

E-Research Is a Fad: Scholarship 2.0, Cyberinfrastructure, and IT Governance

Brad Wheeler

I've seen this movie before, though it was cast in a different setting. When I was an assistant professor of information systems in the business school, the mid to late 1990s seemingly declared that all things that could be digital *would* be digital. "E-commerce" was the rage, and I taught my first MBA e-commerce course in 1995. It was soon supplanted by a more proper e-business moniker a few years later, and companies everywhere started e-business projects or "e" divisions. eBay, Amazon, and Yahoo were start-ups (as were flameouts WebVan, Boo.com, and eToys), and Bill Gates abruptly rechartered Microsoft's entire strategy to embrace this new Internet.

Astute observers of commerce at the time, however, had it right in seeing e-business more clearly. E-business was best understood and pronounced as "business," where the "e" is silent. While speculative equity money hastened the dizzying pace and the Internet provided instant access to the marketplace at scale, the phenomenon was nothing more than pedestrian Schumpeterian creative destruction at work. By 2007, efficient digital connections in the supply chain and with customers became a common and somewhat invisible element of effective commerce.

E-science, e-research, and e-scholarship can expect the same. They are best understood as "e-research," where the "e" will eventually be silent. Scholars in the sciences, humanities, and arts will identify those digital tools that advance their endeavors, that maintain and even enhance rigor and quality, and embed those tools and practices in their daily work. National funding agencies are providing the venture capital to hasten the pace of this evolution, and scholarly conferences and some leading projects already demonstrate efficacy in digital scholarship. "Data-driven science" is debated as a fourth paradigm alongside classic theory, experiment, and computational science.

What then is the role for college and university CIOs and their leadership teams as this e-research movie plays out? Is there a strategic

agenda where attention, advocacy, and action can be effective? In this short essay, I contend that understanding Scholarship 2.0, developing scalable campus cyberinfrastructure, and seriously assessing governance of research IT are essential for great 21st century universities.

Twenty-First Century Scholarship

In 1945, James Bryant Conant, Harvard president, made three assertions regarding the elements for advancing pure and applied science:

There is only one proved method of assisting the advancement of pure science—that of picking men of genius, backing them heavily, and leaving them to direct themselves. There is only one proved method of getting results in applied science—picking men of genius, backing them heavily, and keeping their aim on the target chosen.¹

While Conant wrote of science, with some liberty, let me extend these assertions to the humanities, arts, and all areas of scholarly inquiry. The thesis is that real advancements in knowledge and human creativity come from exceptional minds that have all the tools and resources they need. In the 20th century, this was the great scientific lab working in some isolation at a single university or institute. The enormous expense of the best research tools, however, rationed their scarcity to elite labs with (presumably) the best minds to leverage the investment.

Now seems a prudent time, in this first decade of the 21st century, to question if and how Conant's assertions should steer research investments for this century. We already observe several important trends: First, we have many more faculty—both men and women of genius—at research-intensive institutions worldwide. Their scholarly work products are increasingly multiauthored, and often with scholars from two or more institutions and multiple disciplines. Second, the Internet has hastened the informal and formal communication of scholarly results, and the research process is increasingly a process of contributing to community data repositories and conducting further research from community data. The Human Genome Project, NEESGrid, and the National Virtual Observatory are early exemplars, and the new Large Hadron Collider (\$8 billion) with more than 2,000 collaborating physicists from 34 countries shows community and leveraged investment at scale. Finally, these connectionist endeavors are not only with a research community, but many advances in human knowledge rely on deep interdependencies between disciplines (for example, metabolomics' use of biology and chemistry).

