.caida.org



General Atomics and UC San Diego Launch Fusion Data Science and Digital Engineering Center

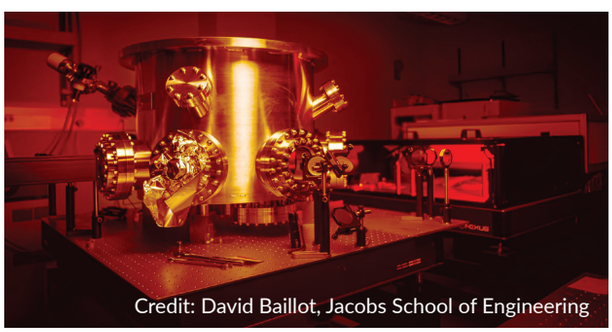
General Atomics (GA) and UC San Diego have launched a Fusion Data Science and Digital Engineering Center in San Diego to strengthen California's leadership in fusion energy innovation. The center aims to advance digital engineering, artificial intelligence, machine learning and high-performance computing to fast-track fusion energy development.

According to Raffi Nazikian, director of fusion data science at General Atomics, achieving fusion energy requires a new approach to fusion data science that leverages AI, ML and HPC to optimize system design and reduce development cycles.

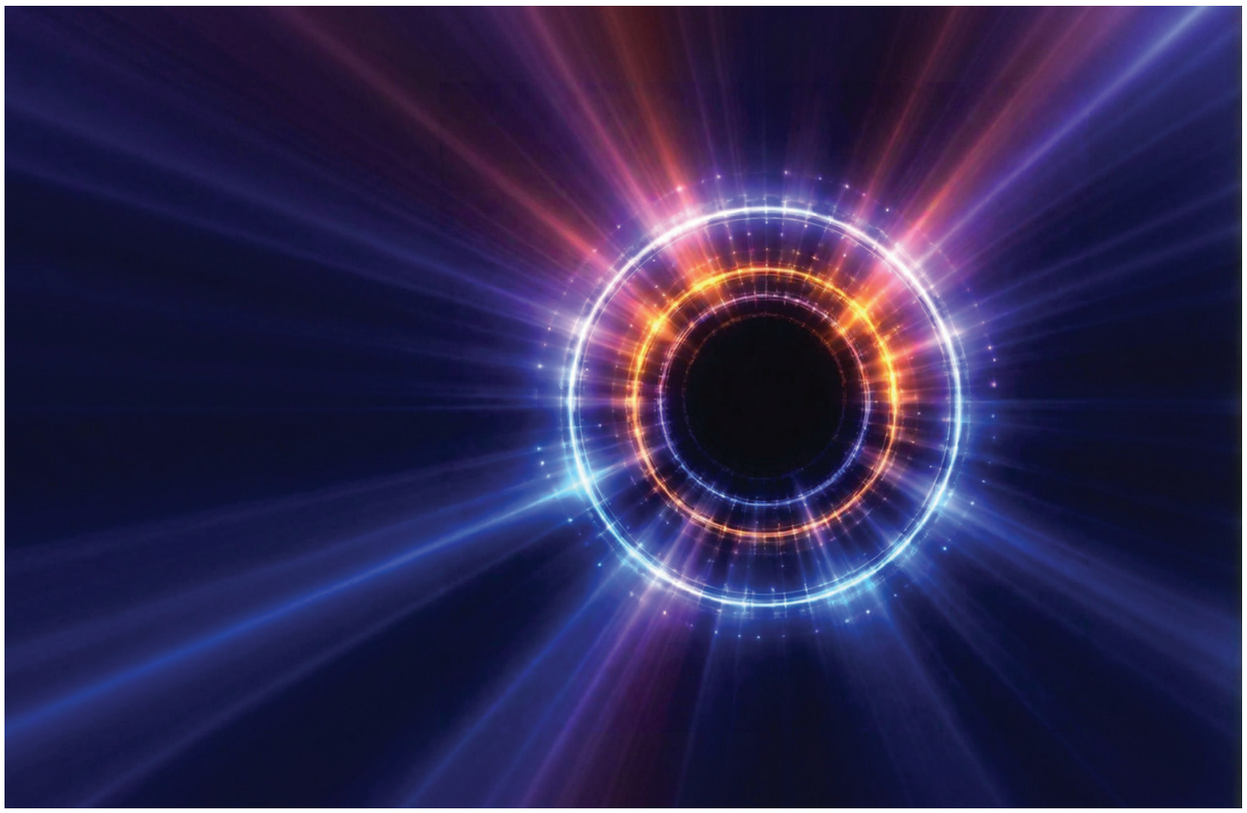
The new center is part of UC San Diego's Fusion Engineering Institute, which brings together students, faculty and external partners to address engineering challenges in fusion energy research and education. The institute connects people and projects across the entire campus, including the Jacobs School of Engineering and SDSC.

The center also builds on the collaboration between GA and UC San Diego, which established SDSC 40 years ago.

