Interest in cyberinfrastructure has accelerated since the 2003 report by the National Science Foundation’s Blue-Ribbon Advisory Panel on Cyberinfrastructure (known as the “Atkins Report”) and the publication of dozens of subsequent reports from across the academic landscape. One of those subsequent reports was the landmark *Our Cultural Commonwealth*, published in 2006 by the American Council for Learned Societies (ACLS). The ACLS report raised many cyberinfrastructure issues for the humanities and social sciences, notably on the “institutional innovations that will allow digital scholarship to be cumulative, collaborative, and synergistic.”

By David Green and Michael Roy

Illustration by Michael Gibbs, © 2008
This seemed a good time for Academic Commons—an online journal designed to surface and debate the technologies that might serve as a “catalyst for reshaping liberal arts education in the 21st century”—to capture a range of perspectives on the challenges and opportunities that cyberinfrastructure presents for the liberal arts and for liberal arts colleges. The December 2007 issue of the journal thus pulled together a range of voices from various parts of the world of liberal arts education to begin a conversation about the difference that cyberinfrastructure might make to liberal arts disciplines and institutions and how those disciplines and institutions might prepare for anticipated changes in scholarly communication. Yet this was by no means the only attempt to advance the work started by the Atkins Report and Our Cultural Commonwealth. Other notable contributions have included issues of CTWatch Quarterly, First Monday, and the Journal of Electronic Publishing and a report from an NSF-JISC workshop.

In this EDUCAUSE Review article, we will summarize what these various communities have been thinking and doing about cyberinfrastructure for the liberal arts, look at some emerging models for transinstitutional collaboration and institution building, and suggest some specific steps that campuses can take to stay engaged, be prepared, and move this agenda forward.

Paradigm Shifts
One of the biggest issues surrounding cyberinfrastructure and the liberal arts is that, overall, a major cultural shift in both the conceptualization and the practice of scholarship is required to take full advantage of what is being offered. Even though National Science Foundation (NSF) Director Arden L. Bement Jr.’s promises radical, even revolutionary, developments in computing, networking, and storage capacities—from grid computing to huge sensor arrays and data storage to new visualization capacities—individual scientists will not, in general, experience a major behavioral shift. True, more scientists may rely on vast data repositories like the Sloan Digital Sky Survey (http://www.sdss.org/) or the GenBank database (http://www.ncbi.nlm.nih.gov/Genbank/) instead of on their own research or experimental data, but for most, the transition will be a quantitative change. But for scholars in the liberal arts—those in the humanities and social sciences, for example—a major paradigm shift is required, both in intellectual outlook and in social organization.

First, although the liberal arts culture is created through scholarly communication—journals, conferences, teaching, the activity of scholarly societies, and the continuing evolution of disciplines—much of the daily activity of the humanities and social sciences is rooted in the assumption that research and publication form essentially an individual rather than a collaborative activity. The tools, the capabilities, and the benefits of larger and deeper engagement with others beckon, but there are few takers. Our Cultural Commonwealth noted, “Lone scholars . . . are working in relative isolation, building their own content and tools, struggling with their own intellectual property issues, and creating their own archiving solutions.” How will the shift to a collaborative approach to research and publication actually happen? Will it be led by the existence of a new generation of easily available, collaborative, increasingly semantic tools that will make the mechanics of finding and working with partners significantly easier? Or will it occur as a result of the pressure of effectively organizing ways of responding to the sheer amount of digital data that is becoming available?

The second part of the necessary paradigm shift is an extension of the first and involves scholars’ preoccupation with the publication of a print monograph, despite declining numbers of print publications and decreasing library purchase budgets. Tenure is still tightly tied to print monograph publication, with little legitimacy of scholarship produced in new media, whether by individuals or in collaboration, and create procedures for evaluating these forms of scholarship. This finding seems particularly dismal considering that the promise of cyber-scholarship lies less in the confining boundaries of the printed page and more in the fluid exchanges and availability of primary data through digital means. An example of this availability is the Humanities E-Book (HEB) project of the ACLS (http://www.humanities ebook.org), an online, fully searchable collection of 1,700 books, recommended and reviewed by scholars and featuring unlimited multi-user access. Now adding 500 books each year (available through institutional and individual subscription), this project is definitely a step in the right direction. But what source of inspiration and what other new models will be needed to persuade more departments and more individuals to take the plunge?

