

MU Cyberinfrastructure (CI) Council Progress Report for 2016 and 2017

Summary:

From its beginning nearly five years ago, the University of Missouri's CI Council has been guided by faculty with representation from all the schools and colleges. It is important to acknowledge the commitments of time, talent, vision, and perspective toward defining and articulating MU's research cyberinfrastructure priorities and direction. This group has created MU's first CI plan in 2013, and in 2016 refined the plan to call attention to four critical areas: Bioinformatics and Genomics, Geospatial Sciences, Digital Humanities, and Imaging & Visualization

Through the dedicated efforts of MU's Cyberinfrastructure (CI) Council significant progress has been made toward the vision established in MU's CI Plan Research Computing has been growing to meet the needs of researchers on campus with a number of improvements, new services, and refinements in priorities and direction. Faculty engagement has successfully inspired campus investment in HPC fine-tuned for the exploding demand for high throughput sequencing. A sustainable funding model is emerging, and thankfully campus funds have been directed toward research CI – notable in a time of draconian budgetary cuts. Over the years, MU's CI Council has also organized four campus-wide CI Days designed enhance communication and share plans, priorities and progress. Plans, progress reports and related documents are in the University's digital repository MOspace. Visit <https://mospace.umsystem.edu> and search for CI Council, or access directly at <https://mospace.umsystem.edu/xmlui/handle/10355/62761>

CI Council Responsibilities

1. Understand the demands for CI resources, and recommend priorities for purchase; or if feasible, enable access through a local, regional, or national source.
2. Provide leadership in properly securing research data.
3. Help ensure campus leaders are informed of the needs for CI and how other universities are meeting those needs; in particular, the needs for human resources, not just hardware and software.
4. Ensure MU has a defined CI architecture to explicitly determine what will (and will not be) supported.
5. Establish, refine, and communicate a business model that will provide the sustainable funding necessary for CI resources, likely some level of generally available no-charge resources and a higher level of for-fee services.
6. Facilitate communications, seeking input on research CI needs and priorities, and share CI Council information in MU's schools and colleges.
7. Oversee the allocation and use of high-performance computing and disk storage designated for special projects.
8. Demonstrate best practices and help facilitate and inspire effective collaborations.
9. Review progress toward MU's CI Plan and refine the plan as demand and situations dictate.

Monthly Meetings

MU's CI Council continued to meet for approximately ninety minutes each month – in E205A&B of the Locust Street Building (615 locust St). Attendance and participation has been stellar, and the discussions have been lively and informative. Over time, some membership of the CI Council has changed. For a current list, see: <https://doit.missouri.edu/services/research/ci/cyberinfrastructure-council-membership-list/>

Researcher Presentations

Starting in December 2014, the CI Council added a new item to its monthly meetings – discussions with MU researchers regarding their research and the related needs for cyberinfrastructure. Each researcher was asked to make a brief presentation designed to stimulate discussion. To maximize the consistency and value across presenters, each presenter is asked to dedicate one or more slides to these topics:

- 1) Brief overview of research interest, why innovative, future trajectory and relevant funding
- 2) Information on research collaborators and their resources relevant to your research
- 3) Describe Intellectual Property (IP) issues
- 4) Describe your CI ‘pain points’ and how they affect your research
- 5) If money were no object, what CI environment and resources would you have?