"GA and UC San Diego came together to establish the San Diego Supercomputer Center 40 years ago thanks to funding from the U.S. National Science Foundation. The new center thus builds on a rich history of cooperation and will leverage significant existing investments to establish shared digital infrastructure," said Frank Würthwein, SDSC director. "We look forward to building cyberinfrastructure for fusion engineering as a public-private partnership with federal, state and industry investments."



Credit: David Baillot, Jacobs School of Engineering



A Promising Alternative to Chemotherapy

Scientists are working to develop noninvasive cancer treatments like photodynamic therapy (PDT), which uses light and chemicals to target cancer cells. An international research team from MIT and Imperial College used the Expanse supercomputer at SDSC to study iridium-centered chemicals that could be used for PDT and bioimaging of cells.

Their findings, published in Angewandte Chemie, indicate that some of these chemicals are promising for noninvasive cancer treatments, as they can produce reactive oxygen species that lead to cell death when illuminated with light. The team used computer simulations on Expanse to understand why certain iridium complexes perform better than others. Using Expanse allowed the researchers to complete this part of their work in just three days — thanks to Expanse's fast processors.

"Using our ACCESS allocations on Expanse, we quickly found that the properties of each of the iridium-centered chemicals could be better understood in the context of the experiment," said MIT Chemical Engineering Professor Heather J. Kulik, who has been using NSF ACCESS allocations on Expanse for several years on an array of studies.

She worked with Gianmarco Terrones, an MIT chemical engineering graduate student researcher, to calculate properties of specific chemicals, namely the energies of their excited states and the energies of their orbitals.

The researchers believe that their findings could significantly advance PDT and provide new opportunities for noninvasive cancer treatments.

SDSC Hosts New SDSU Portal for Healthcare

For the past decade, Rodrigo Valdez has been working on various software and web development projects, including a recent project with a team at San Diego State University (SDSU) to develop a portal called HealthLINK Center Health Science Research Portal (HSRP).

The portal provides health science data related to under-resourced healthcare professionals and has been developed in conjunction with the Hubzero team at SDSC. The portal offers a range of population-based research data that can be used for biomedical, behavioral and clinical services.

The SDSC team, including Project Manager Claire Stirm and Senior Software Engineer Nick Kisseberth, worked with Valdez to customize the portal for role-based management. Portal features include an online repository for sharing and archiving research-related resources, including technical reports, lab protocols and training materials.

According to Stirm, the team has enhanced the Hubzero resources to provide curated access to specific user groups. The portal aims to support health and health disparities research by providing a platform for users to access and share resources.

“To ensure that various types of users and professionals can easily access materials on the HSRP portal while respecting role-based access, we have worked with Valdez and the SDSU team to enhance the Hubzero resources so they are curated to a specific access group,” Stirm said. “It is exciting to see the user community grow, and the resources grow on HSRP.”

Using Supercomputers to Help Design Quality Recyclables from Plastic

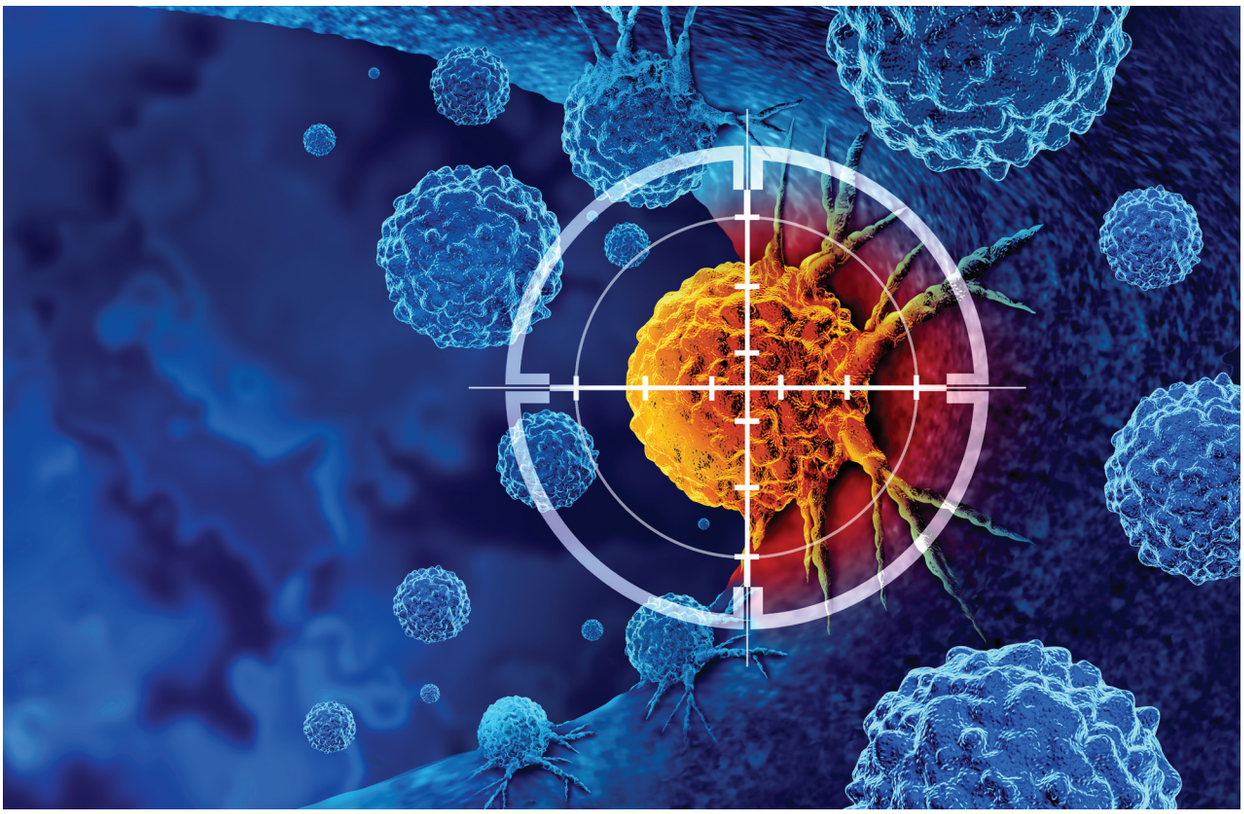
Plastic products are widely used in our daily lives, but their disposal poses a significant threat to the environment and human health. To address this issue, researchers at Northwestern University, led by Senior Associate Dean Linda Broadbelt, are using computational models on the Expanse supercomputer at SDSC to develop a chemical recycling process.

