Voyager Taking Shape for its 2021 Debut

SDSC’s latest supercomputer, an experimental system scheduled for a late 2021 debut, is rapidly taking shape as a high-performance resource for conducting artificial intelligence (AI) research across a wide swath of science and engineering domains. In early April SDSC and Intel’s Habana Labs announced they will use Habana’s cost-effective yet powerful system to accelerate the development of AI algorithms as part of Voyager’s architecture.

Voyager is funded by a five-year, $5 million National Science Foundation (NSF) grant for system build-out, with additional funding for operations and community engagement. It will be the first-of-its-kind available in the NSF resource portfolio. The award is part of a $40 million NSF funding initiative aimed at expanding the agency’s range of innovative computational resources that take advantage of rapidly changing technologies, complementing systems focused on capacity production computing, such as SDSC’s recently launched Expanse supercomputer.

[Continued on page 4]
I’d like to use this column to share some thoughts about why I feel SDSC is well-positioned during the next five to ten years to prosper as a leader in advanced computation, data science and engineering, and software services. I spoke about this at our General Staff e-Meeting in March, but here’s a recap for those of you who might have missed it.

The National Science Foundation’s recent document, ‘Transforming Science Through Cyberinfrastructure – NSF’s Blueprint for a National Cyberinfrastructure (CI) Ecosystem for Science and Engineering in the 21st Century’, signals a new direction by the agency to integrate all areas of science and engineering, research, and education, notably with innovations that lead to applications and sustainable services. The NSF recently shared its vision of this CI ecosystem, which has four key areas – all where SDSC has strong expertise and innovation:

- **Advanced CI Resources & Services**: What we provide to the greater research community with Comet and Expanse, plus our Voyager experimental testbed starting later this year.

- **Data & Software Services**: What we’re already providing with programs such as WIFIRE, CloudBank, EarthCube, and the Sustainable Scientific Software initiatives led by Mike Zentner.

- **Networking & Cybersecurity Services**: Sandeep Chandra is doing that with Sherlock, and CAIDA on the research side.

- **People, Organizations, & Communities**: Again, many examples: Science Gateway Community Institute, EarthCube Office, CIPRES, SeedMeLab, and more.

This new approach makes me feel very optimistic with regard to our national CI mission, since SDSC derives a major portion of its funding from such grants. Add to that SDSC’s industry funding led by Ron Hawkins and our renewed focus to explore other revenue-generating opportunities, and I have no doubt that SDSC will continue to thrive in the years to come.

Please join me in welcoming Cynthia Dillon as Executive Director of our External Relations group, succeeding Jan Zverina who will fully retire in May after 13 years with SDSC. Also, Tom Tate succeeds Jeff Filliez as our Data Center Manager. We’ll feature Tom in an upcoming Innovators staff profile.

My thanks to the entire SDSC staff for your continued high levels of dedication and resilience during the last 14 months. Please remain safe and healthy as we hopefully near the home stretch and return to more normal conditions before too long!

Michael L. Norman
SDSC Director
SDSC Summer & Fall Education Programs Gearing Up - Virtually

SDSC will be offering three education programs this summer and fall, covering students from middle school and high school who may be interested in considering careers in computer science and engineering.

SDSC’s Research Experience for High School students (REHS) celebrates its 12th year this summer. Founded by SDSC Education Manager Ange Mason, the eight-week begins June 21 and will reprise its successful virtual format from last year. Students are mentored by computational research scientists and other SDSC staff on a variety of topics and projects to gain exposure to career options and work-readiness skills through hands-on experience. At the end of the program, students will display their research via posters or short videos via Zoom.

Learn more at https://qrgo.page.link/GLVVX

New for 2021 is a pilot program at six area middle schools. Called ForMIDABLE and based on SDSC’s popular ABLE program for high school students, the goal is to keep younger students interested in STEM (science, technology, engineering, and math) topics prior to high school by empowering them with the resources to better understand computer science and engineering before choosing their high school areas of study. The program will be held via Zoom from June 26-August 21. On Thursdays, students will learn engineering software, and on Saturdays students will be introduced to robotics programming and different engineering fields through workshops, speakers, and demos. Registration opened April 21.