Parallel to the cultural and behavioral shifts needed is the major combined challenge of budget and leadership: where will the funds required for implementing and taking advantage of cyberinfrastructure come from in a community that, in
the words of Gary Wells, has “had to make do with inadequate tools, incompatible standards, tiny budgets and uninterested leaders.” There’s a large gap between rhetoric and vision and what faculty have and do right now. How do we argue the dollar for cyberinfrastructure against the other demands on a limited budget? How can the budget be expanded, especially when there are strong calls to make this cyberinfrastructure both a means for greater collaboration within and among academic disciplines and also a route out to the general public? Who will lead this call to arms?

Fifteen years after the introduction of the World Wide Web, the humanities funding agencies are coming around. After a slow start in supporting digital initiatives, the National Endowment for the Humanities (NEH) has recently announced a suite of smartly conceived new programs (although with no real new dollars for the agency as a whole). The NEH is now fairly vigorously partnering not only with the Institute of Museum and Library Services (IMLS, the only agency mandated to fund digitization) but also with NSF, the Department of Energy (DOE), and the United Kingdom’s Joint Information Systems Committee (JISC). Its new Office for Digital Humanities offers Digital Humanities Start-Up Grants, a new training program (Institutes for Advanced Topics in the Digital Humanities), NEH Fellowships at Digital Humanities Centers (FDHC), and the JISC/NEH Transatlantic Digitalization Collaboration Grants, together with some newly focused support for digital initiatives by older programs. Most revolutionary might be the partnership with the Office of Science at the DOE behind the new Humanities High Performance Computing (HHPC) initiative to explore how high-performance and grid computing can be used in the humanities, including grants for time and training on supercomputers. Such partnering and leveraging with larger, better-endowed agencies, here and abroad, has to be part of the new funding strategy.

The role of the Andrew W. Mellon Foundation cannot be ignored here, since it has proved to be a major agent of change (and not only because its grants in this area—around $250 million—exceed the funding of the National Endowment for the Arts and the NEH combined). Viewing digitization itself only as a means to a greater end and famously insisting on projects’ sustainability, the Mellon Foundation recently funded the eighteen-month collaborative Bamboo Planning Project (issuing from the University of California, Berkeley, and the University of Chicago), an attempt to discover whether focusing on shared tools and services, used by ever-widening communities of practice, will begin to make more of a systemic change. Bamboo, in looking for partners, is insisting on cross-functional support at each of the five to ten institutions that would participate (this would need to include humanities faculty, computer science faculty, the local IT organization, and the administration).

Resources and Repositories
In the first generation of digital humanities, simple digitization sometimes could be an issue. If a work was not already in digital form, the question was how best to go about digitizing it. Often the result was piecemeal digitization, done according to the needs and idiosyncrasies of individual scholars. John Unsworth has noted that one of the first requirements for cyberinfrastructure is simply to have the material ready at hand. This is “content as infrastructure”: the bringing together of standards-based digital representations of the full array of cultural heritage materials in as interoperable, usable, and sustainable a way as possible.” A case study of some of the frustrations that result when material is not easily available in digital form can be seen in the 2006 report “Using Digital Images in Teaching and Learning.” Given the glacial pace with which many institutions were thinking through the required supply and presentational infrastructure for digital images, many teachers assembled their own digital image collections (usually to no standards and with less-than-necessary metadata). Many of the institutions in the study appeared not to understand that digital supply was now essential (rather than a niche interest) and had to be thought through as a campus-wide infrastructure issue. As the art historian Dana Leibsohn has put it, cyberinfrastructure will be “useless if it cannot revolutionize image access and metadata management—art history’s most anxiety-producing fetishes.” Images and their metadata have, after all, proven more difficult than imagined: ARTstor is a more complex proposition than JSTOR, for example. Despite evidence to the contrary, many museums are still afraid they will “give away the store” if they act more like educational institutions than treasure-house gatekeepers. But with today’s massive digitization projects—spurred by Google Book Search and given integrity and breadth by the Open Content Alliance—the tipping point might be close at hand.

