UC San Diego News Center

April 17, 2012 | By Jan Zverina

SDSC Announces 'Center of Excellence' for Predictive Analytics

PACE to Bridge Gap Between Academia, Industry, and Government

The San Diego Supercomputer Center (SDSC) at the University of California, San Diego is launching a new "center of excellence" aimed at leveraging SDSC's data-intensive expertise and resources to help create the next generation of data researchers by leading a collaborative, nationwide education and training effort among academia, industry, and government.

SDSC is providing seed funding for the program, called PACE for Predictive Analytics Center of Excellence. The program's goal is to develop and deploy a comprehensive suite of integrated, sustainable, and secure cyberinfrastructure (CI) services to accelerate research and education in predictive analytics – or the process of using a variety of statistical techniques from modeling, data mining, and game theory to analyze current and historical facts to make predictions, as well as assess risks and opportunities, about future events. Predictive analytics are now being used in a wide variety of fields such as healthcare, pharmaceuticals, financial services, insurance, and telecommunications.

While education and training initiatives are a large part of PACE, the project will also be open to collaborations and projects with industry and government, especially using _Gordon_, a unique, data-intensive supercomputer recently introduced by SDSC and which currently ranks among the 50 fastest supercomputers in the world.

"PACE will be one of several 'centers of excellence' at SDSC that demonstrate the center's expertise and resources in all aspects of big data," said SDSC Director Michael Norman. "We believe that data-enabled science is the beginning of a new scientific era, and we are ready to help academia, industry, and government make significant advances and discoveries in the area of data-intensive research."

Other SDSC centers of excellence include the Center for Large-scale Data Systems Research (CLDS) and the Cooperative Association for Internet Data Analysis (CAIDA), and more will be established as SDSC identifies emerging opportunities to apply its expertise to all areas of big data.

"Big data" is often used by researchers and academics to describe extremely large datasets that are part of the exponential increase in digitally-based information being generated daily by science and society. Many of those datasets that are gathered, stored, or analyzed are so voluminous that most conventional computers and software cannot effectively process them. Big data challenges are pervasive throughout genomics, biological and environmental research, astrophysics, Internet research, and business informatics, just to name a few.

SDSC, which in addition to *Gordon* has a large-capacity multi-tiered data storage system, has been positioning itself as a leading resource in big data management, specifically in the areas of performance modeling, data mining and integration, software development, workflow automation, and more.

Federal Focus on Big Data

SDSC's PACE program comes as the administration's Office of Science and Technology Policy (OSTP) last month announced $200 million in funding for new investments in big data research and development projects with the announcement of its Big Data (PDF) initiative. With support of the National Science Foundation (NSF), and to achieve their goals of leveraging data-intensive tools to aid in the country's research, defense and economic programs, the White House and the OSTP are bringing together six federal agencies or departments, including Homeland Security, the Department of Defense (DoD), the Department of Energy (DOE), the National Institutes of Health (NIH), the Food and Drug Administration (FDA), and the U.S. Geological Survey (USGS).

The key goals of the inter-agency initiative, according to OSTP officials, are to:

- Advance state-of-the-art core technologies needed to collect, store, preserve, manage, analyze and share huge quantities of data.
- Harness these technologies to accelerate the pace of discovery in science and engineering, strengthen our national security and transform teaching and learning.
- Expand the workforce needed to develop and use big data technologies.

"PACE is just one way that UC San Diego is responding to new funding opportunities coming under the government's big data research and development initiative," said UC San Diego Vice Chancellor for Research Sandra A. Brown. "SDSC is to be congratulated for its foresight and leadership in this area."

"As a non-profit public educational organization, PACE will focus on the administration's goal of educating and expanding the human resources needed for big data and predictive analytics," added Natasha Balac, PACE's director and director of data applications and service for SDSC's Cyberinfrastructure Research, Education and Development (CI-RED) group. "By doing so, we will help bridge the gaps between academia, industry, and government organizations by actively pursuing and involving individuals and entities from all three segments."

Guided by industry representatives, PACE will lead collaborative and coordinated nationwide education and training efforts to build a competitive workforce in data management and analysis, in part by developing and promoting a new, multi-level curriculum to involve all individuals in the field of predictive analytics.

In addition to developing standards and methodologies, PACE will serve as a hub for data mining and predictive analytics, while using SDSC's *Gordon* supercomputer and other resources to develop and implement novel, high-performance and scalable data mining tools and techniques. The program is also offering its service as a data mining repository for large datasets.

Other SDSC members of the PACE program include Programmer/Analysts Jo Frabetti and Nicole Wolter, and Research Analyst Paul Rodriguez.

MEDIA CONTACT

Warren R. Froelich, 858 822-3622, froelich@sdsc.edu

UC San Diego's Studio Ten 300 offers radio and television connections for media interviews with our faculty, which can be coordinated via studio@ucsd.edu. To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit https://ucsdnews.ucsd.edu/media-resources/faculty-experts.