Who, then, is the 21st century “them” in Conant’s “back them heavily”? Increasingly, we see that cross-institutional research communities, operating at scale and pace and making use of common and specialized facilities, are becoming an essential means of advancing modern scholarship. Numerous reports from the National Science Foundation, the American Council of Learned Societies, and others call for investments in cyberinfrastructure to support these communities.² Charles Vest uses the term *meta-university* to describe these connectionist trends among universities for research, teaching, and learning: “... we are seeing the early emergence of a meta-university—a transcendent, accessible, empowering, dynamic, communally constructed framework of open materials and platforms on which much of higher education worldwide can be constructed or enhanced.”³

Scholarship 2.0

In the business world, Tim O’Reilly coined the term *Web 2.0* as a revolution that argued for “harnessing collective intelligence” via the Internet as platform and for developing applications that harness network effects through communities, social networking sites, wikis, folksonomies, and other collectivist tools.⁴ Arguably, we can apply the term *Scholarship 2.0* to label a similar adaptation of IT tools as a platform and distributed communities to harness collective intelligence of scholars: “Scholarship 2.0: The rise of IT, digital repositories, and electronic collaboration in achieving and improving the quality of the scholarly endeavors of research, teaching and learning, and service.”⁵

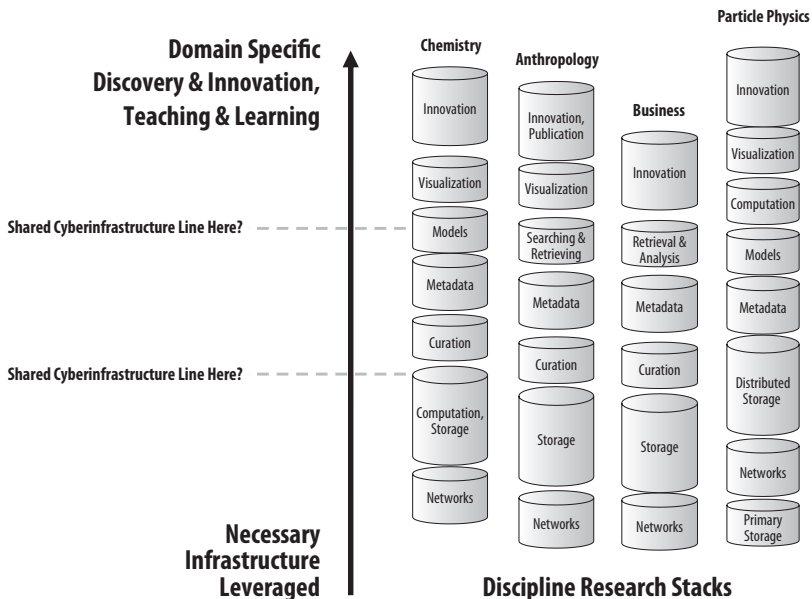
The form of Scholarship 2.0 and its pace of adoption differ among disciplinary scholarly communities. For example, some scholarly communities have already transitioned the scholarly record to bypass traditional journal publishers and embrace open access and rapid publication. Other disciplines have not yet chosen or found a way to do so. At Indiana University (IU), research in ethnomusicology is leading the way with sophisticated video markup, annotation, and indexing as a form of digital scholarship in the humanities, but many other humanist fields have only modest experience with (or perceived need for) new digital tools. Many medical sciences are now moving rapidly from research paradigms of laboratory experimentation to extensive data modeling and simulation for understanding disease and potential treatment.

Institutional Cyberinfrastructure

It is tempting to look at national and international endeavors for cyberinfrastructure development as the province of national funding agencies. Not so. Each institution—especially the research university—is a microcosm of the cyberinfrastructure challenge, and there is a need for alignment of local cyberinfrastructure investments with national and disciplinary investments.

Figure 1, drawing on the notion of the software stack from computing, illustrates a conceptual relationship of disciplinary research stacks and institutional cyberinfrastructure. Each research discipline varies in its need for components of cyberinfrastructure (a nonexhaustive list). Some research disciplines may need dedicated lambdas of optical networking for moving massive quantities of data. Others need lots of high-performance computation cycles or massive storage. Others may need only trivial uses of networks and storage but have extensive needs for sophisticated metadata consultation or specialized visualization devices. The blend of these needs changes over time as a discipline evolves in its use of Scholarship 2.0 practices. For example, while simulation researchers have long been voracious consumers of computational cycles, they have now turned their appetites to storage and networks. One researcher recently told me

Figure 1. Scholarly Infrastructure



he had to dumb-down the data rate of his simulations as the network and storage were incapable of handling the volume.

