Transforming Knowledge Communities through Cyberinfrastructure

Daniel E. Atkins, Professor
University of Michigan, EECS & School of Information
Ann Arbor, Michigan 48103 USA
atkins@umich.edu

Abstract

A new age is dawning in the conduct of knowledge-based activities, pushed by continuing progress in computing, information, and communication technology; and pulled by the expanding complexity, scope, and scale of today's challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive *cyberinfrastructure* on which to build new types of knowledge communities and to pursue the creation, dissemination, preservation, and reuse of knowledge in new ways with increased efficacy. Digital library R&D is already situated within the goal of a global information infrastructure and should now evolve to support the even bolder aspiration of global, functionally complete, cyberinfrastructure-enabled knowledge communities (CKCs).

Keywords

Digital libraries, cyberinfrastructure, knowledge communities, grid communities, collaboratories, global information infrastructure, virtual communities

1. Introduction

The organizers of this conference have very appropriately linked the topic of *digital libraries* with the concept of *knowledge communities* – the building blocks of a *networked information society.* While multi-faceted "digital library" R&D should continue, particularly in a socio-technical context, there is a growing consensus that research, development, innovative deployment, and institutional transformation in this area should be situated in the more comprehensive context of a global, ICT-based infrastructure – *cyberinfrastructure.* This infrastructure is a platform to serve human society in the creation, re-use, preservation, and dissemination of knowledge. The focus of this paper is to elaborate, advocate, and illustrate this theme.

2. Decade of Digital Library Initiatives

As well developed by Christine Borgman in her book *From Gutenberg to the Global Information Infrastructure* [1], the challenges and promise of digital libraries should be embedded in the concept of a global information infrastructure (GII). My assertion in this paper is that the GII should now be situated as a component of an even bigger concept that I will refer to as *cyberinfrastructure-enabled knowledge communities* (CKC). A knowledge community is a group of people engaged in creation, preservation and/or dissemination of knowledge. We are not envisioning one huge knowledge community, but rather huge numbers of specialized, interoperable knowledge communities often with overlapping participants. The value of a university, for example, lies in the complex relationship it creates between knowledge, communities, and credentials [2].

Borgman also includes an excellent historical survey of digital library research including a pointer to the important *Sourcebook on Digital Libraries* compiled by Ed Fox in the early 1990s [3] as a resource for a series of workshops that, under the stewardship of Y.T. Chen, Steve Griffin, and later Michael Lesk at the U.S. National Science Foundation, launched the Digital Library Initiatives [4] [5]. These R&D initiatives linked with similar activities in Europe and Asia have produced insights, architectures, tools, unique digital collections, and an interesting array of evocative test beds – living specifications – for digital library systems.

International digital library research initiatives also seeded the establishment of digital library production activities in universities, catalyzed formation of content federations, and helped launch spectacular commercial successes including Google. To borrow from the title of Y. T. Chien's paper in these proceedings, digital libraries have become "disruptive technologies and have seeded innovation."

These initiatives have also created an international interdisciplinary R&D community drawn from computer and information science, social science, librarians, publishers, and a host of disciplinary specialists from the sciences, engineering, and humanities. Regular digital library R&D initiatives, publications and conferences series, including this one, exist in all regions of the world and can be easily found through web searching and/or *D-Lib Magazine* [6].

Another important contribution from the decade of digital library research is deeper conviction by researchers, practitioner, and sponsors that the design and evaluation of digital libraries must blend the talents and perspectives of technologists, the ultimate disciplinary users, and social informatics specialists who are skilled in design for usability and assessment of longitudinal impact. Teams of these complementary specialists need to develop mutual respect and build mutual self-interest around iterative cycles of research, design, implementation, and evaluation. The approach needs to be human-centered and promote co-lateral learning between the contributing disciplines and practices. The DLI activities funded by NSF have, for example, included participation of people from 35 different academic disciplines [7].

This social informatics perspective is elaborated in Borgman's GII book [1] and aspects of it are treated in her invited paper in this conference, *The Interaction of Community and Individual Practices in the Design of a Digital Library.* This socio-technical systems approach is even more crucial in the design, understanding, and diffusion of CKCs.

