

March 2014 Progress Report

Background

Cyberinfrastructure is broadly defined as the human and technological support framework for advanced data acquisition, data storage, data management, data integration, data mining, data visualization, data curation and other computing and information processing services within the research environment.

The MU Cyberinfrastructure (CI) Council was formed in the realization that the University could not achieve its goal of retaining AAU membership without adequate research cyberinfrastructure. MU's existing research support computing resources, purchased through partnerships with research investigators and a FY04 Congressional Earmark, are insufficient.

MU's CI Council is comprised of faculty and staff representing each of MU Schools and Colleges. They provide input and guidance on the cyberinfrastructure needed to support the wide variety of research and discovery activities across the University of Missouri. Initial members were identified with input from Research Deans of each school/college in December 2012. The first meeting was held in January 2013, and was followed by a series of biweekly meetings. Currently, the Council meets about once a month and welcomes interested visitors. The input of CI members is critical for planning required strategic investments and ongoing support for cyberinfrastructure, as well as the services and resources provided by the Division of IT's Research Support Computing organization.

For a listing of current CI Council members, visit Cyberinfrastructure Council Membership List. If you have questions or concerns about MU's cyberinfrastructure or the CI Council, please email MUClsurvey@missouri.edu.

Progress Towards Goals

Campus Cyberinfrastructure Plan

- Began with careful review of CI Plans from NSF's Campus Cyberinfrastructure-Infrastructure, Innovation and Engineering (CC*IIE) program proposals provided by the University of North Carolina, University of Michigan, University of Chicago, Virginia Tech University, University of Wisconsin-Madison, Florida State University, University of Florida, Duke University, University of Colorado-Boulder, Wayne State University, and Texas A&M University.
- Subcommittees worked through the summer of 2013 to draft a cyberinfrastructure plan for MU.
- Numerous details were gathered to support the plan. See: CI Resources
- Gary Allen and CI Council members presented the plan to key constituents and groups in August, September, and October.
- A draft plan was presented and discussed at the inaugural CI Day on October 10, 2013.
- CI Council approved the plan December 2013. See CI Plan
- Continuing living plan to support relevant funding requests and new research developments (February 2014, a version was developed for a proposal to NSF's CC*IIE Program).

Communications

- April 2013 Survey – Created a survey in April 2013, to gather information on how faculty and graduate students conduct research, their data storage and processing needs, the limitations of MU's current cyberinfrastructure for meeting these needs, and their awareness of available resources. See "Survey," bottom of doit.missouri.edu/research/ci/.
- The planning process was a critical component of the group's communications.
- Web page for CI Council was created: doit.missouri.edu/research/ci/.
- CI Day was held October 10, 2013. Event archives can be found in MOspace: mospace.umsystem.edu/xmlui/handle/10355/39019.
- Focused on particular subsets of researchers as appropriate to convene a High Performance Computing (HPC) users group and conduct a focused survey of HPC users to better inform procurement decisions.
- Beginning with MU's first CI Day, MU's Information Experience (IE) Lab has taken a scientific approach to documenting and communicating to MU faculty, staff, and students in order to raise awareness of and increase comfort with research CI.
- Communications also focused on the criticality of and issues surrounding the security of research systems and data. Discussions to raise awareness of serious risks, problems encountered elsewhere, and the necessary actions were beneficial. During this year, the health sciences units in particular moved forward with establishing an infrastructure to segregate and secure patient data, and has been actively communicating and promoting the use of said environment, thus raising awareness of and compliance with information security requirements.
- Computer Science and Engineering and "domain" researchers helped articulate the need for two NSF proposals: A Major Research Instrumentation (MRI) proposal (Shyu, PI) and a CC*IIE proposal (Calyam, PI).
- In partnership with the DNA Core Facility, established MU's first Data Retention Policy. See: biotech.missouri.edu/dnacore/corepolicies.html.

Increase Funding for CI

- Campus administration expressed the intention to increase funding for research CI, including: \$500,000 for disk storage plus \$1 million in one-time funds, and increased rate funding of \$500,000 per year for four years to eventually add \$2 million per year of ongoing funding.
- The Campus CI Plan was the guiding force for the January MRI proposal (\$717,104).
- CI council outcomes directed MU's portion of the UM-System funding request for research computing (MU portion \$90,000 recurring and \$500,000 one-time).
- The Campus CI Plan was required for NSF CC*IIE proposal (Calyam, PI – \$399,775 for 2 years). A focused version from the approved plan and the Details Behind the Plan document was created specifically for this funding proposal.
- Improved the quality of data PIs provide about research computing when they submit a proposal.
- Federal funding rules are changing and more research computing costs will be allowable as a direct cost of the project rather than only through reimbursement of facilities and administration costs.

Hardware Acquisition and Improvement

- Surveys of faculty and graduate students, and High Performance Computing (HPC) focused surveys for subset of faculty were instrumental in intelligent selection of priority purchases.
- A replacement for the Clark system has been purchased and is being installed. The purchase of upgrades to the Lewis cluster is currently in progress.
- High Memory HPC for genomics studies is planned in consultation with these researchers.
- CI Council investigations and discussions indicate a variety of data storage solutions will be necessary and discoverability across the array of storage needs will be critical.

Improvements to Software

- The CI Council participated in an evaluation of Box cloud storage in Fall 2013. It was made available as a campus-wide service for students, faculty and staff at no charge in February 2014.
- Digital Humanities Commons, funded by Mizzou Advantage, can help identify and prioritize the digital humanities' needs. MU Libraries is leading discussions of possible Digital Asset Management system(s).

Increase Staff Support to Research Computing

- Created the new position of Director, Research Computing Support Services and hired Timothy Middelkoop.
- Committed full-time staff support to test bed project behind Major Research Instrumentation proposal to NSF.
- Staff support was a key component of MU's portion of UM-System Funding Request.
- Prasad Calyam's CC*IIE proposal was entirely for staff support in the form of a CI Engineer.

Master's Degree in Data Science and Analytics

- February 2014, Chi-Ren Shyu was the PI on a proposal that received \$80,000 by Mizzou Advantage to develop six courses and to develop online delivery for four existing courses.
- Proposal for the new degree program is in progress. It must first cross MU, and then UM-system hurdles, before submission to Missouri Coordinating Board for Higher Education (CBHE) for approval.

Addressing Research Data's Complex Challenges

The US government and federal funding agencies are making it increasingly clear that universities will bear the responsibility for assuring data generated by research projects are accessible to the public. Collaborations are emerging to help address this challenge. For example, on March 28, 2014, the Association of Research Libraries was awarded a joint \$1 million grant from the Institute of Museum and Library Services and the Alfred P. Sloan Foundation to develop and launch the Shared Access Research Ecosystem (SHARE) Notification Service. SHARE is a collaborative initiative of ARL, the Association of American Universities (AAU), and the Association of Public and Land-grant Universities to ensure the preservation of, access to, and reuse of research findings and reports. This is a challenge with daunting implications. However, MU has several advantages upon which to build, including:

- Existing collaborations of MU libraries and the Division of IT position.
- MU's tradition of investigator participation in funding research computing.
- Engagement of IT pros from MU's schools and colleges.
- Leadership position in the State and the Great Plains region.