Left to their own devices and with sufficient funding, researchers would rightly seek to solve their IT needs by purchasing and providing services for their projects or academic departments—Conant’s “leave them to direct themselves.” The proliferation of small (and growing) computing clusters is one example of this, and costs for electricity, cooling, and equipment life cycle funding are often overlooked until they become problematic. Likewise, the human staffing for system administration or metadata development often falls between the cracks. In many cases, these critical roles become filled by doctoral students, postdocs, or even faculty, and this is, at best, a misallocation of skills and time.

Figure 1 asserts that there are favorable economies of leverage—getting more for less—in the lower parts of the research stack (networks, computation, and so forth), where economies of scale, scope (serving multiple disciplines), and career path development can retain the best technical talent. The effort of scholars is best applied at the top of the stack—the edge—where their creativity and innovation are the very basis of research. Thus, using doctoral students for advanced server administration and security patching does not advance scholarship and may, in fact, leave a project’s IT investments vulnerable to an ever-growing array of security threats.

The critical question for CIOs and research leaders is to begin thinking of disciplinary IT needs as shared cyberinfrastructure investments that can be leveraged across the scholarly endeavors of an institution.⁶ Such thinking should go beyond the obvious boxes and wires of IT to include help and consulting for effective IT use, security and system administration, and other hidden but very real needs and costs. For example, IU has decided that extensive data storage is common cyberinfrastructure. IU’s Massive Data Storage System, which provisions more than 4 petabytes of storage and is automatically replicated at two sites for continuous backup, is available to any graduate student or faculty member without chargeback. Thus, digital texts for the humanities sit alongside financial databases and genomics data as a leveraged service to all disciplines. The service has full help desk and consultant support.

Figure 1 also illustrates the institutional cyberinfrastructure question by asking where the line is drawn for common resources. Drawing the line lower, near networks, leaves the remainder of the research stacks to the individual scholars or their academic unit. Drawing the conceptual line higher means that more capabilities are provided in some shared-services model but at the risk of possibly impeding disciplinary creativity.

It is important to note that not all of the shared services should come from an IT organization. Cyberinfrastructure thinking enlightens many opportunities for partnerships with university libraries or academic units (a school or department), or even for leverage via interinstitutional consortia.

While Figure 1 presents some conceptual elegance to the notion of leverage (that is, that it costs considerably less to provide 2 petabytes of fully supported storage than to provide four 500-terabyte systems), there is often less elegance in funding models. For many institutions, this is a chicken-and-egg situation, where the money needed to provide quality, shared services is fragmented in the research projects and academic units. Even if the scholars actually would prefer to use common institutional infrastructure, they can't trust it until it exists and has proved itself. It can't exist until sufficient money is aggregated to fund it and competently manage it to demonstrate service quality.⁷

This may seem a typical university turf and funding battle, but that simplistic view fails to understand the opportunities and challenges of interconnected, 21st century scholarship. Even rich institutions with seemingly unlimited resources to support scholars and projects may find their efficacy in 21st century communities held back by 20th century thinking veiled by largesse of purse.

IT Governance for Institutional Cyberinfrastructure

For many institutions, progress on an institutional cyberinfrastructure strategy can be meaningfully advanced only as part of a holistic and rigorous discussion of IT governance. Though the term *governance* is often used loosely with many implied meanings, Peter Weill and Jeanne W. Ross define IT governance as “specifying the decision rights and accountability framework to encourage desirable behavior in using IT,” and much of this section is adapted from their book, *IT Governance*.⁸

All institutions have IT governance, but some have much more effective IT governance than others—anarchy is also a form of IT governance. Five types of IT decisions are presented in sequence as part of establishing effective IT governance for research IT:

Decision 1: IT principles—clarifying the institutional role of IT

Decision 2: IT architecture—defining integration and standardization requirements (if any)

Decision 3: IT infrastructure—determining common, shared, and leveraged services

Decision 4: Disciplinary application needs—specifying the disciplinary need for IT applications (either purchased or internally developed)

Decision 5: IT investment and prioritization—choosing which initiatives to fund and how much to spend

In this brief essay, I'll address only decisions 1, 3, and 5, without reference to who makes these decisions, as that will vary greatly with local context. Ideally, IT principles should flow from institutional strategy, with a clear linkage between the two. Practically, one of the first explicit decisions should be a philosophy for research technologies. Examples of IT principles include the following:

Principle 1: Each research project should bear its own full costs.