3. Other Trends Enabling CKCs

But in addition to the international digital library initiatives of the past decade there are other forces and advances that merge to empower CKCs. We review some of these in this section.

3.1 Core technology

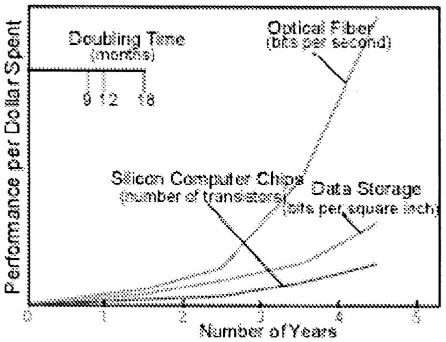

Figure 1. Performance/$ doubling times.

As shown in Figure 1 from reference [8] the electo-optical-magnetic technology components underlying digital processing, storage, and communication technology continues an incredible path of exponential growth in performance and attendant decrease in cost. The performance per dollar spent for silicon computer chips continues to double every 18 months (Moore's curve); the doubling time for data storage density in 12 months; and the doubling time for optical fiber bit transmission is only 9 months. Optical technology is now far outpacing silicon chips and data storage and is the basis for the increasing prevalence of geographically distributed, but functionally tightly couple systems and organizations – grids, collaboratories, and digital library federations. Wireless technology, at increasing bandwidth, is also adding flexibility, mobility, and pervasiveness of access.

3.2 Dominance of digital information

The information explosion continues. The 2003 version of the How *Much Information* report from the UC-Berkeley School of Information and Systems [9] estimates that in recent years the stock of information in the world has been growing at 30% per year and most of it in digital format. The worldwide production of original information in 2002 was in the range 3.4 to 5.4 exabytes with the upper estimate assuming no storage compression. An exabyte is a million terabytes. The print collection of the U.S. Library of Congress is about 10 terabytes, and therefore the annual growth in new digital information is equivalent to adding about half a million Libraries of Congress per year.

3.3 The World Wide Web

In the last decade, the WWW has moved from a tool for particle physicists into most every facet of society. The OCLC Web Characterization Project [10] conducts an annual Web sample to analyze trends in the size and content of the Web. The number of unique websites as of 2002 (latest data posted) is 8.7 million. The size also includes country and language statistics as well as linkage patterns.

A concise summary of trends and other references is given in [11]. The CyberAtlas website [12] of statistics on Internet size and use includes projections for 2004 of close to 1 billion Internet users worldwide. The web continues to evolve in size and function. The original designer Tim Berners-Lee anticipates the "semantic web" [13] as the next big step.

3.4 High-performance computing and computational science

The TOP500 project [14] was started in 1993 to provide a reliable basis for tracking and detecting trends in high-performance computing. Twice a year, a list of the sites operating the 500 most powerful computer systems is assembled and released. The November 2003 rankings placed the NEC machine at the Japanese Earth Simulator Center at Yokohama in first place (36 TFlops); the ASCI Q at the U.S. Los Alamos National Laboratory in second place (14 TFlops), and a homegrown cluster of 1100 Apple MacIntosh G5s at Virginia Tech, USA in third place (10 TFlops). TeraFlops is a million-million, or a trillion (10^{12}) floating-point arithmetic operations per second. We have moved into the era of local "clusters" and wide area "grids" [15] of machines that collectively provide terascale computation [16] and are well on the way to the petaFLOPS (10^{15}) realm.

The significance of this computational horsepower is most vivid in the high fidelity simulation of complex natural systems such as the Earth's global environment [17] or the Universe. Computation is now the basis for a third approach, in addition to experimentation and theory, to scientific discovery, engineering and design.

3.5 Collaboratories, Grids, KDI, E-science, Cyberscience

The concept of a *co-laboratory,* or *collaboratory* —a laboratory without walls built upon distributed information technology—was defined at an invitational National Science Foundation (NSF) workshop at Rockefeller University [18] in 1989 and later elaborated and sanctioned in a National Research Council report [19] published in 1993. Interdisciplinary research to create, deploy, and evaluate specific collaboratories has been funded by several agencies. Overviews of collaboratories and efforts to create general principles of collaboratory design and analysis are available at [20, 21].

