Workshop 2 Summary:

The key points that surfaced during workshop 2 are:

- There are opportunities for small institutions to access resources at national centers at minimal cost. It is important for liberal arts consortia to develop relationships with such centers.
- Faculty need more knowledge of how they might use HPC in the humanities.
- Liberal arts faculty are excited about being able to collaborate on HPC Humanities projects.

Workshop 2 Overview:

Workshop 2, held on September 29, 2009, addressed the question of applications of high performance computing in the humanities and social sciences. It included the following presentations (most are available at http://cs.union.edu/~barrv/Grants/Teagle/teagle-overview.html#W1):

Presenters:

Orville Vernon Burton, Coastal Carolina University and founding Director of I-CHASS at the University of Illinois discussed the history and challenges of the use of advanced computing in the humanities and social sciences, and the work of the Institute for Computing, Humanities, Arts and Social Sciences at the University of Illinois at Urbana-Champaign.

Dean Rehberger, Michigan State University, explored the uses of high-performance computing in the humanities and argued for the importance of moving it from a "curiosity" to a more important role in teaching, learning, and research in the humanities.

Geraldine Heng, The University of Texas at Austin, discussed the opportunities available to faculty and students to become involved in national digital humanities initiatives.

Doug Reside, University of Maryland examined the ways in which Digital Humanities centers, which have thus far largely been focused on careful work on small data sets, are beginning to make use of the power of high performance computing.

Key points from the presenters:

Orville Vernon Burton

Key points:
- National centers, such as I-CHASS provide access to hardware, software and expertise.
• Cyberinfrastructure is advanced, collaborative, data acquisition and management services made available to researchers through high-performance networks.

• Cyberenvironments couple traditional desktop scientific and engineering applications with national and global cyber-resources to provide researchers an unprecedented capability to assemble, integrate, automate, and manage complex, collaborative projects.

• NEH is sponsoring collaborations between humanists and scientists.

• Obtaining funding for work in the humanities is a potential roadblock to moving forward.

• The humanities need to develop a culture of “tolerance of failure” if digital projects are to proceed.

• Key report: Our Cultural Commonwealth, The Report of the American Council on Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences, 2006. This report describes and analyzes the current state of high performance computing in the humanities and social sciences and looks at future needs and strategies.

• NCSA National Center for Supercomputing Applications (http://www.ncsa.illinois.edu/) runs workshops and collaborative opportunities.

• “How can HPC be used to move forth the frontiers of knowledge” “Not to ask the same questions we would ask only with larger data sets but to really ask different questions.”

• We need to be concerned about HPC resources only being a tool available to elite institutions.

• Hot areas for humanists and social scientists include data mining and mapping.

• Work in the humanities needs to be faculty and student driven and collaborative.

Dean Rehberger

Key points:

• Major question for Matrix is “How to make archives and objects within them more useful?”

• Labor at MATRIX is largely undergraduates.

• Some examples of Matrix resources:
  o American Black Journal
  o CVET
  o Studs Terkel
  o HNET – consortium of listservs
  o The Quilt Index
  o Explore PA History
  o Open source repository: KORA
  o Distributed rendering of animations

• The Quilt Index is an example of a humanities project that involves the creation of a large database.

• Need partnerships between computer scientists and humanists.

• Proliferation of media access now also requires the ability to index video and audio to get to specific portions. For example: Search across multiple media collections for “Nixon, angry”.


Geraldine Heng

Key points:
- Key resource is HASTAC (http://www.hastac.org/) a collaboratory for Humanities, Arts, Science and Technology. In particular, the HASTAC Scholars program for students.
- A key concept is data mining to create a culture of evidence. There is the need to better understand the concept of data mining to use advanced technology in the humanities.
- Consortial working groups are valuable to help faculty learn to apply high performance tools.
- A joint project with I-CHASS is interested in participation by liberal arts colleges. Liberal arts consortia should send working groups to Institutes and workshops.
- Another interesting idea was the Quilt Index, a case where the traditional focus of art history is circumvented to allow people to question what is art and who are artists. This suggests the possibility of using the web to establish a community of scholars and students so far under-served.
- Important resources:
  - Global Middle Ages Project
  - MappaMundie
  - SEASR
  - Class of civilizations project
  - Bibliopedia
    - Looking for patterns in huge sets of data.
    - Learning how to phrase questions in the most productive way.
    - Steep learning curve.
- Teragrid funding is available for humanities and social sciences.

Doug Reside

Key points:
- Use of digital resources in humanities areas is not necessarily the same as the need for high performance computing.
- Lev Manovich’s project looking at cultural patterns through the collections of the Internet Archive is an example of a current humanities project that really needs HPC. Characteristics include; really large data sets, large amounts of visual data, pattern analysis tools that are tailored to questions asked.
- Project Bamboo, funded by Mellon, is getting more organized and looks like an upcoming resource for digital humanities scholarship.
- There are four levels of activity in the use of advanced computing for the humanities: content, transcription and encoding of the content for digital use, widgets used to provide interfaces between the user and the encoded data, and user generated data.
- The “Cloud of Widgets” level at MITH is one of the layers of technology used in accessing and mining huge data sets. These tools and recipes for interacting with content in collections continue to be modified and new ones created are by MITH and by others.
- One of the major questions at MITH is “How to preserve born digital files?” How do we preserve the virtual environments and experiences that result from Humanities digital
scholarship? We have focused on the collections and objects within them but the process for maintaining outcomes of working with collections are not as evolved.

- Digital humanities scholars need to be involved in the theory of editing if not actually with some level of technical skill themselves.

**Key points from closing discussion:**

- Among the things to watch include: adding to the current emphasis on data mining and pattern analysis the goal of predicting as well. For example, in data mining and pattern analysis of conflict studies can we predict where the next conflict will arise?
- Need to include our students – undergraduate and graduate – in the research we are doing in humanities. Their understanding of disciplinary methods combined with their adaptability to “media” texts will open new ways of thinking about our scholarship.
- Have “student conferences” between campuses and invite HASTAC Scholars.
- Connect undergraduate liberal arts schools and the outcomes of projects from grants at research universities. The large project outcomes could feed directly into curricular components if connections are made and fostered.