MU CI Council - Researcher Presentations & Discussions

November 2015 – April 2016 monthly meetings focused on briefings on the focus areas of the update to MU’s CI Plan. The various planning reports and documents can be found in MU’s digital repository, MOspace, at: https://mospace.umsystem.edu/xmlui/handle/10355/48293	
May 2016	Kannappan Palaniappan, Lapiere Associate Professor of Computer Science Briefing on regarding his research in video analytics and cloud computing
August 2016	Marjorie Skubic, Professor in the Electrical and Computer Engineering Department and Computer Science; and Director of MU’s Center for Eldercare and Rehabilitation Technology. Eldertech and the decade of research using sensors to monitor health status and improve health outcomes
September 2016	When the CI Council was formed in 2013, the group established the direction that we need to ‘grow our own’ data scientists to meet the expected demand. Mizzou Advantage funding helped advance the idea and Chi-Ren Shyu provided a report on the significant progress has been made toward that goal. The presentation included a briefing and lively conversation with Grant Scott and Sean Goggins who are co-directors of MU’s new Master’s Degree in Data Science and Analytics see: https://dsa.missouri.edu/ This new degree program is underway with 16 students enrolled this fall, and a planned total enrollment of 50–100 students. Please help spread the word of this new degree program.
October 2016	Dr. James D. Schiffbauer, Assistant Professor in Geological Sciences shared his innovative research in digital paleobiology. See: paleo.missouri.edu .
February 2017	Rainer Glaser, Professor of Chemistry Dr. Glaser shared information on his research, teaching, and service activities and the importance of the synergy between these facets of his work. Dr. Glaser is a theoretical organic chemist and his group frequently publishes theoretical results before any experimentation begins
March 2017	Researcher conversation with Dr. Lloyd Sumner, Professor of Biochemistry and Director of MU’s Metabolomics center UMMC). Dr. Sumner came to Mizzou from the Samuel Roberts Noble Foundation in 2016 as a signature hire. He shared his experience as a new hire and described his immediate and longer-term CI needs. He established a new U of Missouri Metabolomics Center (UMMC) and recently hired Dr. Zhentian Lei as Assistant Director. UMMC’s mission is to develop, translate and apply cutting-edge metabolomics technologies in a collaborative manner to study plants, microbes and animals.

May 2017	Dave Larsen, Professor of Forestry in MU’s School of Natural Resources. Dr. Larsen will be sharing his research endeavors and his perspectives on research cyberinfrastructure. Dr. Larsen was a member of the Research subcommittee of MU's IT Transition Project (<i>ITTP</i>), and has agreed to share his thoughts on that process and what has transpired since. The charge of the Research Subcommittee was to: “ identify best practices, unmet needs, and inefficient uses of research computing infrastructure and services. This subcommittee will establish standards as well as processes to ensure broad, sustainable engagement of the research community in an effort to provide necessary services and reduce duplication.” For more information see https://transition.missouri.edu/research/
August 2017	Preparations for CI Day scheduled for October 4, 2017.
September 2017	Topic: “Building and Supporting Science Gateways to Expand the Capabilities of Campus Research Computing” Dr. Marlon Pierce, Director of the Science Gateways Research Center, part of Indiana University’s Pervasive Technology Institute. Pierce is manager for the XSEDE Science Gateways Program, Co-Principal Investigator and Extended Developer Support services lead for the NSF-funded Science Gateways Community Institute, and Principal Investigator of the NSF-funded Science Gateways Platform as a service (SciGaP.org) project, which develops the open source Apache Airavata project. Pierce is a member of the Apache Software Foundation.

CI Day 2017

The University of Missouri’s fourth annual CI Day was held on October 4, 2017 from 8:30 am to 4 pm in MU’s Memorial Union. In addition to two excellent speakers, the day consisted of updates on progress toward MU’s cyberinfrastructure (CI) planning progress, and a presentation on cyber security by representatives of the Federal Bureau of Investigation (FBI). Dell Corporation provided funding for the refreshments and a box lunch. The keynote speakers were:

- Irene Qualters, Director of the Office of Advanced Cyberinfrastructure at the National Science Foundation
- Mark McIntosh who serves as Vice Chancellor and Vice President directing the MU Office of Research, Graduate Studies and Economic Development and the UM-System Office of Research and Economic Development

Eighty (80) attended the event. In addition to significant participation from the Division of IT and the Distributed IT professionals who work within MU’s Schools and Colleges, the event was attended by faculty and staff from: Arts and Science, Business, CAFNR, Education, Engineering, Journalism, Libraries, Nursing, Medicine, Vet Med, and a number of students.

Dr. McIntosh affirmed MU’s CI Plan in his CI Day presentation. He agreed with the need to make CI sustainable and that doing so will require university funding, student fees and external project funding. He was instrumental in a recent investment by MU campus administration that provides researchers exciting new HPC resources fine-tuned to support high throughput sequencing.

As envisioned in MU’s CI Plan, Research Computing Support Services, within the Division of IT, has designed and implemented an architecture that supports integration and is extensible, and research allocations can be based on the common good with benefits for many. Dr. McIntosh encouraged the CI Council to reach to faculty, students, and research staff to assure they have input in helping prioritize purchase decisions and ensure good stewardship of limited funding and resources.