In 1999 Foster and Kesselman [22] published the concept of the grid, "a new concept for computing infrastructure." The concept focused initially on the aggregation of distributed computers to provide re-allocation and load balancing for computers analogous to that of the electricity power grid. The concept has evolved to include all of the functions of a collaboratory – linking computation, information, people, and instruments. An expanded version of the Grid book was published in 2003 [23].

In the late 1990s the U.S. National Science Foundation launched a Knowledge and Distributed Intelligence (KDI) initiative to explore reshaping relationships among people and organizations, and transforming the processes of discovery, learning, and communication based in part on the explosive growth in computer power and connectivity [24]. Many of the funded projects focused on distributed and multidisciplinary collaboration facilitated by information and communication technology (ICT).

Fortunately, this initiative consisting of 71 projects funded over two years has been studied by Cummings and Kiesler [25]. This recent study showed that there is still much to be done to successfully coordinated ICT-linked projects distributed over geographic and organizational distance (distance matters), but it also showed that multidisciplinary projects were superior to unidisciplinary in producing new ideas.

Other R&D to adopt ICT to enable scientific research and education is being done under the label *e-science* [26] or *cyberscience* [27].

3.6 Middleware Initiatives

Middleware is software that connects otherwise separate resources across networks. More specifically, the term refers to a layer of software services between the network and applications for managing security, access, and information exchange. It is designed to let people transparently use and share distributed resources, such as computers, data, networks, and instruments and to develop effective collaboration and communications and other advanced services to expedite research and education. Middleware provides a working architecture and approach that can be used by a large set of Internet and network users.

The growing interest in creating networked environments that span both geographic and organization domains is driving R&D in middleware [28]. Examples include Globus [29] and the Internet2 Shibboleth Project [30] that are together developing architectures, policy structures, practical technologies, and open source implementations to support controlled sharing in the higher education world. The emergent Web Services [31] and more ambitious Semantic Web [13] activities are also contributing important components of middleware.

3.7 Embedded networked sensing systems

Micro-electro mechanical systems (MEMS), nanotechnology, wireless communication, and robotics are enabling creation of large-scale distributed systems composed of thousands of smart sensors, actuators, and autonomous systems embedded in the physical world. [32] These systems, sometimes called the smart sensor web [33], are showing the potential to revolutionize the way we observe our physical world for scientific understanding as well as monitoring built infrastructure (buildings, bridges, roads, etc.) for maintenance and security.

3.8 Visualization and the Human Interface

All of the above has also been accompanied by a growth in ubiquity and modality for interaction between humans and ICT-based systems. Displays range from walls to palm size devices and graphical interfaces navigated with a mouse are being supplemented with voice and haptic interaction. An except from the website of the Electronic Visualization Laboratory at University of Illinois-Chicago [34], developers of the *CAVE®* and *ImmersaDesk®* virtual reality systems provides a glimpse of trends in this area:

> *Current research focus is tele-immersion - users in different locations around the world collaborate over high-speed networks in shared, virtual environments. Related research interests include scientific visualization, paradigms for information display, network performance monitoring and measurement, distributed computing, sonification, human/computer interfaces, new methodologies for informal science and engineering education, every citizen interfaces, and abstract math visualization.*

3.9 Ubiquitous Computing

The final trend we note here, an unifying one sometimes called the *third wave* in computing, is the concept of ubiquitous computing fathered by the late Mark Weiser at XEROX PARC [35]. Ubiquitous computing is roughly the opposite of virtual reality as explored in the EVL mentioned above. Where virtual reality puts people inside a computer-generated world, ubiquitous computing forces the computer to live out in the world with people. Ubiquitous computing is a very difficult integration of human factors, computer science, engineering, and social sciences.