Another aspect of “content as infrastructure” was the subject of the 2007 NSF-JISC workshop on the implications of the rapid multiplication of digital content on network infrastructure. A central issue that emerged during the workshop was scholars’ underuse of network-accessible digital repositories. There are several reasons for this, but what was apparent, echoing the earlier period of digitization, was the need for a coordinated, standards-based approach (as well as for the greater involvement in the process by the scholars for whom the repositories were being built). One goal of the emergent agenda was for digital content to be “routinely and rigorously collected, curated, managed, and preserved” in order to be consistently discoverable and usable “via common infrastructure and tools through space and time, and across
disciplines, stages of research, and modes of human expression.” The workshop outlined an international seven-year program to achieve such an infrastructure.\textsuperscript{14} Intellectual property concerns are a continuing hurdle, with online solutions still only partial and inconclusive. Material under copyright is hard to find and to use in digital form; the traditional sense of “fair use” is difficult for many to translate to the digital stage. Unsworth has cited intellectual property as “the primary data-resource-constraint in the humanities,” noting that robust solutions were for him the “primary ‘cyberinfrastructure’ research agenda for the humanities and social sciences.”\textsuperscript{15} Commenting from the perspective of one working in the sciences and the humanities, Michael Lesk has noted that much of the cyberinfrastructure-related discussion he hears in the humanities is not so much “about how to manage the data or what to do with it, but what you are allowed to do with it.”\textsuperscript{16} Some combination of technical, social, and legal answers is surely called for here.

\section*{Approaches to Cyberscholarship}

Moving to the possibilities of cyberscholarship, defined by the authors of the NSF-JISC workshop report as “new forms of research and scholarship that are qualitatively different from traditional ways of using academic publications and research data,”\textsuperscript{17} we see two very broad approaches, with the second generally building on the first.

The first approach to cyberscholarship in the liberal arts is data-driven, relying on the algorithmic sorting, sifting, and selecting of materials within the ever-growing digital repositories. This approach has its roots in much of the earliest literary and linguistic computing work, going back to 1949, when Roberto Busa began his \textit{Index Thomisticus}, a lemmatized concordance to the works of Saint Thomas Aquinas and related authors. That tradition of text analysis resulted, through intense social and intellectual organization, in such core tools and standards as the Text Encoding Initiative’s \textit{Guidelines for Electronic Text Encoding and Interchange}. Today, rather than analyze one or more texts on a single computer, we can harness the power of the evolving cyberinfrastructure and advanced tools for very fast searching, sifting, and analysis across different kinds of data sets and collections of often disparate materials in different repositories. And here lies the concern behind the NSF-JISC workshop: that to date, work both in organizing individual digital repositories and in managing and coordinating these repositories internationally has fallen behind and needs some focused attention. Within this context, it’s worth pointing out that the size of liberal arts databases is, of course, exploding exponentially. Cathy Davidson notes that although today the Sloan Digital Sky Survey uses 40 terabytes of data, the USC Shoah Foundation Institute for Visual History and Education’s archive of nearly 52,000 videotaped testimonies from Holocaust survivors and other witnesses requires 200 terabytes of compressed data.\textsuperscript{18}

The second approach to cyberscholarship in the liberal arts derives less from the vector of the growth in the sheer power of computers and the size of the archives and more from the cultural shaping of computer networking and developments on the social networking road to the semantic web. This approach was outlined by Janet Murray in her \textit{Academic Commons} article.\textsuperscript{19} A large part of Murray’s approach has to do precisely with encouraging and observing the emergence of new web-based genres. Cyberscholarship will discover its rules and emerge as its own genre, which Murray defines as the “elaboration of a cognitive scaffold for shared knowledge creation.” Online games have emerged as one of the first web genres, partly because of their highly developed rule structures, and may provide one model. Another key route to new procedures could be provided by the semantic web, in which chunks of coded knowledge from different communities and disciplines can be “understood” by computers and related one to another. However, the original semantic web route imagined by Tim Berners-Lee, a vision in which scholars in different disciplines agree on common tagging vocabularies, does not seem to be working (and probably would never work for those in the liberal arts).
A more likely route to semantic solutions might be through the evolution of "smart" social networking and collaboration tools. Zotero (http://www.zotero.org), for example, has progressed rapidly from being a desktop citation and annotation tool to a collaborative object discovery and sharing tool. Likewise, the 2008 edition of The Horizon Report outlines the next generation of social networking: "social operating systems." Recognizing one’s “social graph”—that is, one’s web of social and intellectual activity—social operating tools will uncover the people associated with particular topics or digital objects. Berners-Lee has now famously said that just as the Internet connects computers, and the web connects documents, so the “Giant Global Graph” connects not so much people as the people-centered world of content.