Learn more at https://qrgo.page.link/Wgn3G

The Mentor Assistance Program (MAP) program opens its application process on April 14. Now in its seventh year and to begin in Fall 2021, MAP is a campus-wide program which engages high school students in a mentoring relationship with an expert from a broad array of science disciplines. MAP’s mission is to provide a pathway for student researchers to gain access to UC San Diego faculty, post-doctoral fellows, and staff to mentor students in a field of their own interest. Students may communicate with a mentor via email or indicate an interest in visiting the mentor’s lab or work environment.

Learn more at https://qrgo.page.link/uvfcV
What and who inspired you to go into the geosciences?

Like a lot of people, I was inspired to go into geosciences by my natural curiosity of wanting to know how things work, as well as my love for the outdoors. Geosciences was a great fit for this type of curiosity because of the many disciplines that are involved along with the complexity and scale of the earth’s systems.

Can you describe for us the TZVOLCANO project?

We recently implemented real-time Global Navigation Satellite System sensors on the East African rift called Ol Doinyo Lengai that allow us to collect one data point per second. TZVOLCANO is the accompanying data broker that we created to collect and share the sensor information with the broader scientific community.

What do you think of the current state of diversity in the geosciences profession?

I would say the current state of diversity in the geosciences profession is not good. In many programs there is a clear lack of representation and this goes back to the founding of these programs. The good news is that there seems to be a desire to change this and the visibility of Black scientists and their achievements have helped incredibly.

What does Black History Month mean to you?

To me, Black History Month is a specific time when I devote an extra amount of time to learning more about the history of my family, as well as celebrate what previous Black heroes have done to overcome terrible situations and active resistance to equal rights.

Read the full interview at https://qr.go.page.link/U1yjJ

Voyager Taking Shape for its 2021 Debut

[Continued from page 1]

“With innovative solutions optimized for deep learning operations and AI workloads, Habana accelerators are excellent choices to power Voyager’s forthcoming AI research,” said Amit Majumdar, principal investigator who also leads SDSC’s Data Enabled Scientific Computing Division. Technology partner Supermicro will provide the Habana-based AI systems for Voyager.

Majumdar is working with co-PIs who include SDSC’s Mai Nguyen and Robert Sinkovits, as well as Computational Chemist and Biophysicist Rommie Amaro from the Department of Chemistry and Biochemistry, and Physicist Javier Duarte from the Department of Physics in UC San Diego’s Division of Physical Sciences. SDSC Deputy Director Shawn Strande is the project manager.

The NSF award is structured as a three-year “test bed” phase, followed by a two-year phase in which the system is expected to be more widely available using an NSF-approved allocation process.

Read more at https://qr.go.page.link/g53ky and https://qr.go.page.link/9ZzCw
Advancing Discovery Through Innovation

SDSC’s Comet supercomputer continues to serve as a vital research resource across a wide variety of science domains. Below are some recent highlights:

Illustrating the Mechanical Process of Cancer Growth
Stanford University and Purdue University researchers recently used Comet to run simulations showing how cancer cells divide in primary and metastatic tumors. The findings could lead to new therapies to treat tumors, according to the researchers. Prior to this research, there had been no study of how cells physically divide in collagen gels.

Better Understanding Alzheimer’s
"UCLA Professor Daniel Tward and his collaborators at Johns Hopkins University used Comet and other supercomputing resources to collect and share quantitative information from human brain images, including subtle changes in shape and cortical thickness and how it affects the early stages of Alzheimer’s disease. That cortex has been believed to be the first area impacted by Alzheimer’s disease, which until now could not be diagnosed until autopsy results were available.