4. The Cyberinfrastructure Movement

4.1 Background

The intersection of trends described in the last section has created the opportunity to develop a comprehensive new infrastructure for the support of knowledge-based activities. Growing numbers of researchers are seeing the potential for transformative impact – even revolution – on the conduct of research and allied learning. Recent efforts towards this have arisen, for example, from e-science activities in the UK [26] and EU [36], a variety of pilot projects sponsored by the U.S. NSF under the Information Technology Research (ITR) program, and advanced simulation capabilities such as publicized through the Japanese Earth Simulator Project [17]. Examples of these and other projects are referenced in [37].

In 2001 leaders of the U.S. National Science Foundation adopted the term *cyberinfrastructure.* The term *infrastructure* emerged in the 1920s to refer to the roads, bridges, rail lines, and similar public works that are required for an industrial economy to function. The new term *cyberinfrastructure* denotes systems of information and communication technologies together with trained human resources and supporting service organizations that are increasingly required for the creation, dissemination, and preservation of data, information, and knowledge in the "digital age." Traditional infrastructure is required for an industrial economy; cyberinfrastructure is required for a knowledge economy.

Infrastructure is informally defined as systems and services you notice only when they breakdown. Although it is taken for granted and suffers from low status with respect to the activities it serves, the provision of infrastructure is one of the most complex and costly investments made by society. On page 18 of Borgman [1] is a multi-dimensional answer to the question "What is infrastructure?" The most relevant thing to note for our purposes is that infrastructure is not just equipment or even a technical framework – it is the complex interaction of technology, social and work practices, and standards.

Although the word "cyberinfrastructure" does not roll off the tongue, it is a useful at present to emphasize two things: 1) that we must now invest in ICT as institutionalized, sustained, evolving but robust infrastructure; and 2) that "cyber" has some important different properties than traditional "built" infrastructure. On the negative side, it generally depreciates much more rapidly that bricks-and-mortar infrastructure, but on the positive is generally more sharable and general purpose. Cyberinfrastructure and its use is both an object of research as well as an enabler of research

4.2 The NSF Cyberinfrastructure Report

In February 2003 an advisory panel on cyberinfrastructure appointed by the NSF and chaired by the author of this paper issued a report entitled *Revolutionizing Science and Engineering through Cyberinfrastructure* [38] based upon extensive testimony and survey of the NSF research community. The principal finding from he report is

> *that a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information and communication technology; and pulled by the expanding complexity, scope, and scale of today's problems. The capacity of this technology has crossed thresholds that now make possible a comprehensive "cyberinfrastructure" on which to build new types of scientific and engineering knowledge environments and organizations and to pursue research in new ways and with increased efficacy. The cost of not doing this is high, both in opportunities lost and through increasing fragmentation and balkanization of the research communities.*
>
> *Such environments and organizations, enabled by cyberinfrastructure, are increasingly required to address national and global priorities such as global climate change, protecting our natural environment, applying genomics-proteomics to human health, maintaining national security, mastering the world of nanotechnology, and predicting and protecting against natural and human disasters; as well as to address some of our most fundamental intellectual questions such as the formation of the universe and the fundamental character of matter.*

Figure 2 is a schematic of the general functions included in a cyberinfrastructure stack beginning at the bottom with networking, operating systems, and middleware providing the generic capabilities for management, transport, and federation of systems and services (tools) described in the five columns. The goal is to allow customized CKCs to be created efficiently and effectively using facilities, tools, and toolkits provided at the cyberinfrastructure layer to federate the requisite resources.

Physical, intellectual, and long-term access to data, information, and knowledge (eventually both explicit and tacit) is a primary service required in cyberinfrastructure. Digital libraries and indeed global information infrastructure is subsumed here.

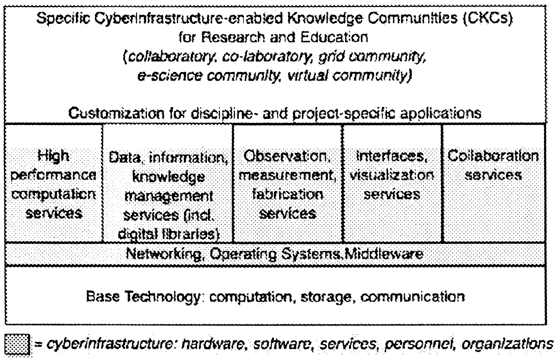

Figure 2. Integrated cyberinfrastructure for knowledge communities.

