

THE *Current*

March 12, 2014

Julie Cohen

NCEAS Upgrades Widely Used Scientific Data Repository

With small labs, field stations and individual researchers collectively producing the majority of scientific data, the task of storing, sharing and finding the millions of smaller datasets requires a widely available, flexible and robust long-term data management solution. This is especially true now that the National Science Foundation (NSF) — and a growing number of scientific journals — require authors to openly store and share their research data.

In response, UC Santa Barbara's [National Center for Ecological Analysis and Synthesis \(NCEAS\)](#) has released a major upgrade to the KNB Data Repository (formerly the Knowledge Network for Biocomplexity). The upgrade improves access to and better supports the data management needs of ecological, environmental and earth science labs and individual researchers.

The repository stores data related to a diverse range of topics, from [Influenza A subtypes in wild birds](#) to [decadal scale changes in coral reefs in the United States Virgin Islands](#) to [60 years of plankton data from Lake Baikal](#). Thousands of individual researchers, dozens of field stations and even large research organizations, such as the [Partnership for Interdisciplinary Studies of Coastal Oceans \(PISCO\)](#) and the [Long Term Ecological Research \(LTER\) Network](#), use the KNB to collaborate with colleagues and preserve data for the benefit of science.

The major overhaul to the KNB improves data access by making the repository more responsive with an intuitive, multifaceted search interface that is exponentially faster than the previous version. Queries across the entire repository now take less than a second, which makes finding data and potential collaborators faster and easier than ever.

The upgrade enables researchers to assign digital object identifiers (DOIs) to their data, so their work can be cited easily in science journals and credited when other scientists use their data. Designating a DOI is as simple as clicking the publish button, which makes the dataset publicly available and registers the DOI. Researchers can choose to share their data with only a small group of collaborators before releasing the information publicly prior to publication.

The system also features a new user interface and an improved search function, which make the repository easier to use. “The new KNB interface is very impressive,” said Margaret O’Brien, a research scientist at the Marine Science Institute and information manager for the Santa Barbara Coastal LTER project. “Showing how many datasets are available is very useful, and the search filter showing how many datasets to expect will help users tailor their input.”

As one of the founding member nodes of the DataONE network, the KNB Data Repository contributes to the diverse collection of data within the network, ensuring reliable, distributed storage of valuable research data for decades to come. NCEAS researchers Matthew Jones and Mark Schildhauer created the repository in collaboration with the LTER Network.

The KNB is built on the [Metacat](#) data repository software system, which is open source and freely available for other research groups to use to deploy their own repositories and link them into the DataONE federation. The KNB Data Repository is available free of charge for researchers, and scientists have stored tens of thousands of datasets on the service.

The NSF originally funded the project in 1998. Funding for enhancements came from NSF, the Andrew W. Mellon Foundation and the Gordon and Betty Moore 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.