

Towards a Collaborative Research Agenda for the Humanities, Arts, and Social Sciences in the Digital Age

Mark Kornbluh, Michigan State University

Dean Rehberger, Michigan State University

Vernon Burton, University of Illinois at Urbana-Champaign; NCSA

Kevin Franklin, University of Illinois at Urbana-Champaign, NCSA

The creative partnership between computer science, the humanities, and the social sciences—the core of what we now call “humanities technology”—is the cornerstone of the digital revolution. Knowledge is useless without meaning, and meaning is the essence of the humanities and social sciences. Humanities technology emerged in the 1960s as an interdisciplinary effort by humanists and social scientists to harness the power of the computer for their studies. The early pioneers used computers for textual and quantitative analysis, to provide new insights and new ways to teach. But it is with the advent of the Internet and the ensuing digital revolution of the last decade, which was initiated by the development of the MOSAIC web browser at the National Center for Supercomputing Applications (NCSA) at the University of Illinois, that humanities, as well as social science, technology has truly come into its own. In a world where information can be reduced into bits and bytes and communicated instantaneously, humanities and social science technology has rapidly emerged as a necessary and fundamentally interdisciplinary method of archiving, analyzing, and interpreting human activity and the human record. Humanities technology can, for the first time in world history, securely preserve and provide broad democratic access to the documents, images, languages, sound, and film that constitute the human record and facilitate its understanding. Social Science technology can now allow us to analyze, model, and even predict human social behavior on a scale that was unimaginable just a few years before.

The humanities encompass “those academic disciplines which study people—their ideas, their history, their literature, their artifacts, and their values.”¹ Humanities scholarship and sources, which comprise the disciplines of history, languages, letters, the humanistic social sciences such as anthropology, and such interdisciplinary fields as folklore, ethnic, and women’s studies, are accessible for the most part only in museums, classrooms, and libraries. Given such diverse fields, many research and technological barriers to universally accessible humanities knowledge exist, because of limits on storage and bandwidth and because technologies used at different institutions to create digital humanities projects are often incompatible.² Yet technologies such as pattern recognition and XML are helping to finally break down these barriers. Perhaps most important, the manuscripts, visual art, sound, video, and printed text that form the basis of humanities research have long defied the neat categories of scientific method.

Quantitative social scientific scholars have long utilized mainframes and personal computers for statistical analysis and other types of data processing, but qualitative social scientists have

1 Maryland Humanities Council, “What are the Humanities?” <http://www.gnet.net/MHC/definitn.htm>

2 See Paul Shabjee, “Primary Multimedia Objects and 'Educational Metadata': A Fundamental Dilemma for Developers of Multimedia Archives. (D-Lib Magazine, March 2000 <http://www.dlib.org/dlib/june02/shabjee/06shabjee.html>)

largely been in the same situation as the humanists described in the previous paragraph, and are subject to many of the same conditions. The situation is, however, rapidly changing for quantitative social scientists as well. Huge data sets comprised of data from large multi-year surveys or the census, email archives, large samples of internet sites over time, and video recordings require high performance computing and data management. New methods of analysis such as visualization and data-, text-, and image-mining are required to make sense of these large, complex datasets or to supplement more traditional statistical analysis. Simulation and mathematical modeling are becoming more widely utilized, often drawing on datasets for parameters or starting conditions, and these too often require supercomputing or grid computing support. Quantitative social science also has important contributions to make to public policy deliberations and to education via tools that can enable laypersons to interact with social science data and models.

Although computing and the information sciences have primarily concerned themselves with science, engineering, and business applications, “trends in computing and communications illuminate new opportunities associated with electronic ‘content’: visualization of all kinds of information, digitization of music and other audio material, electronic publishing with increasingly elaborate formats, and more complex paradigms for finding, selecting, displaying, and making use of greater quantities and varieties of information.”³ In addition, humanities and social science technology addresses the fundamental paradox of language and technology: the need to preserve vernacular and traditional modes of expression while providing methods of universal translation and communication.