Figure 3 illustrates two complementary dimensions of an advanced cyberinfrastructure program. One dimension is technology capacity—a measure of the processing power, data volumes, and data rates that can be handled in the IT-based research environment. The other dimension is the *functional comprehensiveness* of the resources needed by a research team or organization to do its work. This is a qualitative measure indicating, for example, what percentage of colleagues, relevant scientific literature, archival data, instruments, and other critical facilities are easily available through the cyberinfrastructure-based environment, i.e., through the virtual organization—the collaboratory—the grid community.

The NSF Panel identified projects heading towards creating functionally complete virtual organizations – cyberinfrastructure-enabled communities including all the people, literature, data, instruments, computational models, etc. need for their work. Achieving that goal, as is the case in creating a digital library federation involves mastery of a complex set of technical and social factors. The goal of an advance cyberinfrastructure program should be to explore a range of projects (in the grey area of Fig. 2) – some pushing the limits of technology and others pushing on such things as ubiquity, breadth of participation, easy of use, and functional completeness. Methods of human-centered design explored by the digital libraries research community are critical to creating successful CKCs.

The opportunity now before the research community is to establish new working environments that are advanced, innovative, and innovating in terms of both capacity and comprehensiveness. The shaded region represents such a region and the general goal of an ACP is to move both *within* and *among* broad fields of science into these regions, and to move the regions up and to the right. As the combined region capacity and functional comprehensiveness increases and is adopted more broadly, the payoff will likely derive from enhancing both "depth" and "breadth" approaches to discovery.

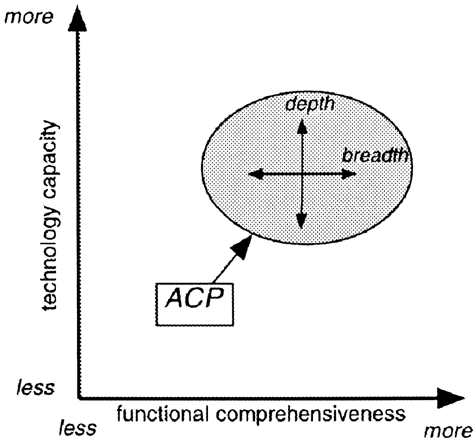

Figure 3 - Increasing technology capacity and functional comprehensiveness of cyberinfrastructure.

4.3 Broader implications

This paper has focused most directly on CKCs conducting scientific and engineering research. We have noted the growing needs of many research fields to work in more interdisciplinary global teams using larger, richer, and more diverse sources of expertise, information, observation, and computational tools. These needs coincide with a state of information technology that has now crossed thresholds of capacity and function that help meet these needs through cyberinfrastructure that enables virtual federation of resources to create new forms of knowledge work environments. Success at this in the science and engineering research enterprise would have enormous impact on other knowledge-intensive activities including commerce and education.

The concept of advanced cyberinfrastructure and its use as a platform for virtual organizations of many types, durations, and missions obviously has application beyond scientific and engineering research. The author is aware of initiatives to explore cyberinfrastructure for the humanities [39], for rapid-response to natural and man-made disasters [40], for transforming the future of higher education including the notion of global universities [41-43], and for supporting engaged institutions [44] to enhance civil society.

The sound bite version of the findings from the NSF Cyberinfrastructure Report might be *CKCs have new options for what is done, how it is done, and who participates*. The *who* participates – the democratizing potential of cyberinfrastructure, the opportunities for "digital inclusion" – is for many one of the most compelling potentials for CKCs.

5. Some Implications for the Digital Libraries Community

Libraries – traditional, hybrid, or digital only – are important institutions in the infrastructure for a knowledge society and now need to be conceived as part of a cyber-infrastructure –enabled communities served by resources (people, information, facilities) distributed over global networks, with effective intellectual team work flowing through all four variations of same-different time and place. In closing we mention a few of the challenges and opportunities this presents for the digital library community including researchers, developers, operational entities, policy-makers, and funders. I offer complementary thoughts on this topic specific to the research library community in [45].