This process involves breaking down plastics into their basic building blocks, called monomers, and the team has published their latest models in the Cell Reports: Physical Science journal.

Broadbelt's work focuses on understanding the "ceiling temperature" (Tc) of polymers, which determines their ability to be recycled.

"By combining experimental data with Expanse simulations, we can better understand Tc and predict key thermodynamic parameters," Broadbelt explained. "When we can see the details of how polymers behave under different conditions, we can design more sustainable materials."

SDSC's Expanse supercomputer



Enabling Breakthrough Immunotherapy Study

Researchers at the La Jolla Institute for Immunology made a groundbreaking discovery that helps understand why the immune system is not as aggressive toward self-antigens. This discovery, published in Nature Immunology, could lead to new ways to improve cancer immunotherapy.

The researchers used the Expanse supercomputer at SDSC to study how cancer tumors react when the immune system tries to fight them. They found that two proteins, PD-1 and CD73, play a significant role in suppressing the activity of immune cells called CD4+ T cells, which are responsible for responding to self-antigens.

By blocking these proteins, the researchers were able to boost the response of self-specific CD4+ T cells, allowing them to behave more like foreign-specific cells.

This finding was made possible through a collaboration between researchers with different areas of expertise and funding from organizations such as the U.S. National Institutes of Health (NIH) and the U.S. National Science Foundation (NSF).

“Thanks to funding from NIH, NSF and additional funding sources, we were able to make a significant impact on an important problem using a relatively modest amount of computation,” said Bob Sinkovits, a senior computational scientist and Expanse co-principal investigator at SDSC. “The collaboration brought together teams with complementary areas of expertise to achieve results that might not have otherwise been possible, hopefully one day leading to improvements in the treatments of cancers that exploit tolerance mechanisms to evade immune detection.”

Virtual Environments for Science

SDSC makes data and computationally intensive capabilities accessible to millions of researchers, educators and students across disciplines through science gateways, secure cloud enclaves, research software development, and immersive visualizations.

Scientists Combine Immersive Technology, Collaboration to Address Challenges

A team of scientists and science communicators affiliated with SDSC and the Center for Coastal Climate Resilience at UC Santa Cruz have developed an initiative called EcoViz focused on immersive technology — advanced hardware and software to simulate environments and experiences — and collaboration to address environmental challenges ranging from fire management to reef restoration and more.

Jessica Kendall-Bar, an Eric and Wendy Schmidt AI in Science Postdoctoral Fellow from Scripps Institution of Oceanography, along with SDSC Chief Data Science Officer Ilkay Altintas and researchers from UC Santa Cruz's Center for Coastal Climate Resilience, presented EcoViz at the 2024 IEEE VIS Workshop on Visualization for Climate Action and Sustainability.

To demonstrate the initiative aimed at making climate data clearer and more engaging, the team developed three major visualization use cases, each tailored to illustrate complex, time-dependent environmental changes. Using a mix of cinematic techniques, scientific data and interactive formats, EcoViz illustrates complex datasets ranging from coastal flood protection benefits and marine animal behavior to wildfire management.

According to the presenters, some visualizations were most effective as narrative-driven animations, while others gained traction through being interactive by allowing users to explore geospatial data and experience ecosystems virtually through immersive technology.

Altintas, who co-leads the Schmidt AI in Science Fellowship and is the Director of the Cyberinfrastructure and Convergence Research Division at SDSC, said that by combining visual creativity with scientific rigor, EcoViz is helping to transform raw climate data into stories that can inspire informed policy and resilient climate solutions.

"As we integrate AI into scientific workflows that form the basis of decision-making, visualizations are key to making our models interpretable," she said.