Principle 2: Technology resources for research are efficiently matched to projects through user chargeback. (User fees fund research technologies.)

Principle 3: Research technologies are provided in abundance and with few limitations to encourage use in all forms of scholarship. (Research technologies are centrally funded or via broad-based tax.)

Principle 4: Academic units and projects are responsible for all specialized technology needs beyond basic networks, computation, and modest storage. (Research technologies have a blended funding model.)

IT principles, whether established through explicit process or tacit neglect, establish much of the domain for decision 3 of IT infrastructure (that is, the shared cyberinfrastructure line in Figure 1). Principle 1 above would effectively limit the creation of common infrastructure as rich projects and disciplines as they would likely do their own thing. Likewise, principle 2 would make it difficult for less well funded disciplines or graduate students to have access to advanced research technologies. Principle 3 overtly endorses cross subsidization for the breadth of the institution. Principle 4 invokes ongoing turf and service coordination challenges among units over what is common infrastructure or disciplinary applications.

Finally, the IT principles also substantially shape decision 5 regarding IT investment priorities and how much to spend. IT principle 3, a philosophy of abundance, necessitates substantial IT expenditures with coordination of timing, policy, and support.

CIO Engagement and Advocacy

University CIOs have an opportunity and responsibility to help university leaders frame cyberinfrastructure investment philosophies and decisions. For some, the opportunity arises through the pain of no space for fragmented research computing clusters or the tragedy of a research data security breach. Others may be able to frame the discussion as part of a strategic plan that looks to the essential elements of leading 21st century institutions.

Whatever the impetus for local action, there is no doubt that Scholarship 2.0 practices and needs for leveraged cyberinfrastructure will continue to advance for the foreseeable future. CIOs should engage in relationship building, educational efforts with administrative colleagues, and extensive data gathering to understand the current expenditures, value, and challenges that researchers have with IT. For many institutions, a comprehensive review of funding may reveal that the institution does not have a funding problem for research technologies; rather, there is a complex coordination problem of where funding goes and where it is needed for leveraged cyberinfrastructure.

There are also considerable time lags and path dependencies in developing leveraged shared services, effective partnerships with other administrative units (for example, libraries), and skilled staff with disciplinary expertise.

Conclusion

This essay argues that notions of e-research and e-science are transient ideas as scholars embrace and mainstream a range of digital tools for their work. The practice of digital scholarship may be labeled as Scholarship 2.0 to reinforce that it is first and foremost the classic and rigorous form of scholarship, but it now further harnesses distributed intelligence of global scholarly communities via new collaborative tools. Conant's wise advice to advance pure science by engaging scholars of genius and leaving them to direct themselves remains unchanged. The "back them heavily" admonition, however, has in part shifted to multi-institutional scholarly communities that make extensive use of local, national, and international cyberinfrastructure. A portion of that financial backing—especially to advance the whole of a university—must now be targeted at leveraged cyberinfrastructure investments. Effective and purposeful IT governance is an essential tool for framing discussions of institutional cyberinfrastructure.