For Murray, the answer to understanding all the data we will have at our fingertips will not be through “algorithmic tricks,” because so much of the human experiential context of data, including our values and assumptions, cannot be encoded and included in systems. Murray states that rather than an instrument for mining, the computer should be seen as a “facilitator of a vast social process of meaning-making.”

Cyberscholarship will differ, in its processes and its products, from traditional analog scholarship. The increased social and collaborative basis of resource discovery, analysis, discussion, and argument will be joined by the availability of resources in a greater range of formats beyond text. Images, moving images, sound recordings, maps, and GIS will be able to come into play, while the use of APIs and “mash-ups” will enable a huge flexibility in and variety of presentational forms.

In the Academic Commons special issue, one of the conclusions of the art historians’ roundtable discussion of the future of the discipline was the expectation that in producing the results of scholarship, they would need to be thinking about formats “more interesting than the book.” From Robert Darnton’s cogent presentation of the pyramid model of the electronic book’s ability to layer argument, evidence, and primary data, to the by-now classic work of William Thomas and Edward Ayers in producing an electronic article that would “fuse the electronic article’s form with its argument” to the “multimodal” production of Vectors: Journal of Culture and Technology in a Dynamic Vernacular (http://www.vectorsjournal.org), new and “more interesting” models and formats are indeed being formulated and produced.

The puzzle is, What is it going to take for models to take hold? Will it be the network of digital humanities centers that are now very consciously beginning to work more closely together? Will it be radically increased funding from government and private agencies? Will it be the vision and practical accomplishments of the Bamboo Planning Project? And who will take the initiative? Will faculty lead and institutions accommodate, or will institutions lead?

New Institutional Arrangements and New Kinds of Institutions

For cyberinfrastructure to take root in the liberal arts, a key component, identified in Our Cultural Commonwealth and elsewhere, is staffing. A brief stroll down the lane of innovative work in the liberal arts clearly demonstrates this point. Looking at the teams at places like the Perseus Project (Tufts University), the Institute for Advanced Technology in the Humanities (the University of Virginia), the Center for Digital Research and Scholarship (Columbia University Libraries), the Scholarly Technology Group (Brown University), MATRIX (Michigan State University), the Institute for Multimedia Literacy (University of Southern California), the Center for History and New Media (George Mason University), or Connexions (Rice University), anyone can see that it “takes a village” to produce this type of cyberscholarship. Who lives and works in such a village? Faculty are there, but so too are software programmers, designers, project managers, digitization specialists, copyright lawyers—and the list goes on. One obvious worry is that this sort of endeavor is so
expensive that it will become the exclusive enclave of the richest of institutions, a result that would undermine the populist hopes expressed both in the Atkins Report and in Our Cultural Commonwealth. If liberal arts cyberscholarship is in fact the way forward, then a path must be forged to allow this resource-intensive approach to be possible at the broadest number of locations.

One possible path involves the development and deployment of templates that allow for the creation of scholarship by tapping into libraries of digital objects and creating nonlinear, hypermedia, network-based documents. Projects like the New Media Consortium's Pachyderm, the Institute for the Future of the Book's Sophie, and Michigan State University's Project Builder allow us to imagine a future in which faculty working in relative isolation can collaboratively produce new scholarship using these networked authoring environments without the need for a team of expensive, local specialists. Certainly those working in scholarly publishing efforts like Vectors and Gutenberg-e also understand that without better authoring tools, the per-unit cost of scholarship will be one (although not the most pressing) of the factors that inhibit broader adoption. Beyond template-based authoring systems, liberal arts cyberscholarship is made possible more broadly through the development of tools that allow for collaborative work across distances. George Mason's Zotero and MIT's Cross Media Annotation System point the way here.

