

NEWS

Initiative to Establish Research Data Alliance Moves Forward

PAGE 354

With the amount of scientific data being gathered growing exponentially—due to the increasing number of sensor networks, scientific instruments, computer models, and other tools—there are enormous scientific opportunities but also significant challenges, including how to effectively share research data within and across scientific communities. A 29 August symposium on global scientific data infrastructures, sponsored by the National Academy of Sciences' Board on Research Data and Information (BRDI), addressed this issue while focusing on a new initiative to establish a Research Data Alliance (RDA).

The alliance already has received about \$3 million in funding from two U.S. agencies—the National Science Foundation (NSF) and the National Institute of Standards and Technology—the Australian government, and the European Union to move forward with funding a handful of nongovernmental structures (NGS) to work on the initiative, according to Alan Blatecky, director of the NSF's Office of Cyberinfrastructure. The initiative—which he said would include a council, plenary, secretariat, NGSs, and working groups—is on a fast track, with the first NGS meeting planned for 2–3 October in Washington, D. C., and the first

meeting of the RDA scheduled for March 2013 in Europe.

At the symposium, Blatecky outlined some major hurdles to establishing an open global research infrastructure, including a general lack of understanding about the urgent need for it, and waiting for an all-encompassing solution to enable data sharing rather than moving forward sooner with good but perhaps not perfect measures.

What's needed, he said, is to establish the RDA to promote and enable the exchange of information and to “stop talking about data sharing and start sharing data.” He said that while some scientific communities already have a well-established capacity for data sharing, others lack that, and there is also the need for data sharing across scientific disciplines.

BRDI cochair Fran Berman, professor of computer science at the Rensselaer Polytechnic Institute, said RDA should not be thought of as something that replicates what is already working and effective for specific communities. “Think about it as an organization that facilitates working communities that need a little push in order to get something tangible that will help them make progress,” she said.

Kelvin Droegemeier, a BRDI member and vice president for research and chair emeritus of applied meteorology at the University of Oklahoma, noted that with so many

science communities now working with and producing huge amounts of data, RDA could serve as a facilitator for data sharing and other purposes; empower interdisciplinary scholarship; and help connect policies, practices, and standards with institutional policies at universities.

“The challenge is, How do we coordinate all the data repositories and data sources to where we can all collaborate using one another's data?” Droegemeier told *Eos*. “What's interesting is bringing in the social science aspects of research because ultimately the geophysical phenomena—like hurricanes, tornadoes, and earthquakes—affect people. We can't study these problems from a purely physical science point of view. We have to do it in an integrative way.”

“We're in the middle of a very rapid technological change in sophistication, and in the volume, and in the complexity of data,” said Robert Chen, director of the Center for International Earth Science Information Network at Columbia University. “We're flooded with data, and we're looking ahead to a flood of data.”

Chen said that institutional silos that make it difficult to share data need to be broken down by building digital bridges between the silos. He also noted other data challenges, including deciding what data need to be stored, questions of liability related to the collection of data, the possibility of data being misinterpreted by people unfamiliar with it, and privacy concerns. However, Chen stressed that “big data isn't just important for scientists. It's also important for decision makers, it's important for everyday life.”

For more information, see http://sites.nationalacademies.org/PGA/brdi/PGA_070715.

—RANDY SHOWSTACK, Staff Writer