5.1 Onwards to the GII

First, the vision of a global information infrastructure that emerged with the massification of the Internet is far from reality and continues as an important overarching goal for the digital library community. A recent NSF sponsored workshop on research directions for digital libraries [7] reaffirms this and uses the phases *ubiquitous knowledge environments* and *information ether*. The notion of ubiquitous also includes increasing the scope and scale of information resources and services.

5.2 Scholarly communication systems

One trend worthy of more emphasis is to situate digital library R&D in the context of systems that support new modes of scholarly inquiry and communication – often just called "scholarly communication". This topic is being approached from several potentially complementary perspectives including:

1. A new system is needed as a response to "runaway" costs and copyright policies of commercial academic publishers.
2. Digital technology disaggregates the traditional scholarly communication chain and allows new players to assume new functions (e.g. publishers as libraries, libraries as publishers).
3. The move towards federated "digital repositories" [46] in universities could create an infrastructure for open publication of scholarly artifacts. These systems could be the platform for new approaches to review and credentialing based upon the social life of the information, i.e. use, rating, and reputation systems.

See [47]for more on these topics from a recent U.S. National Academies symposium on the future of scholarly communication.

5.3 Increased scope, scale, and persistence of collections

The creation of CKCs in scientific research is creating new demands for curation of large and diverse data sets and the need for more research on the automatic production of metadata. Scientists want to capture provenance or "audit-trials" on data streams as they undergo processing into information and knowledge. Engineers want to automatically capture, store, and retrieve design rationale.

The enormous increase in computational power is enabling scientists to create comprehensive models of complex natural systems but this requires the interoperation of both computational models as a well as associated data sets. Challenges of federation, interoperability, production of quality metadata, terms and conditions for use, credentialing, and long-term stewardship of data sets all require the attention of the digital library community.

To the extent that the process of knowledge creation occurs in CKCs we have the possibility to capturing not only the end products of the work but also the process by which the results are produced. All of the synchronized streams of interaction between participants in a collaborative session could be captured as a "session object" and made available through digital libraries or archives. In the UARC and SPARC experimental collaboratory projects [48, 49], for example, the data gathering campaigns from observational stations in Greenland were captured as session objects that could be revisited, annotated, and reused by other scientists at other times in other locations. The data flows through emerging scientific CKCs will be astronomical in rate and volume.

Guaranteed long-term access to digital collections a critical need and the lack of it is a critical barrier to broader adoption of cyberinfrastructure in many fields. The need for high quality curation of enormous and complex data sets is gaining importance in science in part because of the growing power of data mining techniques that motivate reuse of the data. The lack of guaranteed access over the ages to both bits and their rendering environment is probably the single greatest barrier to the adoption of new multimedia digital genres in much of the humanities. The field of history, for example, depends critically on access to the human record over centuries.

5.4 Shortage of human attention

The expanding creation and use of CKCs brings with it not only the threat of more information overload but now also an increased participation overload. Through CKCs people can participate in more communities in various roles. They can be a teacher in one community and a student or merely an observer in another. An amateur astronomer can make seminal discoveries using the Virtual Observatory [50]. A professor can easily and occasionally assist and share his virtual laboratory with a high school science class (see for example the Bugscope Project [51]).

The availability of human attention is not growing and thus the same technology that is creating the overload needs to be use to relieve it. The enhancement of human productivity and the stewardship and augmentation of human attention needs to continue as a top priority in the digital library research community. This goal includes a broad array of research topics including more powerful searching, abstracting, knowledge and process representation, intelligent agent technology, visualization, and usability. The Dl workshop makes similar recommendations in terms of enhancing human productivity. They include 1) employing context at the individual, community, and societal levels to improve performance, 2) developing algorithms and strategies for transforming data into actionable information, 3) demonstrating the integration of information spaces into everyday life, and 4) improving availability, accessibility and, thereby, productivity.