Dr. McIntosh also affirmed the planning principle that researchers should work with funding agencies to budget for and include CI services as a direct cost. Through a shared “investor model“ researchers can easily assure they have the cyberinfrastructure their projects need when they need them, and when they don’t they are shared. Resource descriptions and budget Information and project planning assistance is available to help researchers understand and incorporate their CI needs into external funding proposals. Researchers are encouraged to contact Dr. Timothy Middelkoop, Director Research Computing Support Services at: middelkoopt@missouri.edu

For more about MU’s CI Council, see: <https://doit.missouri.edu/services/research/ci/>
 MU’s CI Plan is available at: <https://doit.missouri.edu/services/research/ci/cyberinfrastructure-plan/>
 The CI Day schedule, presentations, evaluation and related materials can be found in the University of Missouri’s digital repository, MOspace, at: <https://mospace.umsystem.edu/xmlui/handle/10355/62761>

Improvements to Services and Resources:

Significant improvements have been made each semester. The table below provides a summary and additional details follow. In addition, Dr. Timothy Middelkoop prepares an update each semester. **Those are included as Appendix A to this progress report**, or can be found in the News area of the research computing web site. See: <https://doit.missouri.edu/research/research-computing-news/>

<p>Summer 2016: Decommissioned Lewis2 (8 racks). Decommissioned 3Par/IBRIX (100TB in 2 racks). Recommissioned 16 racks, Upgraded provisioning infrastructure (Lewis 3). Added Jacob and Susie (Double the team size). Dr. Middelkoop convened a statewide organization focused on research CI. See: http://www.showmeci.org/</p>
<p>Fall 2016: Added teaching cluster. Upgrade configuration management infrastructure (Lewis 4 – October)</p>
<p>Spring 2017: Commissioned Secure4. Commissioned new HTC storage service with 300TB (Now 1PB). Setup HPC storage evaluation environment (3 racks). Add Biocompute partition. Migrate MRI RC GPU’s to Lewis. Migrate MRI RC partition to Lewis. Move 100B Gbps connection from Internet2 to the GPN. Upgrade over 80 software packages. Complete the Lewis upgrade.</p>
<p>Summer 2017: Commission General Purpose Research Network service leveraging RNet building switches. Co-developed UMKC researcher managed backup.</p>
<p>Fall 2017: Commission new Science Gateway environment (oVirt). Full utilization (~ 6 months from upgrade to 5000+ cores). Posted CI Engineer Position. CI Day October 4, 2017.</p>

Calendar Year 2016 Achievements and Activities

Throughout all of 2016, Lewis 3 Training was provided from 11am-noon on Wednesdays. This training covered the basics of the new hardware, how to use the new scheduler, how to use secure shell key-based authentication, and any other questions trainees had.

January 2016 Research Computing Support Services announced an academic site license for MobaXterm software (<http://mobaxterm.mobatek.net/>). There is no-charge for researchers and students on the University of Missouri Columbia campus for research and teaching purposes. MobaXterm is the preferred way to access the Lewis3 cluster from a Windows desktop. One can download and install the software by downloading the installer from <https://missouri.box.com/rcss-mobaxterm>.

On Wednesday, April 13, MU’s third annual CI Day was held at MU’s Memorial Union. The theme was *Collaboration: The Future of Research*. The keynote speaker was Henry Neeman Assistant Vice

President, Information Technology at Oklahoma University and Research Strategy Advisor, and the Director of the [OU Supercomputing Center for Education & Research](#) (OSCER). Henry also provided feedback and advice to the CI Council made up of faculty and IT professional representatives from each of MU's schools and colleges. The goal of CI Day was to raise awareness and assure that MU researchers can effectively compete on the national and international levels. Research Computing Support Services strives to promote a research environment that supports advanced data acquisition, storage, management, integration, mining, visualization, and other computing and information processing services, as well as the human resources necessary to make CI useful and effective. Faculty input is vital to our success in these efforts and participation in this event includes opportunities to discuss with peers how to grow the MU research computing community.