Moment of Opportunity for the CIC:

Digital humanities and social science scholarship and research is at a turning point. Earlier this year, the ACLS released its long-awaited report, “Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences.” (The committee was chaired by the University of Illinois’ John Unsworth.) While the “convergence of advances in digital technology and humanistic scholarship is not new,” the report concludes that a “revolution similar to the transformation of science and engineering ... is inevitable for the humanities and social sciences.”⁴

The report urges universities, colleges, and funding agencies to invest in cyberinfrastructure in the humanities and social sciences as a matter of strategic priority. Key federal agencies have already responded to this recommendation with major new initiatives. The National Endowment for the Humanities has launched a major cross-cutting digital humanities initiative. National Science Foundation workshops have been held to begin to plan for new funding in this area and both agencies are coordinating with the Institute for Museum and Library Studies, major private

³ “Computing and the Humanities: Summary of a Roundtable Meeting.” ACLS Occasional Paper No. 41, 1998. <http://www.acls.org/op41-i.htm>

⁴ “Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences,” 2007, <http://www.acls.org/cyberinfrastructure/OurCulturalCommonwealth.pdf>

fundors (especially the Mellon Foundation), and large IT corporations (especially Google) to coordinate significant increases in funding for humanities and social science technology.

Most important, the ACLS report calls for the creation of National Digital Humanities Centers to support scholarship that contributes to and exploits cyberinfrastructure. Both NSF for the social sciences and NEH for the humanities have indicated a strong interest in funding a few such centers to pioneer the field.

The CIC is ideally situated to take the lead in developing a broad-based trans-institutional center for digital humanities and social sciences. CIC institutions currently house much of the leading work in digital humanities and social sciences. MATRIX, at Michigan State University, is the largest and best funded digital humanities center in the country; the Illinois Center for Computing in Humanities, Arts, and Social Science (I-CHASS), at the University of Illinois, is uniquely situated in partnership with NCSA and thus able to harness the most advanced cyberinfrastructure. CIC institutions house the most forward-looking Schools of Information in the country, schools that are training the next generation of information scientists with a keen interest in working in the humanities and social sciences, and recently announced that they would partner on a single integrated digital library system. Indeed, every CIC institution hosts significant digital humanities and social science projects and tentative collaborations are already underway. Michigan State University and Northwestern University are working to digitize the oral arguments before the U.S. Supreme Court. Michigan State University and the University of Illinois are working to establish an NSF-supported Social Science Learning Center, which will include the rich resources available at the University of Michigan's Interuniversity Consortium for Political and Social Research (ICPSR). The Universities of Illinois and Minnesota are together developing a next-generation online research system for social science using advanced cyberinfrastructure that will also be central to the virtual component of the proposed Social Science Learning Center. Michigan State University and the University of Illinois are in the early stages of planning a National Center for the Teaching and Learning of History. These developments are all examples of the enormous potential for collaboration across CIC institutions in this arena.

Towards a Collaborative Research and Teaching Agenda

The ACLS report sets a research agenda that includes the development of robust tools and the creation of extensive and reusable digital collections for the digital humanities and social sciences. To move forward, a new cyberinfrastructure needs to be developed and made available to scholars in the humanities and social sciences. High-powered, high-speed computing is necessary for these data-rich fields of research. New software tools need to be developed and new human skills learned. Across the CIC, projects are already underway to:

- mine and evaluate vast stores of digital objects (whether text-, image-, audio-, or video-based);
- model and simulate the complexities of the social life and human mind;
- analyze and explore a variety of data types;
- develop secondary repositories that support user-generated metadata and facilitate social tagging;

- create and disseminate complex learning objects.

The CIC should move forward boldly to inventory these efforts and build collaboration and coordination across institutions that harness the computing power and expertise already within the CIC. In the digital humanities and social sciences, advances in research, teaching, and public outreach have always gone hand-in-hand. CIC efforts should be focused on retaining these connections between advanced research, teaching, and public engagement.