Our Local Impact
Top 10 REASONS why SDSC is of value to UC San Diego/San Diego

SDSC provides unique value to its home base of UC San Diego and its neighbors in the greater San Diego area. For the campus, SDSC is helping to assure UC San Diego's place among the great research universities in the world, on the cusp of the next great era of intellectual discovery: data-enabled science and engineering. For the local community, SDSC is working to establish research, education, and training partnerships with those needing the knowledge and skills to tackle problems posed by this new era, while providing expertise and resources to help protect the region's health and welfare.

As SDSC approaches its third decade, it does so as a leader in data cyberinfrastructure and technologies, and as a strong, collaborative institution with a clear and strategic focus to help solve the fundamental problems facing science and society. As such, SDSC provides value to UC San Diego and the greater San Diego region in the following 10 ways:

1. **Extensive research partnerships and collaborations**
   SDSC researchers collaborate with a large number of UC San Diego faculty and staff researchers on numerous projects led through SDSC and other campus units. Some 44 unique UC San Diego researchers in 23 UC San Diego departments/units submitted at least one extramural grant with SDSC's PIs over the past six years. Conversely, SDSC PIs collaborated on more than 50 extramural awards with a total exceeding $300 million to 21 other UC San Diego departments.


Geisel Library, UC San Diego

2. **A magnet for faculty recruitment**
   SDSC has helped UC San Diego attract some of the best and brightest faculty, who have partnered with SDSC to bring large-scale, competitive awards to the university. Recruitments have included Phil Bourne, who brought the Protein Data Bank to UC San Diego; J. Andrew McCammon, holder of the Joseph E. Meyer Chair of Theoretical Chemistry; Michael Norman, a distinguished professor of physics at UC San Diego, a globally recognized astrophysicist and SDSC director; and Shankar Subramaniam, developer of the Biology Workbench and partner in the UCSD/Nature Signaling Gateway, among others.

3. ## Research cyberinfrastructure (RCI)

SDSC is a partner with several UC San Diego entities to offer research cyberinfrastructure to UC San Diego faculty and staff. The core elements consist of: co-location facilities housed in SDSC's data center; the SDSC Cloud, believed to be the largest academic cloud storage facility anywhere; digital curation and data services, provided by UCSD Libraries; the *Triton Shared Computing Cluster* (*TSCC*), serving researchers at UC San Diego and any of the other UC campuses as well as external academic, non-profit, and corporate users; and a research network, provided by the Administrative Computing and Telecommunications. (ACT).

4. ## Pioneering design for high-performance computing

SDSC houses *Gordon*, the first data-intensive supercomputer anywhere to use large amounts of flash-based memory—making it "the largest thumbdrive in the world." The result of a five-year $20 million grant from the NSF, *Gordon* has 300 trillion bytes of flash memory and 64 I/O nodes, making it ideal for data mining and exploration. SDSC also has a *Comet* sighting, thanks to a recent $12 million NSF award to launch in 2015 this new petascale supercomputer targeted to researchers in what's called the "long tail of science", which makes up about 99% of HPC users (see page 42).

5. ## Technology innovation

SDSC researchers have pioneered several significant software packages including ROCKS, which provides a blueprint for the construction of cluster computers for multiple laboratories around the globe; the Storage Research Broker, serving as "middleware" to hold together data cache sites for Big Data projects; iRODS, the open-source Integrated Rule-Oriented Data System, which represented a dramatic new approach to digital data management; and KEPLER, a scientific workflow automation system.

San Diego skyline. Image credit Tim McNew.

6. ## The "Go To" place for "Big Data"

SDSC is at the vanguard of the emerging 4th methodology of science known as data-enabled science and engineering, which seeks to apply massive data sets—commonly referred to as Big Data—to scientific discovery. Though the *Gordon* project puts SDSC at the vanguard, the Center's resources and expertise extend beyond the deployment and operation of this system for national users. SDSC's Big Data portfolio includes application team development, software development, user training and outreach, scientific collaborations, high-speed 100Gpbs (Gigabits per second) networking, training for today's academic and commercial researchers, and education for tomorrow's leaders.

