UC SANTA BARBARA

THE Current

November 18, 2009
Margaret Connors

UCSB's NCEAS is Partner in DataONE, a New Global Data Access Network

Researchers at the University of California, including UC Santa Barbara's National Center for Ecological Analysis and Synthesis (NCEAS), have partnered with dozens of other universities and agencies to create DataONE (http://dataone.org), a global data access and preservation network for earth and environmental scientists that will support breakthroughs in environmental research.

DataONE (Data Observation Network for Earth) is one of two $20 million awards made this year as part of the National Science Foundation's (NSF) DataNet program. Universities and government agencies coalesced to address the mounting need for organizing and serving up vast amounts of highly diverse and inter-related but often incompatible scientific data.

Resulting studies will range from research that illuminates fundamental environmental processes to the identification of environmental problems and potential solutions.

NCEAS, the Department of Computer Science and Genome Center at UC Davis, and the California Digital Library at the UC Office of the President are integrally involved in the NSF DataONE initiative.

Across these UC partners, the award will drive advanced research and data acquisition, storage, mining, integration, and visualization for DataONE. The

resulting computing and processing "cyberinfrastructure" will be made permanently available for use by the broader UC community and international science communities. DataONE is led by the University of New Mexico, and includes additional partner organizations across the United States, as well as from Europe, Africa, South America, Asia, and Australia.

"Scientists have spent hundreds of years collecting environmental data – measuring temperature, counting fish and butterflies," says Stephanie Hampton, deputy director of NCEAS. "We already know quite a lot, when you estimate the volume of scientific data that must exist out there, but the challenge is to find those data sets and then put them together in a manner that helps to address the important questions for science and society. DataONE will be that portal for environmental data."

The DataONE team will study how a vast digital data network can provide secure and permanent access into the future, and also encourage scientists to share their data. The team will help determine data and data citation standards and will create the tools for organizing, managing, and publishing data.

As one of five DataNet collaborations envisioned by the NSF (http://www.nsf.gov/pubs/2007/nsf07601/nsf07601.htm), DataONE will build a set of geographically distributed Coordinating Nodes that play an important role in facilitating all of the activities of the global network.

The initial three Coordinating Nodes will be at UCSB (housed at the Davidson Library), at the University of New Mexico, and at the University of Tennessee/Oak Ridge National Laboratory.

"Institutions have made extensive investments in infrastructure for managing data at their local institutions and in discipline-specific consortia, but these systems generally don't interoperate," says Matthew Jones, director of Informatics Research and Development at NCEAS. "DataONE will provide a critically needed interoperability layer that will allow scientists from diverse domains to collaborate on pressing environmental science challenges."

Scientific data integration and management also occupy computer science researchers who develop methods and tools that support all stages of the data life cycle. "Effective annotation and integration of data, and efficient management of data lineage information are hot research topics in the database and scientific

workflows communities," says Bertram Ludaescher, professor of computer science at UC Davis, whose team specializes in scientific workflow and data integration technologies, and storage and querying of data provenance.

Libraries have traditionally played a critical role in preserving and providing access to scholarly materials, and recently have begun to focus on the complex challenges associated with managing scientific data.

"Libraries don't have the capacity to address these challenges individually," says Patricia Cruse, director of the UC Curation Center at the California Digital Library. "We need to partner with researchers, information technologists, and domain specialists to address these complex problems."

DataONE includes experts from library, computer, and environmental sciences explicitly to bridge these worlds and to develop an infrastructure to serve science for many decades to come.

About the UC Curation Center and the California Digital Library

The UC Curation Center (UC3) of the California Digital Library (CDL) was established in 2009.

UC3 is a central preservation and curation service provider addressing the system-wide needs of the 10 campuses of the University of California, one of the pre-eminent public universities of the world.

The California Digital Library provides digital library development and support for the University of California libraries and the communities they serve.

For further information contact Patricia Cruse, director, UC Curation Center, at patricia.cruse@ucop.edu or (510) 987-9016.

About the National Center for Ecological Analysis and Synthesis

NCEAS was established in 1995. The organization has hosted more than 4,000 scientists from over 50 countries, and supported more than 430 collaborative projects in ecology and related fields. NCEAS scientists develop new techniques in informatics and apply general knowledge of ecological systems to specific issues, such as the loss of biotic diversity, global change, and sustainability of marine

ecosystems. NCEAS is among the top 1 percent of 38,000 institutions evaluated for scientific impact in environmental research. NCEAS is funded by the NSF, the State of California, the University of California and numerous other donors.

For further information, contact Stephanie Hampton, deputy director of NCEAS, at hampton@nceas.ucsb.edu or (805) 892-2505, or Matt Jones, director of Informatics Research and Development, at jones@nceas.ucsb.edu.

Related Links

NCEAS

DataONE

National Science Foundation

About UC Santa Barbara

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.