The National Data Platform is designed to increase access to and usage of large datasets, such as those from solar simulations created by Curt Dodds and colleagues at the University of Hawaii's Institute for Astronomy. Image: Daniel K. Inouye Solar Telescope. Credit: DKIST/NSO

National Data Platform Recognized as a Project to Watch

In 2024, UC San Diego and the University of Utah announced a national-scale pilot project, called the National Data Platform (NDP), aimed at a service ecosystem to provide broad access to and use of scientific data across a range of communities.

Since then, NDP, led by SDSC and Utah's Scientific Computing and Imaging Institute (SCI), in partnership with the EarthScope Consortium, has been recognized in the Ninth Annual BigDATAwire Readers' and Editors' Choice Awards for their leadership in the NDP, citing it as one of the "Top 3 Big Data and AI Open Source Projects to Watch." Additionally, NDP received an Editor's Choice Award from HPCwire.

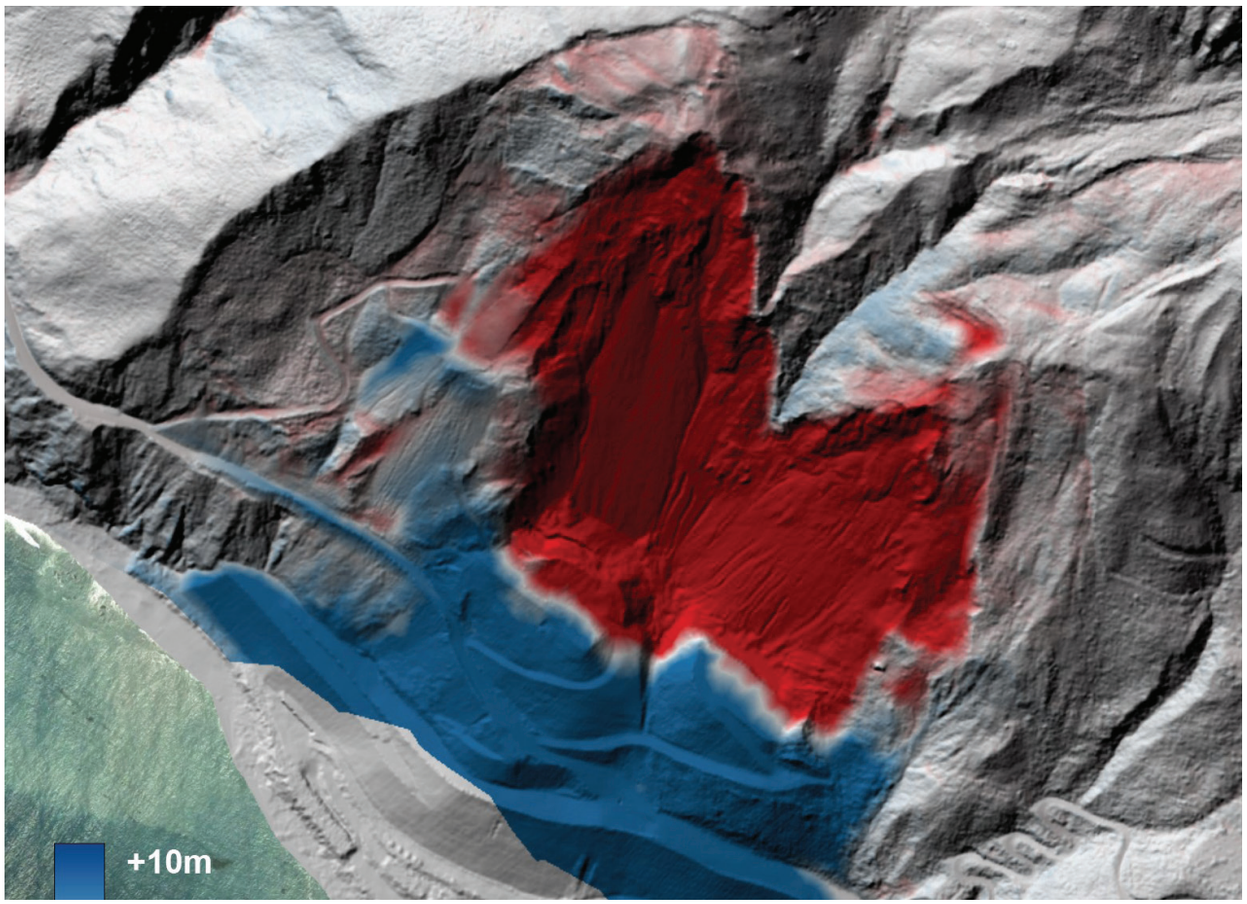
According to Ilkay Altintas, SDSC's Chief Data Science Officer, these awards reflect the dedication of the community of collaborators who share a vision of democratizing access to AI-ready data and computing resources.