Advancing Computational Chemistry Using New Artificial Intelligence (AI) Methods
Even though computational chemistry represents a challenging arena for machine learning, researchers from the Massachusetts Institute of Technology have made it easier by developing an artificial intelligence approach to detect electron correlation, or the interaction between a system’s electrons. While that correlation is vital, it’s also expensive to calculate in quantum chemistry. Using Comet and Bridges at the Pittsburgh Supercomputing Center, their work makes electron correlation detection more tractable while improving the number of materials that can be analyzed.

SeedMeLab and nanoHUB Win R&D 100 Awards
In late 2020 SDSC’s SeedMeLab, and a joint effort led by Gerhard Klimeck at Purdue University called nanoHUB, won awards in the Software/Services category of R&D World magazine’s R&D 100 Awards. Congratulations to SDSC Visualization Group Lead Amit Chourasia (SeedMeLab) and Michael Zentner, SDSC Division Director of Sustainable Scientific Software group (nanoHUB).

SeedMeLab is a cloud service to access, manage, and share data with context, while promoting productivity via discussion and visualization. The publication called it “an innovative way to grapple with rising data growth where new tools are essential to harness and share insights from them.”

Started in 2002, nanoHUB has grown from a small group of early adopters to about 18,000 users who annually perform more than 1 million simulations using ~600 online apps driven by scientific software, and nearly 2 million unique visitors accessing nanoHUB for more than 6,000 other resources.

Michael Zentner, who joined SDSC in 2019 following nine years with Purdue, brought the technical operations of the nanoHUB.org site to SDSC as part of moving the HUBzero infrastructure from Purdue to SDSC. R&D World said nanoHUB "accelerates innovation by removing traditional barriers... allowing software developers to easily transform their products into apps that run in the cloud with a simple interface presented through a web browser."

Both SeedMeLab and HUBZero are available as open source software and ready-to-use paid services from SDSC.

Read more at https://seedmelab.org and https://nanohub.org
Meet Cynthia Dillon
SDSC’s New External Relations Director
By Jan Zverina

Cynthia Dillon joined SDSC as the Center’s new Executive Director of External Relations, effective February 1. As a Director of Communications with UC San Diego’s University Communications since mid-2015, Cynthia helped raise the profile of academic divisions, including physical sciences, with her skills in strategic communications, public and media relations, writing and editing, and branding and marketing. Cynthia has a B.A. in journalism from Sacramento State University and an M.A. in anthropology from San Diego State University. She succeeds Jan Zverina, who plans to remain with SDSC on a part-time basis until mid-May to assist in the leadership transition. Jan interviewed Cynthia for this issue of Innovators.

Q: First of all, welcome to SDSC! While it’s only been a short while so far, can you describe how your previous roles at UC San Diego have been beneficial to your new position?

Thank you! It’s nice to join the team and contribute. Having had a career in nonprofit leadership, management, and teaching before joining UC San Diego, my roles on campus introduced me to the university’s system of leadership and administration. They also put me in touch with remarkable thought-leaders and cutting-edge researchers whose work was exciting to learn about, enjoyable to write about, and satisfying to share across media channels. My role at SDSC is a blend of my professional strengths grown from experiences both off and on campus.

Q: What attracted you to UC San Diego to begin with?

I always enjoyed school—especially my academic experiences as an undergraduate and as a graduate student. These experiences ultimately led to my work as a museum and library director. Additionally, having taught anthropology part-time in the community college district, I knew I would appreciate working full time at the university.
Q: Your educational background has an emphasis on both journalism and anthropology. How has that influenced your career choices?
Both journalism and anthropology involve a lot of writing—something I love to do. They also are both about documenting the human story. Part of having a skill in the business of strategic storytelling is being able to see the various pathways, how they intersect, and where they might lead. As Robert Frost said in The Road Not Taken, “way leads onto way.” Looking back at my career to this point, it makes sense that I am right here right now.

Q: Joining SDSC obviously comes at a time when almost all of us are still working remotely. Did you find that to be more of a challenge in some ways? Initially I thought the remote situation might be a disadvantage as a newcomer, but in actuality it’s fine. I do not feel disconnected, and I attribute that to everyone’s warm welcome and willingness to meet on Zoom or to take a phone call. That said, I do look forward to seeing people in person—many for the first time—when the time is right for it.