Although these developments on various campuses seem promising, this path—with innovation taking place on one campus and then being diffused and amplified over the network—is familiar and suffers from familiar problems. How do these sorts of projects take root, go viral, and become part of a national or international infrastructure? One model is through privatization. Good ideas from higher education become attractive to venture capitalists, who take the idea private and sustain it by creating a product that they then sell for a profit. There are countless examples of academic software packages that have followed this model. Another is the open source model. It is possible to imagine a universe in which liberal arts cyberinfrastructure technologies are sustained through open source communities, much as happened with Linux, Apache, and Firefox. A variation, exemplified by projects like DSpace and Sakai, is what we like to call the “pay-to-have-a-say” open source model, in which members of the core development work need to make a financial commitment to the effort.

Yet another model is the creation of new or newly imagined transinstitutional associations that can help reduce the risks and investments that individual institutions must take to get started and that can facilitate the building of discipline-based communities of practice. One example is the National Institute for Technology and Liberal Education (NITLE), which offers an institutional repository service that members can choose to subscribe to. With a relatively small investment (less than $10,000), an institution can quickly have access not only to an instance of DSpace that it did not have to install itself but also to a community of practice that can help the institution figure out how to organize its efforts. NITLE offers similar services for Sakai and Moodle. The key difference between this model and a more traditional outsourcing model is that NITLE takes on the responsibility not only of brokering the service from a vendor but also of facilitating the work of the community of practice. Although it remains to be seen if this model would scale down to more discipline-specific applications and services, the model seems important to consider as an alternative to rolling the dice on whether privatization or an open source community will develop around a particular project. Our Cultural Commonwealth calls for the creation of a series of national humanities and social science computing centers, one of which has started at the University of Illinois. It remains an open question as to whether or not such centers will have the ability to serve the broad array of interests and needs of the nation while living within the constraints of a single institution.
One way to most concretely think about what one of these new models might mean on any given campus is to consider the following question: Would it be better to invest in hiring, training, and supporting a java programmer for the institution, at approximately $100,000 per year, or to invest in funding, at about $20,000 per year, a consensual pool of money that will collectively support a team (distributed) that works on common problems? Although each of these models—privatization, open source, pay-to-have-a-say open source, and members-only or emergent transinstitutional associations—has its place in this emerging landscape, the key shift in thinking must be away from what can be done locally on an individual campus and toward how the campus can be connected to other campuses and how it can contribute to the refining of these new ways of doing scholarship and teaching. No single institution will ever have all the necessary resources to support these efforts on its own. The software, texts, media, expertise, and communities of scholars are moving very quickly to reside on the network—and not necessarily on local networks. Even though this idea may sound a bit like Nicolas Carr’s “IT is irrelevant” argument, with all important IT (and library) functions moving to the cloud, it is a bit different. Carr may be right that some core infrastructure (e-mail, productivity applications, ERP) will end up being outsourced to commercial providers, but those particular pieces of infrastructure that are core to scholarship and teaching seem more likely to end up in a different cloud or set of clouds.

**Steps to Take**

In lieu of a grand conclusion, we would like to revisit the eight recommendations published in Our Cultural Commonwealth and add our own suggestions for steps that liberal arts colleges might take to work toward those goals.

1. **Invest in cyberinfrastructure for the humanities and social sciences, as a matter of strategic priority.** Many liberal arts colleges do not have the discretionary funds to reallocate resources to support this work. They can, however, make liberal arts cyberinfrastructure a focus of professional development and future hiring, to be prepared once this future becomes a reality. For example, if cyberinfrastructure for the humanities and social sciences is thought of as part of a larger shift in library collection—from collection development to content curation—then liberal arts colleges will need to develop the capacity to curate digital content and to make that activity a priority.

2. **Develop public and institutional policies that foster openness and access.** Open access and revised approaches to intellectual property are key components in this effort, since scholars need to have as complete a library as possible of primary and secondary materials in digital form. Liberal arts colleges can track the work of the Create Change (http://createchange.org) educational initiative and follow its recommendations about changing the campus culture to promote open access publishing, although the problem of tenure remains vexing. Colleges can also promote conversations about copyright and fair use that will improve access to the twentieth- and twenty-first-century cultural materials that many humanists and social scientists need to carry out their scholarly and teaching work.