The $6 million NDP pilot, funded by the U.S. National Science Foundation, serves as a federated and extensible data and service ecosystem to foster innovation, discoveries and collaboration through the broad access and use of science data and leveraging existing national cyberinfrastructure capabilities. Such access and use helps to ensure responsible data-driven research to address urgent national and global issues such as climate change and environmental sustainability through AI-integrated solutions.

Additionally, with the increasing potential of artificial intelligence to enhance and accelerate solutions to many scientific and societal problems, broad access to AI-ready data repositories is essential in developing and deploying responsible AI models and enabling everyone to be a part of AI-integrated solutions.

"NDP aims to bridge the gaps between data innovations and computing infrastructure through the combination of a data hub and an extensible service platform. Carefully designed workflows based on user needs assessment aim to make it possible for everyone to participate in AI-integrated solutions for research discoveries and global societal challenges," said Altintas, NDP principal investigator.

SDSC Director Frank Würthwein stated in an article last year that "NDP builds a data and knowledge curation layer on top of low level content delivery networks like the Open Science Data Federation, thus leveraging prior and contemporary investments in cyberinfrastructure across dozens of academic institutions."



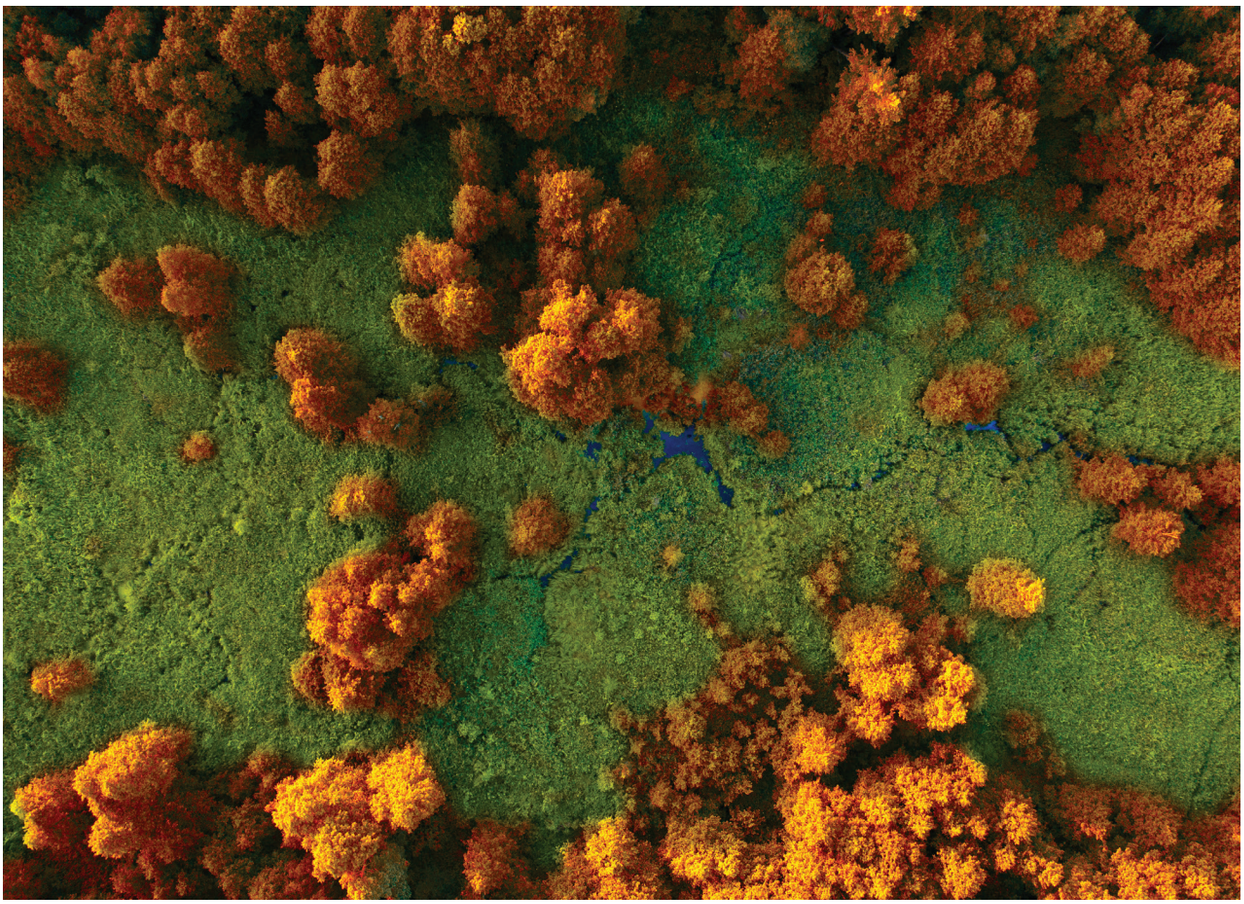
OpenTopo Supports AI-Ready Access to Topographic Data for Research and Education

The NSF has renewed funding for OpenTopography, a platform that provides open access to high-resolution topographic data. The $4.18 million award supports the next generation of the project over four years, focusing on making data AI-ready for emerging use cases in Earth sciences.