7. ## Network to the nation

Since opening its doors nearly three decades ago, SDSC has conceived, nurtured, and raised multiple national partnerships and collaborations with individuals and communities across a wide spectrum of disciplines and fields of study. Today, SDSC is the only supercomputer center in the Western part of the nation participating in the National Science Foundation's Extreme Science and Engineering Discovery Environment (XSEDE) as a service provider for advanced cyberinfrastructure services for the U.S. open research community. SDSC also provided the lynchpin for UC San Diego's partnership with the Open Science Grid (OSG), a multi-disciplinary consortium funded by the U.S. Department of Energy and the NSF.

8. ## Education for the next generation

SDSC is a founding partner in UC San Diego's Computational Science, Math and Engineering (CSME) program and serves as a "real world" training ground for more than 80 students each year, many of whom have taken advanced cyberinfrastructure skills to the private sector or to traditional academic environs. SDSC also invites students and teachers from the Greater San Diego region to experience computing at its highest levels during summer programs. In addition, the Center organizes a unique volunteer internship program for undergraduate students designed to provide valuable "on the job" experience in a cross-section of projects in computational and data research. Under a grant from The Parker Foundation, in partnership with select local middle schools and CONNECT's Entrepreneurs for Young Innovators program, SDSC is working to provide a new series of computer science workshops aimed at getting minority and female students involved in computing (see sidebar, page 35).

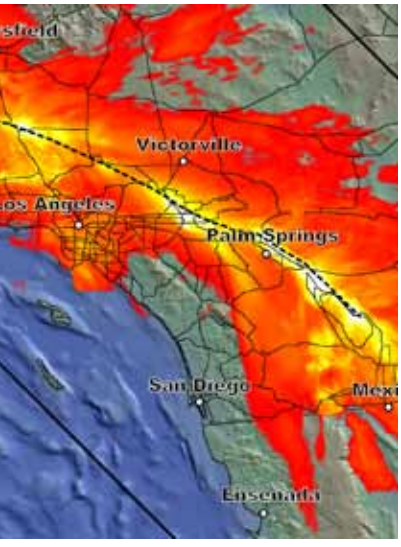

Image from a simulation of a magnitude 8 earthquake along the San Andreas Fault. Image credit: Amit Chourasia., SDSC

9. Outreach to the corporate community

SDSC has a long history of collaborating with and delivering value to industry, leveraging science and technology to deliver solutions for real-world problems. Among other programs, SDSC created the Predictive Analytics Center of Excellence (PACE) and its "boot camps" to accelerate research and education in predictive analytics for the academic and corporate communities. The Center for Large-scale Data Systems Research (CLDS) provides an opportunity for industry to collaborate with SDSC researchers on a large range of issues and challenges facing information-intensive organizations in this Big Data era. Finally, SDSC's Industrial Partners Program (IPP) provides member companies with a framework for interacting with Center researchers and staff, exchanging information, receiving education and training, and developing collaborations.

10. Cyberinfrastructure to save lives and property from the "Big One"

Seismologists at SDSC, San Diego State University (SDSU), and the Southern California Earthquake Center (SCEC) at the University of Southern California (USC) recently created the largest-ever simulation of a Magnitude 8.0 M8) earthquake, along primarily the southern section of the San Andreas Fault. About 25 million people reside in that area, which extends as far south as Yuma, Arizona, and Ensenada, Mexico, and runs up through southern California as far north as Fresno. SDSC provided the high-performance computing and scientific visualization for the simulation. The research was selected as a finalist for the Gordon Bell prize, awarded annually for outstanding scientific achievement in high-performance computing applications. The work represents a major breakthrough in seismology both in terms of computational size and scalability. It also opens up new territory for earthquake science and engineering with the goal of reducing the potential for loss of life and property.

DEVELOPING A "TRULY INTELLIGENT" SMART GRID TO MANAGE ENERGY CONSUMPTION

Natasha Balac is director of SDSC's Predictive Analytics Center of Excellence (PACE). Balac has developed a "sustainable communities" infrastructure proposal for downtown San Diego, in part to reduce power consumption.