May, 2016 As suggested by the Henry Neeman, keynote speaker at the April 13, CI Day. Dr. Middelkoop convened a statewide group designed to share cyberinfrastructure, information, education, and resources across the Show Me State." This group has continued to meet regularly through 2016 and 2017. For recent progress and additional information, see: <http://www.showmeci.org/>

July 2016 two new staff members joined Research Computing Support Services:

- As a **Cyberinfrastructure Engineer**, Jacob Gotberg will help identify, prioritize, and optimize the computational resources at MU to assist a wide variety of researchers in achieving their computational goals. An important part of his job will be to implement Advanced Cyberinfrastructure Research and Educational Facilitation (ACI-REF) best practices. He will be traveling to conferences and training to help bring awareness and understanding of internal and external ACI resources to potential users of those resources.
- As a **Business Technology Analyst**, Susie Meloro will play a critical role in working closely with faculty and researchers to help identify needs and requirements for their research activities. She will also help them access cyberinfrastructure assets and services, including high-speed research networks and high-performance computing clusters.

July 17-21, 2016 XSEDE16 Conference Dr. Timothy Middelkoop, Director of Research Computing Support and Services, attended the fifth annual XSEDE conference which showcased the discoveries, innovations, challenges, and achievements of those who use and support XSEDE resources and services. The Extreme Science and Engineering Discovery Environment ([XSEDE](#)) is a five-year, NSF-funded project that provides the most advanced, powerful, and robust collection of integrated advanced digital resources and services in the world. It is a single virtual system that scientists can use to interactively share computing resources, data, and expertise. As a "Campus Champion", Dr. Middelkoop serves as a conduit between the XSEDE community and the University of Missouri. He facilitates educational and training opportunities for our researchers, faculty, students, and staff and helps them make effective use of local, regional, and national digital resources and services. The first five years of the XSEDE project has been so successful that the NSF has renewed funding for another five years. XSEDE 2.0 will offer even more services to the community that will provide the backbone to a national cyberinfrastructure for researchers, scientists, and engineers.

September 2016 Dr. Timothy Middelkoop and the Research Computing Support Services staff participated in Research & Technology Development Conference (RTD2016) Missouri University of Science & Technology in Rolla, MO on September 12 and 13. Research Computing Support Services announced the arrival of 18 new High Performance Computing nodes to be integrated into Lewis4. The HPC update to MU researchers detailed the schedule and expectations of the HPC enhancements which included upgrading and expanding two existing clusters, building 4 new specialized clusters, and decommissioning old equipment and infrastructure. The upgrade included the addition of more than 100

new nodes (more than 2800 cores and 34 TB of RAM!) and restructuring the networking, storage, and other infrastructure.

Oct. 21, 2016 The University of Missouri is one of the initial collaborating institutions participating in a Research Coordination Network project called, “RCN: Advancing Research and Education Through a National Network of Campus Research Computing Infrastructures – The CaRC Consortium”. The consortium was recently awarded a \$750,000 grant by the National Science Foundation to set up a national forum for the exchange and dissemination of best practices, expertise, and technologies to enable the advancement of campus-based research computing activities. Clemson University, who is leading the effort, issued this press release: <https://newsstand.clemson.edu/clemson-nsf-carc-consortium/>

November 13-18, 2016, the Research Computing Support Services team attended Supercomputing16 (SC16), the premier international conference for high-performance computing, networking, storage and analysis. This conference showcased the many ways these technologies lead to advances in scientific discovery, research, education, and commerce. SC2016 attracted more than 10,000 attendees from around the world, bringing together an exclusive group of scientists, researchers, lab directors, and funding program directors from the HPC and networking industry. As in previous years, MU shared booth space with the Great Plains Network.

Calendar Year 2017 Achievements and Activities

Throughout all of 2017, Training was provided from 11am-noon on Wednesdays. This training covered the basics of the new hardware, how to use the new scheduler, how to use secure shell key-based authentication, and any other questions trainees had. In August 2017 the training was expanded to include New User Training from 10-11 am in addition to Lewis Training from 11 - noon. New User training covers the basics of getting an account to running simple jobs on Lewis. Lewis training covers more in-depth topics, questions, and hands-on help with tuning jobs.