Q: SDSC researchers cover a wide range of science domains. Then there are the researchers across campus who rely on SDSC’s resources and computational expertise. Beyond that, many researchers at universities around the world use our computational and data resources. How do you see External Relations best serving all those areas?
Our team outreach extends from individuals at SDSC to our campus collaborators and their affiliates; to UC-system collaborators and those at other universities, centers, and labs; to our partners at the NSF and other supporting institutions; to our industry partners; our program participants; our web services clients; the press; and our social media followers. Any human relationship is developed over time and through effective communication. It is our job to make those connections in the ways we are uniquely positioned to do so as communications and multimedia professionals. SDSC has a strong ER team and tools, thanks in large part to your leadership tenure.

Q: One final question we always ask: What are some of your interests outside of work?
Outside of work, I serve as a board member on a couple of civic organizations. I live near a canyon, so I hike a lot. I enjoy spending time with family and friends—these days with CDC guidelines strongly in place. I also enjoy binge-watching a good series and reading about science.

ALSO WELCOME...

Please also welcome Tom Tate, who succeeds Jeff Filliez as SDSC Data Center Manager. Tom joins us from a commercial colocation facility in Orange County called Ntirety, where he worked for 10 years, first as an operations engineer and the last five as the Data Center Manager. Tom has hit the ground running and is eager and excited to bring his expertise to SDSC. “We’re thrilled to welcome Tom to SDSC,” said Brian Balderston, Director of Infrastructure for SDSC’s Research Data Services Division. “His expertise will be a great addition to our team.” Tom will be featured in a future Innovators staff profile.

STAY CONNECTED

Check out SDSC’s Training Videos, Social Media Feeds, and More!
Did you know that SDSC has a YouTube channel? Just click on the YouTube icon at the bottom of the SDSC home page under the ‘Get Connected’ section or go to...

https://www.youtube.com/SanDiegoSupercomputerCenter

There you’ll find a wide variety of topical videos, with playlists for Events, Training, Grand Challenges, and more.

twitter.com/SDSC_UCSD
instagram.com/SDSC_UCSD
facebook.com/SanDiegoSupercomputerCenter
youtube.com/SanDiegoSupercomputerCenter
linkedin.com/company/san-diego-supercomputer-center
Governments are taking advantage of advances in scientific instrumentation, sensors, and computational and storage technology to amass and analyze large amounts of data ("population-scale" data) about their citizens with the objective of better understanding and improving human health. One such initiative is the United Kingdom’s "UK Biobank" project, a large-scale biomedical database and research resource containing in-depth genetic and health information from half-a-million UK participants.

The database is regularly augmented with additional data and is globally accessible to approved researchers undertaking vital research into the most common and life-threatening diseases. It is a major contributor to the advancement of modern medicine and treatment and has enabled several scientific discoveries that improve human health. UK Biobank comprises a wide range of data including whole-exome and whole-genome sequencing on all participants, imaging data, biochemistry markers, phenotypic information, etc. The data is made available to bona fide researchers worldwide through an application process.

SDSC is working with an industrial research partner to support studies of UK Biobank data, leveraging SDSC’s advanced cyberinfrastructure, including the new Expanse supercomputer and SDSC’s high-performance Universal Scale Storage (USS) platform. The USS storage, based on technology from Qumulo, has been expanded to over 3 petabytes to support staging of data sets for analysis. Researchers will use Expanse, supplied by Dell and comprising over 100,000 CPU and GPU processor cores, to conduct genomics and other analyses of interest. SDSC’s computational scientists and bioinformatics experts will provide assistance in optimizing computational and storage workflows to accomplish the studies. SDSC expects the knowledge gained from this effort will benefit other population-scale health initiatives in the years to come.

To learn more about this initiative or how SDSC can work with you to support your computational research, please visit industry.sdsc.edu.

Ron Hawkins
SDSC Director of Industry Relations