OpenTopography is a comprehensive platform for open access to high-resolution (one meter pixel resolution or better) and global topographic data that is utilized by an international community of researchers, educators, government agencies, industry and hobbyists. It is operated by SDSC; EarthScope Consortium, a non-profit university-governed consortium facilitating geoscience research and education; and Arizona State University's School of Earth and Space Exploration.

The project aims to expand pathways for efficient machine-readable data access, ensure availability of detailed metadata and promote trustworthiness and reproducibility. This will enable researchers to leverage essential data products and technologies to address complex Earth science challenges with greater precision.

"A key aspect in harnessing the power of artificial intelligence and machine learning in Earth sciences is ensuring that data facilities make their data AI-ready for emerging use cases," said Viswanath Nandigam, the project's principal investigator and director for the Advanced Cyberinfrastructure Development Lab within the SDSC CICORE Division. "AI-ready data are essential for advancing transformative AI research, as it provides more accurate and trustworthy insights to address complex Earth science challenges with greater precision."



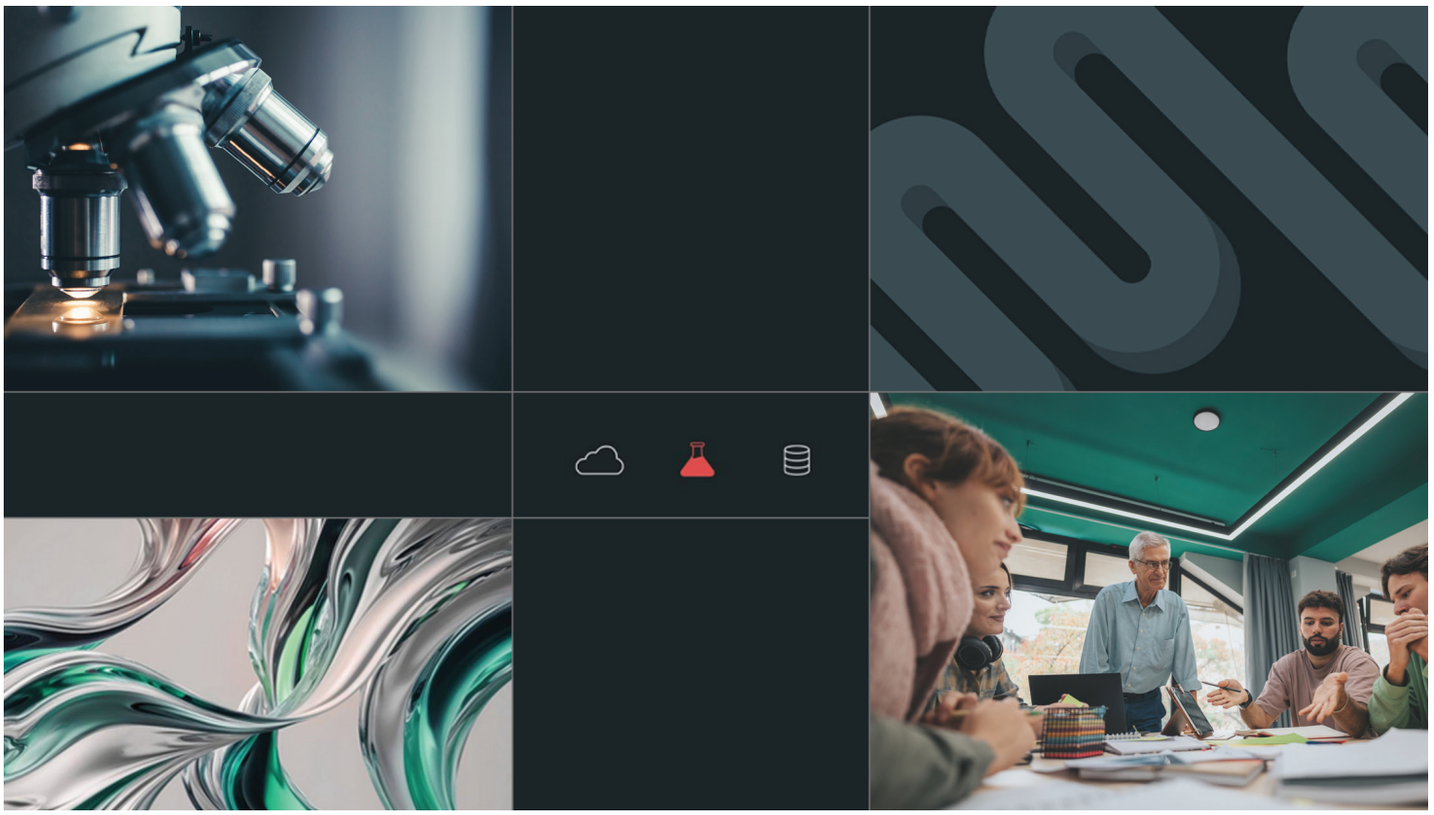
Researchers Develop Next-Generation Forest Mapping and Monitoring

Researchers at the University of California San Diego, University of Florida and Arizona State University (ASU) were awarded $3.28 million from the U.S. National Science Foundation (NSF) to build OpenForest4D – a web-based cyberinfrastructure platform for next-generation 4D forest mapping and monitoring.