Beginning in **February 2017** Dr. Timothy Middelkoop began a series of traveling office hours meeting with researchers at UMKC, UMSL, and MU Council of Research Administrators and the deans and representatives of several of MU schools and colleges - Engineering Medicine, Law, and Nursing so far. CI Council representatives are engaged in meetings with the deans, and meetings with faculty groups are encouraged.

March 2017 The new High-Throughput Computing Storage (HTC) system is now available on the Lewis cluster. It is designed for researchers who need large amounts of long-term storage for high-throughput computation at a low cost. Storage is allocated in blocks and each storage block has its own quota mount point. Storage blocks are prepaid in full in 10TB increments for 5 years. The FY17 rate is \$1,200/10TB/60 Months. MU’s CI Council helped develop and approve the relevant policies:

- Storage Policy: <http://docs.rnet.missouri.edu/Policy/storage>
- HTC Storage Policy: <http://docs.rnet.missouri.edu/Policy/htc-storage>

May 2017 Upgrades to the HPC (High Performance Computing) teaching cluster *tc.rnet.missouri.edu* to integrate it into the current Lewis infrastructure and to add more capacity.

June 2017 Announced plans to apply a set of urgent security patches that required a system reboot. More information at <https://access.redhat.com/security/vulnerabilities/stackguard>. Plan to start patching on Wednesday 6/28 at 8 am and completed by the end of the day - unless an exploit goes public before then at which point there may be an emergency patch depending on the situation. Linux users should also ensure that their systems are up to date as well. Jobs will continue to run as long as they are scheduled to

complete by 8 am on Wednesday. The plan included preserving the queue, so pending jobs will start after patching is complete.

September 2017 Training opportunities expanded to include a Software Carpentry Workshop on campus November 7th & 8th from 8am – 4:30pm. Software Carpentry aims to help researchers get their work done in less time and with less pain by raising awareness of basic research computing skills. For more about Software Carpentry, see: <https://software-carpentry.org/about/> The focus of the workshop is Automating tasks with the Unix Shell, Programming with R, and Version control with Git. Space is limited. Additional information at MU's workshop webpage: <https://mizzou-rcss.github.io/2017-11-07-mizzou/>

September 2017 Announced a new Cyberinfrastructure Engineer position that will work closely with researchers to help them evaluate their research computation needs and to recommend solutions that allow them to take full advantage of the resources available. Mizzou's cyberinfrastructure engineers develop scientific workflows to advance research in all areas, including bioinformatics, machine learning, material science, and computational chemistry. This new cyberinfrastructure engineer, will work on a broad range of projects across campus with opportunities to specialize in a number of domain areas, and will provide end-user support services to faculty, staff, and students in the use of research computing hardware and software as well as scientific applications and scientific workflows. This position's responsibilities will include, creating and maintaining training materials for research computing related software and hardware; conducting training; evaluating, installing, upgrading, and maintaining scientific software; and consulting with researchers on the development and migration of scientific workflows to a high-performance research computing environment. This position is within the Research Computing Support Services (RCSS) group in the Division of IT, which manages MU's shared research computing, storage, and networking infrastructure; and provides computing support, training, and consulting to MU's research community. RCSS works to ensure that the campus' research computing infrastructure meets the growing needs of the community. For additional information including Essential Functions and Minimal Qualification see: <https://doit.missouri.edu/services/research/research-computing-news/cyberinfrastructure-engineer-position/>

October 2017 Fourth Annual CI Day was held on October 4, 2017 (see above). The current storage system is having trouble keeping up with all the new cores and demanding applications. Upgrade to Lewis data storage performance was accomplished. Plan outage at end of November to switch over to new system. The move will be transparent, and everyone's /home, /data, and /group files will be moved to the new system. The new storage system is a parallel filesystem (zLustre) built for HPC applications and will be using the latest and fastest servers and networking (100Gbps). The new filesystem will be 960TB of raw storage (about 600TB usable) and is very scalable. The upgraded system will have a theoretical maximum of 40GB/s throughput compared to 5GB/s and should go from hundreds of MB/s typical throughput to GB/s for applications. In many applications, this can mean a 10x speedup in performance.

November and December 2017

No meeting was planned for November. The December 11 meeting was cancelled due to a flood in the Locust Street Building's second floor restroom.