6. References

1. Borgman, C.L., *From Gutenberg to the global information infrastructure : access to information in the networked world.* Digital libraries and electronic publishing. 2000, Cambridge, Mass.: MIT Press. xviii, 324.
2. Brown, J.S. and P. Duguid, *Universities in the Digital Age,* http://www2.parc.com/ops/members/brown/papers/university.html. 1995.
3. Fox, E.A., *Sourcebook on Digital Libraries: Report for the National Science Foundation* http://fox.cs.vt.edu/DLSB.html. 1993, VPI&SU Computer Science Dept.: Blacksburg, Virginia. p. 439.
4. NSF, *Digital Libraries Initiative Phase 1 (DLI1)* http://www.dli2.nsf.gov/dlione/.
5. NSF, *Digital Libraries Initiative Phase 2 (DLI2)* http://www.dli2.nsf.gov/.
6. D-LIB, *D-Lib Magazine* http://www.dlib.org/.
7. Larsen, R. and H. Wactlar, *Draft Final Report from Wave of the Future: NSF Post Digital Library Futures Workshop,* http://www.sis.pitt.edu/~dlwkshop/. 2003.
8. Stix, G., *The Triumph of Light.* Scientific American, 2001.
9. Lyman, P. and H.R. Varian, *How Much Information? 2003* http://www.sims.berkeley.edu/research/projects/how-much-info-2003/. 2003, School of Information Management and Systems: Berkeley.
10. OCLC, *OCLC Web Characterization Project* http://www.oclc.org/research/projects/archive/wcp/default.htm. 2002, OCLC.
11. O'Neill, E.T., B.F. Lavoie, and R. Bennett, *Trends in the Evolution of the Public Web 1998-2002,* http://www.dlib.org/dlib/april03/lavoie/04lavoie.html. D-Lib Magazine, 2003. **9**(4).
12. *CyberAtlas: Internet Statistics and Market Research for Web Marketers:* http://cyberatlas.internet.com/. 2004.
13. W3C, *Semantic Web Activity website,* http://www.w3.org/2001/sw/.
14. *Top 500 Supercomputer Sites,* http://www.top500.org/.
15. *Global Grid Forum Website,* http://www.gridforum.org/.
16. *Teragrid Project website,* http://www.teragrid.org/.
17. *Earth Simulation Center website,* http://www.es.jamstec.go.jp/esc/eng/. 2002.
18. Wulf, W.A., *The National Collaboratory - A White Paper, Appendix A in Towards a National Collaboratory, the unpublished report of an invitational workshop held at Rockefeller University, March 17-18, 1989 (Joshua Lederberg and Keith Uncapher, co-chairs.).* 1989.
19. NRC, *National collaboratories : applying information technology for scientific research.* 1993, National Academy Press: Washington, D.C. p. xi, 105.
20. Finholt, T.A., Olson, G. M., *From laboratories to collaboratories: A new organizational form for scientific collaboration.* Psychological Science, 1997. **8**(1): p. 28-35.
21. *Science of Collaboratories Project Website,* http://www.scienceofcollaboratories.org/.
22. Foster, I. and C. Kesselman, *The grid : blueprint for a new computing infrastructure.* 1999, San Francisco: Morgan Kaufmann Publishers. xxiv, 677.
23. Foster, I. and C. Kesselman, *The grid : blueprint for a new computing infrastructure.* 2nd ed. 2003, San Francisco, CA: Morgan Kaufmann, Elsevier Science. 700.
24. *NSF Knowledge & Distributed Intelligence (KDI) Inititative,* http://www.ehr.nsf.gov/kdi/default.htm. 1999.
25. Cummings, J. and S. Kiesler, *KDI Initiative: Multidisciplinary scientific collaborations.*

2003, U.S. National Science Foundation: Arlington, VA.