The project aims to improve the mapping and monitoring of global forest ecosystems by combining the latest remote sensing data and artificial intelligence (AI) models to estimate forest structure and above-ground biomass over time. This helps generate research-grade estimates across various timescales, providing valuable insights into forest ecosystems.

The OpenForest4D team, led by principal investigators Viswanath Nandigam (SDSC), Carlos Silva (Florida) and Chelsea Scott (ASU), aims to develop a cyberinfrastructure framework to calculate forest structure and above-ground biomass estimates using remote sensing data. Currently, there is no accessible platform for efficient data processing, but the team plans to bridge the expertise gap between domain and data sciences, enabling researchers to leverage essential data products and technologies.

According to Silva, advances in AI will enable academics to conduct transformative research in forest sciences and ecology. Scott added that OpenForest4D will be a valuable STEM educational resource, providing educational resources and live webinars to develop a productive STEM workforce. The project will democratize scientific computing, allowing users of all expertise levels to access advanced forestry analysis tools, which is crucial as wildfires.



Stack Science Contributes Innovative CI and Tech Expertise to Research and Teaching

In a rare change for the organization, SDSC has merged two of its Centers of Excellence, Sherlock Regulated Data Management & Secure Cloud and Science Gateways & Research Software, to create a new division called Stack Science, bringing new capabilities to the communities SDSC serves.

According to SDSC Director Frank Würthwein, the merger is an example of how changes can help achieve organizational efficiencies and strategic goals.

“Such a merger is rare at SDSC, and with Stack Science, we break new ground,” Würthwein said. “The new division is an example of how such changes can be instrumental in realizing organizational efficiencies and strategic goals. We strongly believe this merger opens up new avenues of opportunity to pursue.”

The new division combines technology expertise and science focus, suggesting a layered approach to build upward. It also brings together the Sherlock Division, which focused on infrastructure and cybersecurity, and the Scientific Software Solutions (S3) Division, which centered around software.

This integration enables SDSC to provide end-to-end cyberinfrastructure capabilities covering the full stack.

“The S3 and Sherlock divisions were among the top 10 accomplishments by SDSC in the last decade. S3 was impactful for the access to advanced computational resources it provided to millions of researchers through sustained leadership in science gateways. Sherlock established a one-of-a-kind, nationally recognized center of excellence in regulatory data management,” noted Würthwein, adding that both divisions were aligned in their missions to serve the research and scientific community. “Both were independent and successful divisions. Both practiced a solution-focused, customer-centric and service-oriented approach to engagement.”

Stack Science Offers Secure Enclaves

Over the past year, the Stack Science Division at SDSC has expanded its partner institutions while supporting academic, government and industry partners with compliant and secure environments for sensitive data through its Sherlock services. These services provide tailored solutions to protect varied data, including those supporting healthcare research and government initiatives. Through these efforts, Stack Science continues to advance its mission of delivering trusted, secure enclaves.

Building on its commitment to provide secure and compliant data solutions, Stack Science is forging new partnerships, such as with Microelectronics Center of North Carolina (MCNC), to develop the MCNC Secure Enclave, a secure data environment, enhancing cybersecurity and compliance. As a leader in research and education networking, MCNC established the North Carolina Research and Education Network (NCREN) and seeks a long-term collaboration to leverage Stack Science's expertise. This partnership further enables Stack Science to create a replicable and scalable template for other research and education networks nationwide.

Beyond partnerships, Stack Science continues to enhance its own security framework. In alignment with these efforts, Stack Science has made significant strides toward NIST 800-171 and CMMC compliance for its Sherlock Government Cloud service, reinforcing its commitment to protecting sensitive data. With the Department of Defense set to include CMMC requirements in contracts, Stack Science is well-positioned to support partners requiring a CMMC-compliant environment. Additionally, Stack Science completed a SOC 2, Type2 assessment for its Sherlock Cloud service offering.