3. **Promote cooperation between the public and private sectors.** Most liberal arts colleges do not develop technology, nor do they get involved in technology-transfer efforts. That said, efforts like the Bamboo Planning Project have recognized the important role that such scholars/teachers can play in the definition of the needs of liberal arts cyberinfrastructure. Finding ways to participate in these planning efforts will ensure that the emerging tools
and standards can serve the broadest population of students, scholars, and teachers.

4. **Cultivate leadership in support of cyberinfrastructure from within the humanities and social sciences.** Finding ways to educate provosts and presidents about these issues is important. So too is finding those faculty who hold leadership roles in their professional associations and making sure that they understand what the stakes are. Representatives of liberal arts colleges can attend conference sessions that highlight recent developments, taking the news back to campus. Fascinating conferences-within-a-conference take place at the Modern Language Association conferences, the American Historical Association conferences, and other meetings where cutting-edge work is presented and analyzed. Colleges can encourage faculty who attend these conferences to go to these sessions.

5. **Encourage digital scholarship.** Providing ready access to emerging tools and templates, supporting digitization efforts, and having staff available to support these sorts of efforts form one obvious tack. Another harder but probably more important strategy is to find a way to encourage interesting but difficult campus conversations about cyberscholarship and the tenure process for faculty working in the humanities and social sciences. Colleges can also track granting opportunities that provide resources for faculty to experiment and innovate in this area, although without a reward system structured to encourage this kind of work, grants alone will not do.

6. **Establish national centers to support scholarship that contributes to and exploits cyberinfrastructure.** As noted above, we would amend this recommendation to include other types of consortial arrangements that might not fit exactly into what the authors of *Our Cultural Commonwealth* were suggesting. As these centers and other consortia come online, liberal arts colleges can explore what these arrangements offer for collaborative work and can consider how their offerings map onto local needs and interests. Colleges can
make their local needs and interests known well by having frank conversations with consortial partners about what is needed to share expertise within the consortium, allowing for faculty on one campus to rely on the services of staff on another campus. And if a college doesn't have consortial partners, it should get some!

7. **Develop and maintain open standards and robust tools.** Most liberal arts colleges do not develop software and are often left out of the standards-making process. Finding a way into those conversations is vitally important, since these standards and tools will become the core infrastructure that defines how scholarship is made and shared in the coming decades. Joining organizations such as the Coalition for Networked Information (CNI) is one way to participate in those conversations; simply reading through the agendas for CNI meetings is another way to track important developments on this particular front. Those liberal arts colleges that do develop software or data should become familiar and then obsessed with standards both in the data produced locally and in the applications developed and/or deployed.

8. **Create extensive and reusable digital collections.** Most liberal arts colleges have unique collections that, once digitized and cataloged, can become the raw materials of liberal arts cyberscholarship.

In addition, both EDUCAUSE and CLIR are developing communities of practice around cyberinfrastructure in general. Tracking developments and learning about best practices through these professional organizations form another important step to take.

Oh, and did we mention this before? Buy more bandwidth!

---

**Notes**


3. See the Academic Commons web page “About the Commons,” <http://www.academiccommons.org/about/about-academic-commons>.

4. See “Cyberinfrastructure and the Liberal Arts,” special issue edited by David L. Green, *Academic Commons*, December 2007, <http://www.academiccommons.org/commons/annoucement/table-of-contents>, featuring thirteen articles: an art historian’s review of Our Cultural Commonwealth and an interview with that report’s chair, John Unsworth; considerations of how cyberinfrastructure may affect the humanities and the sciences differently; a review, an essay, and a roundtable discussion on different approaches to cyberscholarship; three takes on institutional change, including interviews with two key figures holding out for radical change in university and museum settings and a proposal to extend the function of the college museum into the curriculum; and three essays detailing how cyberinfrastructure may be built at the campus level.


15. Unsworth, “Cyberinfrastructure for the Humanities and Social Sciences.”


