

Grant Boosts UCSD Libraries' Storage and Management of Digital Assets

September 16, 2004

Paul K. Mueller

The science of data management, usually associated with corporate information, is of equal value to modern libraries, which are, after all, highly organized repositories of data - much of it now digital. That's why a recent award of nearly \$311,000 for advanced data management research at the San Diego Supercomputer Center (SDSC) and the libraries of the University of California, San Diego (UCSD) and the Massachusetts Institute of Technology (MIT) is welcome news for both librarians and university computer specialists.

The funding, provided by the National Archives and Records Administration (NARA), will help the UCSD Libraries integrate two different data-preservation architectures, and benefit from the key features of both.

Brian E.C. Schottlaender, UCSD's university librarian, said the project will provide the Libraries with a user-friendly system to collect, manage, store and disseminate its growing collection of digital objects. "We particularly look forward to the availability of a storage infrastructure, using SDSC-developed software, that can scale to accommodate millions of objects regardless of format," he said. "This flexible and expandable storage layer will provide the foundation for future interoperability of digital collections across institutional, national and even international boundaries."

According to Dawn Talbot, senior associate for the Digital Library Program at the UCSD Libraries, the award is part of an ongoing effort to develop prototype persistent archives, an effort both UCSD and MIT have now joined. "This project integrates two totally independent systems," she said, "DSpace, a user-friendly data-management system developed by MIT and Hewlett Packard, and the Storage Resource Broker (SRB) middleware developed at SDSC, which enables many different storage devices to work as one. By integrating the systems, we'll combine the user-friendly front end of DSpace with the 'middle ware' data-management and storage capabilities of the SRB."

Chris Frymann, digital library architect at the UCSD Libraries, says that the result of the work will benefit other libraries and educational institutions, as well.

"While retaining their paper-based storage and cataloging functions, libraries must increasingly also manage many kinds of digital content, including text, images, music, video and other file formats," said Frymann. "By combining DSpace software with the SDSC's SRB data-grid technology, we'll be able to store data in a much wider variety of physically diverse and remotely distributed storage devices. In the context of the traditional library, this is like magically reducing the cost, while increasing the capacity, accessibility and manageability of shelf space."

Even better, said MacKenzie Smith, associate director for technology for the MIT Libraries, both DSpace and SRB are available to educational institutions for use without charge. "DSpace is available to any type of organization as free, open-source software," she said. "And the standards developed by the project will be tested with large-scale storage systems from commercial providers, offering solutions to the full range of organizations which need digital asset-management for long-term storage and preservation."

The libraries' digital collections can be "virtually unlimited" in size, as well as be stored, replicated and made accessible through the resulting "grid technologies," said Talbot.

Reagan W. Moore at the San Diego Supercomputing Center (SDSC) said that MIT, SDSC and the UCSD Libraries will collaborate on the integration. "Some of our specific technical goals are to improve distributed data management support for DSpace processes; to specify policy models for managing archival contexts; to test and evaluate the integration using UCSD collections and, finally, to validate the integration against a NARA collection," he said. "The integration of DSpace technology with the SRB will be tested for use within both the digital library community and preservation community."

The funding for the project will be roughly split between MIT and UCSD. Moore and Richard J. Marciano of UCSD will lead the effort at the SDSC, while Schottlaender will oversee the UCSD Libraries' contributions. Smith will direct MIT's involvement.

Media Contact: Paul K. Mueller, (858) 534-8564