26. *The UK Research Councils e-Science Core Program website, http://www.research-councils.ac.uk/escience/.*
27. Nentwich, M., *Cyberscience; research in the Age of the Internet.* 2003: Austrian Academy of Sciences Press. 490.
28. *NSF Middleware Initiative website, http://www.nsf-middleware.org/.*
29. *Globus World website, http://www.globusworld.org/.*
30. *Shibboleth Project website, http://shibboleth.internet2.edu/.*
31. W3C, *Web Services Activity website, http://www.w3.org/2002/ws/.*
32. *Center for Embedded Networked Sensing, http://cens.ucla.edu/.*
33. Piquepaille, R., *The Smart Sensor Web, http://radio.weblogs.com/0105910/2003/10/05.html#a603*. 2003.
34. DeFanti, T., *Electronic Visualization Laboratory Website, http://www.evl.uic.edu/info/index.php3*.
35. Weiser, M., *Ubiquitous Computing Website, http://www.ubiq.com/hypertext/weiser/UbiHome.html*.
36. *EU Sixth Framework Program (2002-2006) website, http://europa.eu.int/comm/research/fp6/index_en.html*.
37. Atkins, D.E., *Transformation through Cyberinfrastructure-based Knowledge Environments*, in *The Economic and Social Implications of Information Technology (In Press)*, W.H. Dutton, et al., Editors. 2004, The MIT Press: Cambridge, MA.
38. Atkins, D.E., Droegemeier, K.K., Feldman, S. I., Garcia-Molina, H., Klein, M.L., Messerschmitt, D.G., Messina, P., Ostriker, J.P., Wright, M.H., *Final Report of the NSF Blue Ribbon Advisory Panel on Cyberinfrastructure: Revolutionizing Science and Engineering Through Cyberinfrastructure, available at http://www.cise.nsf.gov/sci/reports/toc.cfm*. February 2003.
39. ACLS, *Website for the American Council of Learned Societies, http://www.acls.org/. In February 2003 the ACLS initiated a study of the role of cyberinfrastructure in the humanities sponsored by the Mellon Foundation.*
40. Cal-(IT)2_Project, *Workshop on Cyberinfrastructure Research for Homeland Security, 25-27 Feb. 2003, http://web.calit2.net/RiskReduction*. 2003.
41. NRC, *Preparing for the Revolution: Information Technology and the Future of the Research University.* 2002, Wahington, DC: The National Academies Press, http://www.nap.edu/books/030908640X/html/. 97.
42. Duderstadt, J.J., Atkins, D.E., VanHouweling, D., *Higher education in the digital age: technology issues and strategies for American colleges.* ACE/Praeger series on higher education. 2002, United States: Am Coun on Educ Praeger Greenwood Publ Grp Inc. xi/288.
43. Raschke, C.A., *The digital revolution and the coming of the postmodern university.* 2003, London ; New York: RoutledgeFalmer. x, 129.
44. Holland, B.A., *Characteristics of Engaged Institutions and Sustainable Partnerships and Effective Strategies for Change (http://www.oup.org/researchandpubs/engaged.pdf)*. 2001.
45. Atkins, D.E., *Keeping Academic Libraries at the Center of the University*, in *Digital Libraries: A Vision for the 21st Century*, P. Hodges, et al., Editors. 2003, University of Michigan Library Scholarly Publishing Office, http://www.hti.umich.edu/cgi/t/text/text-idx?type=simple;c=spobooks;cc=spobooks;sid=75bb88ce6fcf8c816736a4857858ba0c;rgn=div1;q1=atkins;view=text;subview=short;sort=occur;idno=bbv9812.0001.001;node=bbv9812.0001.001%3A7: Ann Arbor. p. 58-73.
46. Johnson, R.K., *Instituitonal Repositories: partnering with faculty to enhance scholarly communication (www.dlib.org/dlib/november02/johnson/11johnson.html)*. D-Lib magazine, 2002. **8**(11).
47. National_Academies, *Electronic Scientific, Technical, and Medical Journal Publishing and Its Implications.* 2003, http://www7.nationalacademies.org/cosepup/E-Publishing.html.
48. Olson, G.M., Atkins, D.E., Clauer, R., Finholt, T.A., Jahanian, F., Killeen, T.L., Prakash, A., & Weymouth, T., *The Upper Atmospheric Research Collaboratory (UARC).* Interactions, 1998. **5**(3): p. 48-55.
49. *Space Physics and Aeronomy Research Collaboratory Website, http://intel.si.umich.edu/sparc/.*
50. *National Virtual Observatory website, http://www.us-vo.org/.*
51. *Bugscope Project website, http://bugscope.beckman.uiuc.edu/.*