NIST Practices at SDSC

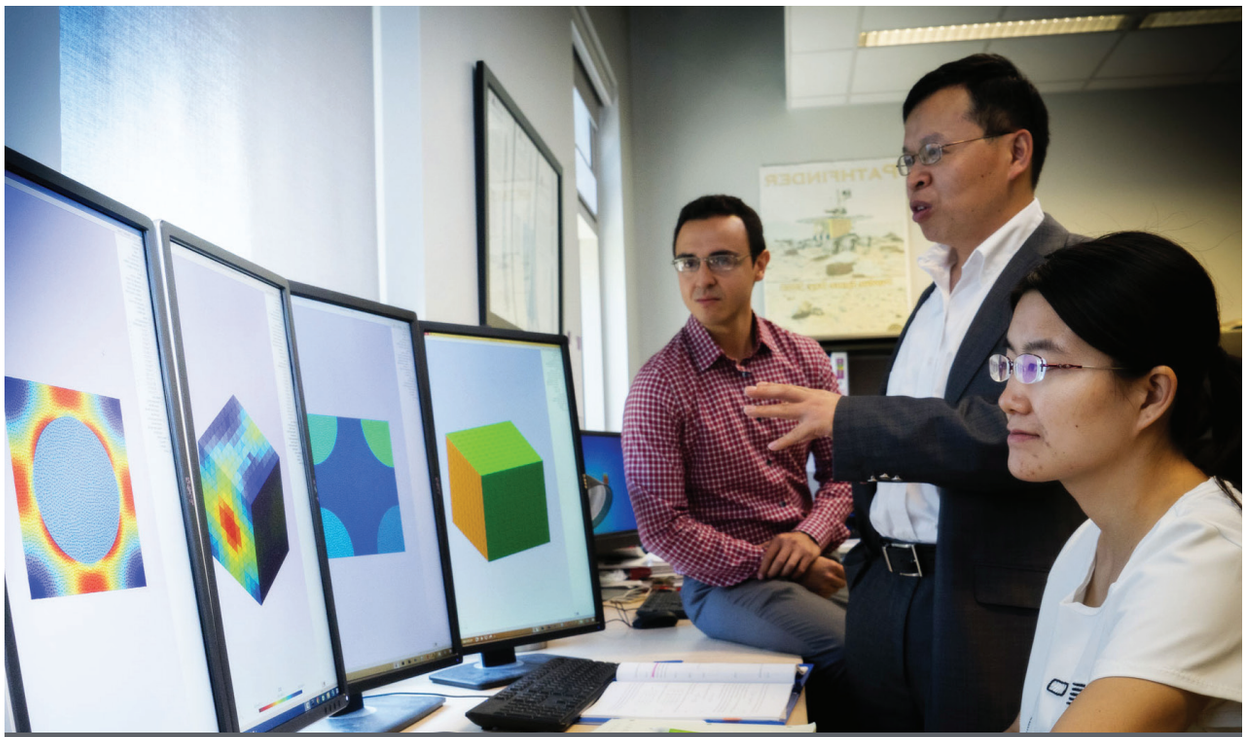
To work with the federal government, contractors must follow guidelines to protect sensitive information, known as Controlled Unclassified Information (CUI). The National Institute of Standards and Technology (NIST) created Special Publication 800-171 to standardize practices for safeguarding CUI. This publication outlines 17 control families, such as Access Control and Incident Response, to protect CUI confidentiality in nonfederal systems.

SHERLOCK

NIST 800-171 requirements apply to systems that process, store or transmit CUI, and are the foundation for the Cybersecurity Maturity Model Certification (CMMC) framework, which the Department of Defense is implementing this year to verify compliance with cybersecurity requirements. SDSC's Sherlock, a provider of secure Cloud infrastructures, has incorporated NIST 800-171 requirements into its services, building on its existing security frameworks. Sherlock's CUI-compliant solution includes robust security controls, such as log collection, encryption and malicious software protection, and is available on-premise and in AWS.

TRITON SHARED COMPUTING CLUSTER (TSCC)

TSCC, UC San Diego's campus research HPC cluster, has initiated efforts to meet MIST 800-171 requirements. This work involves close collaboration with administrative and IT stakeholders from main campus and health sciences, alongside biomedical research groups to ensure NIST 800-171 compliance while maintaining support for secure genomics research.



Science Gateways Ease Access to Multiple Resources

Science gateways are web-based portals or suites of applications that provide scientists with access to cutting-edge research tools, such as telescopes, supercomputers and undersea sensors. These gateways connect an array of resources and make them easily accessible, saving researchers and institutions time and money. For example, a single portal can give thousands of users access to optimized versions of analysis codes, allowing researchers to focus on their scientific goals without needing to know how supercomputers and data cyberinfrastructures work.

Science gateways have shown significant growth in terms of users, processing hours and published research papers. They can also be used for teaching classes, workshops and tutorials without requiring set-up on high-performance computing (HPC) resources or creating new accounts for students and attendees.

The Center of Excellence for Science Gateways (SGX3) at SDSC supports the science gateway community, providing training, consulting and good practices to help researchers and developers design and operate robust, scalable gateways.

"Our services provide all the tools necessary to foster collaborative communities, create science gateways, develop research software and establish secure enclaves to protect data and user workflows. We take a consultative approach, customizing our solutions to meet the unique needs of each stakeholder, all while ensuring our cyberinfrastructure complies with agency requirements," said Stack Science Director Sandeep Chandra.

Sandra Gesing, executive director of the U.S. Research Software Engineer Association and director of SGX3, is involved in national efforts to broaden access to computational tools and data, including gathering feedback for the design of a National Artificial Intelligence Research Resource (NAIRR) portal.

"We believe that science gateways are essential infrastructure," said Gesing, also a senior researcher in the Stack Science Division. "The NAIRR portal should be built for everyone: researchers at large institutions and small colleges alike, students just starting out, and researchers and educators not familiar working with advanced cyberinfrastructure. That's why our first step is to listen."