Endnotes

1. James Bryant Conant, "Letter to the Editor," *New York Times*, August 13, 1945, 18.
2. Anthony J. G. Hey and A. E. Trefethen, "The Data Deluge: An e-Science Perspective," in *Grid Computing: Making the Global Infrastructure a Reality*, eds. Fran Berman, Geoffrey Fox, and Anthony J. G. Hey (New York: John Wiley and Sons, 2003), 809–24; *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (Washington, DC: National Science Foundation, January 2003), <http://www.nsf.gov/cise/sci/reports/atkins.pdf>; *Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences* (New York: American Council of Learned Societies, December 2006), http://www.acls.org/uploadedfiles/publications/programs/our_cultural_commonwealth.pdf; and *Cyberinfrastructure Vision for 21st Century Discovery* (Washington, DC: National Science Foundation, March 2007), http://www.nsf.gov/od/oci/CI_Vision_March07.pdf.
3. Charles Vest, "Content and the Emerging Global Meta-University," *EDUCAUSE Review* (May/June 2006): 30, <http://www.educause.edu/ir/library/pdf/ERM0630.pdf>.
4. Tim O'Reilly, "Web 2.0 Compact Definition: Trying Again," *O'Reilly Radar* (December 10, 2006), http://radar.oreilly.com/archives/2006/12/web_20_compact.html.
5. Brad Wheeler, "Open Source 2010: Reflections on 2007," *EDUCAUSE Review* (January/February 2007), 67, <http://www.educause.edu/ir/library/pdf/ERM0712.pdf>.
6. Bradley C. Wheeler and Michael McRobbie, *Final Report of the Indiana University Cyberinfrastructure Research Taskforce* (Bloomington, IN: The Trustees of Indiana University, 2005), <http://hdl.handle.net/2022/469>; Thomas J. Hacker and Bradley C. Wheeler, "Making Research Cyberinfrastructure a Strategic Choice," *EDUCAUSE Quarterly* 31, no. 1 (2007): 21–29, <http://www.educause.edu/ir/library/pdf/EQM0713.pdf>.
7. Brad Wheeler, "Research Technologies: Edge, Leverage, and Trust," in "The Organization of the Organization: CIOs' Views on the Role of Central IT," *EDUCAUSE Review* (November/December 2007): 44, <http://net.educause.edu/ir/library/pdf/ERM0761.pdf>.
8. Peter Weill and Jeanne W. Ross, *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results* (Boston: Harvard Business School Press, 2005); see also Peter Weill and Jeanne Ross, "A Matrixed Approach to Designing IT Governance," *Sloan Management Review* 46, no. 2 (Winter 2005): 26–34.

Bibliography

Conant, James Bryant. "Letter to the Editor." *New York Times*, August 13, 1945, 18.

Cyberinfrastructure Vision for 21st Century Discovery. Washington, DC: National Science Foundation, March 2007. http://www.nsf.gov/od/oci/CI_Vision_March07.pdf.

Hacker, Thomas J., and Bradley C. Wheeler. "Making Research Cyberinfrastructure a Strategic Choice." *EDUCAUSE Quarterly* 31, no. 1 (2007): 21–29. <http://www.educause.edu/ir/library/pdf/EQM0713.pdf>.

Hey, Anthony J. G., and A. E. Trefethen. "The Data Deluge: An e-Science Perspective." In *Grid Computing: Making the Global Infrastructure a Reality*, edited by Fran Berman, Geoffrey Fox, and Anthony J. G. Hey, 809–24. New York: John Wiley and Sons, 2003.

O'Reilly, Tim. "Web 2.0 Compact Definition: Trying Again." *O'Reilly Radar*, December 10, 2006. http://radar.oreilly.com/archives/2006/12/web_20_compact.html.

Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences. New York: American Council of Learned Societies, December 2006. http://www.acls.org/uploadedfiles/publications/programs/our_cultural_commonwealth.pdf.

Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. Washington, DC: National Science Foundation, January 2003. <http://www.nsf.gov/cise/sci/reports/atkins.pdf>.

Vest, Charles. "Open Content and the Emerging Global Meta-University." *EDUCAUSE Review* (May/June 2006): 18–30. <http://www.educause.edu/ir/library/pdf/ERM0630.pdf>.

Weill, Peter, and Jeanne Ross. "A Matrixed Approach to Designing IT Governance." *Sloan Management Review* 46, no. 2 (Winter 2005): 26–34.

Weill, Peter, and Jeanne W. Ross. *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*. Boston: Harvard Business School Press, 2005.

Wheeler, Brad. "Open Source 2010: Reflections on 2007." *EDUCAUSE Review* (January/February 2007): 48–67. <http://www.educause.edu/ir/library/pdf/ERM0712.pdf>.

Wheeler, Bradley, and Michael McRobbie. *Final Report of the Indiana University Cyberinfrastructure Research Taskforce*. Bloomington, IN: The Trustees of Indiana University, 2005. <http://hdl.handle.net/2022/469>.