Traditionally, electrical utility grids have been designed as a one-way street: to deliver power to the consumer. But with an ever-increasing demand to reduce energy costs and increase efficiencies, utility companies are now looking toward two-way street "smart grids" that not only deliver power, but also send back valuable information about how that power is being used to better manage these resources.

Toward that end, researchers with SDSC's Predictive Analytics Center of Excellence (PACE) and UC San Diego are leveraging traditional meter-based power information, as well as weather, locale, time of day and other potentially significant factors, with predictive analytic tools to develop a "truly intelligent" smart grid for the UC San Diego's 1200-acre campus.

The results are expected to improve operating efficiency, lower costs, and reduce the campus' overall carbon footprint generated by its 45MW peak load capacity.

"Achieving a truly smart energy infrastructure—for energy generation, distribution, and consumption—requires basic and advanced computing research and application," said Natasha Balac, Director of SDSC's PACE.

"The bidirectional flow of both electricity and information in the smart grid—consumption patterns in relation to the production, demand, environmental indicators and cost of energy—is extremely valuable when coupled with data analytics approaches," she added.

A significant amount of campus building and energy data is collected from UC San Diego's smart grid. Data is currently collected from 84,000 distinct, independent data streams (with continuous measures at one minute intervals) from approximately 30 different campus buildings. Data streams, including real-time measurements and set points, are collected from various building management and control systems including HVAC systems, the central utility plant, electric power meters, photovoltaic panels, network model output data, weather stations, and even plug-in electric vehicles. This large, complex dynamic data set holds enormous potential for significant energy savings.

"UCSD's smart grid operators, as many others, have experienced that minimizing power consumptions does not necessarily reduce overall energy costs," said Balac. "To reduce the performance costs we are developing time series models that may be used to develop forward-thinking management policies. Integration of smart grid control, optimization and scheduling dramatically improves the controllers' ability to optimize indigenous resources, import energy, export surpluses and shed loads in a more optimal manner."

"The fundamental goal of this research is to significantly lower energy costs by applying predictive analytics algorithms on the existing master-controller-optimizer," she added.

A key part of the project is a data-mining technique called MineTool-TS (MineTool for Time Series data), enhanced by UC San Diego and SDSC researchers to capture time-lapse information for mining smart grid data.

Also participating in the project are Tamara Sipes and Homa Karimabadi, UCSD Jacobs School of Engineering; Nicole Wolter, Kenneth Nunes, and Robert Sinkovits, all from SDSC.

EXTENDING COMPUTER SCIENCE TO MINORITY AND FEMALE STUDENTS

Diane Baxter is associate director of education at SDSC, with a focus on introducing computational sciences and computational thinking skills to students and teachers in regional, national, and international settings.

SDSC, in a partnership with selected local middle schools and CONNECT's Entrepreneurs for Young Innovators program through a grant from the Parker Foundation, began a series of computer science workshops in 2013 aimed at getting more minority and female students involved in computing.

The workshops were designed to encourage middle school students to take a new Computer Science Principles course when they matriculate to high school. Participating students were from Granger Junior High School and National City Middle School in the Sweetwater Union High School District.

"San Diego's economy depends upon technology and innovation that impact nearly every job sector," said Diane Baxter, SDSC's associate director of education. "These workshops are all about getting more students, especially underrepresented ones, to learn the computational thinking skills that those jobs require. But most of all, we want them to engage in the fun and excitement of computing so they look forward to learning more as they continue their studies."

The middle school program, funded by a consortium of local start-up industries represented by CONNECT, dovetails with a larger program funded by the National Science Foundation (NSF) called ComPASS, for Computing Principles for All Students' Success. That program is focused on training teachers to instruct students in Computer Science Principles.

"Local industries clearly understand the need to engage the full diversity of the region's talent in their workforce," said Karen Winston, vice president of workforce development and STEM initiatives for CONNECT. "They want to be sure that all students are getting the preparation they need to be part of a technology-based, innovation-driven economy in the San Diego region."