Annals of Library and Information Studies 52, 4; 2005; 147-153

CYBERINFRASTRUCTURE: AN OPPORTUNITY FOR EDUCATION AND IMPLICATIONS FOR RESEARCH LIBRARIES

M. Natarajan
Scientist
National Institute of Science Communication and Information Resources (NISCAIR)
14, Satsang Vihar Marg
New Delhi – 110 067
E-mail: m_natarajan@hotmail.com

Deals with the role of research libraries and the concept of Cyberinfrastructure with its definitions. The digital library (DL) movement with few examples in Indian Scenario has been explained. The need for human infrastructure due to the advancement of Cyberinfrastructure has been envisaged. The goals and visions are enumerated with building a successful Cyberinfrastructure. Some of the existing examples are given at International level. The activities and uses of Cyberinfrastructure are given for college level educators; undergraduate and graduate students with the motivation for the database developers are highlighted with the future trend in research libraries. Concluded that Cyberinfrastructure has to be taken as a research challenge and advanced Cyberinfrastructure activity must be brought in for better electronic access.

INTRODUCTION

Research libraries are functioning as the primary institutions for access and effective use of recorded knowledge in support of teaching, research and in some cases community service. They also have core competencies and expertise that is highly relevant to cyberinfrastructure and its challenges. The organizations like archives, historical societies, museums and other institutions devoted for improving access to information also support research and learning. Most research libraries are now well into the reality of the digital information age and are moving at various rates and levels of success to cope with the challenges and opportunities. The publication of library and information science data in electronic format and the archiving of that information in large web-accessible relational databases are some of the more significant and consequential changes to have occurred during the last decade due to the impact of information technology (IT).

Cyberinfrastructure is an emerging concept, which redefines science and engineering in colleges and universities. The development and rapid expansion of cyberinfrastructure has produced increasingly data-rich, linked networks, which open up new possibilities with databases and challenges for using data in library science education. But educators seem to be somewhat divided on the utility of using large online databases. Others have invested considerable time and energy in designing ways to integrate these resources into their teaching, but these advances have not been widely disseminated to the larger educational community. For these and a variety of other reasons, cyberinformatic databases are not being widely used by instructors or by students, especially in the form of active learning experiences in higher education [2]. This paper deals the recent digital library movement along with the current emergence of the “cyberinfrastructure movement” and concluded with some of its implications for education and the future form and function of research libraries.

DEFINITIONAL ANALYSIS

"Cyberinfrastructure" is more than just hardware and software, more than bigger computer boxes and wider pipes and wires connecting them. The term was coined by Ruzena Bajscy of National Science Foundation (NSF) to describe the new research environments in which capabilities of the highest level of computing tools that are available to researchers in an interoperable network. It acts as a more comprehensive infrastructure for research and education, based upon distributed but
federated networks of computers, information resources, online instruments and human interfaces. It allows access to complex services as well as raw computing resources through the network and it is required for a knowledge economy. Cyberinfrastructure consists of advanced computational engines ("Supercomputers"), mass storage, high-performance next-generation networks, digital libraries and databases, sensors and effectors, middleware, application frameworks, collaborative tools and services -- all organized to serve individuals, teams and organizations in ways that revolutionize what they can do, how they do it and who participates. It integrates distributed (Grid) computing, high-speed communications and information management capabilities as well as to remote instruments and visualization 'devices into a single coherent entity. It includes computing cycles and broadband networking, massive storage and managed information, as well as leadership on shared standards, middleware and basic applications for scientific computation. It is focused on sharing, efficiency, making greater capabilities available across the science and engineering research communities. It facilitates new applications and collaboration and interoperability across institutions and disciplines [2].

Therefore Cyberinfrastructure is regarded as a socio-technical construct, that is, it comprises technological infrastructure (such as Internet2 and Grid Computing) and institutional infrastructures (institutions, regulations, standards and policy frameworks). Generic names for such cyberinfrastructure-enabled environments include collaboratory, co-laboratory, grid community / network, virtual science community and e-science community [3].

HUMAN INFRASTRUCTURE

The cyberinfrastructure's human infrastructure is a synergistic collaboration of hundreds of researchers, programmers, software developers, tool builders and others who understand the difficulties of developing applications and software for a complex, distributed and dynamic environment. These people are able to work together to develop the software infrastructure, tools and applications of the cyberinfrastructure. They provide the critical human network required to prototype, integrate, harden and nurture ideas from concept to maturity. The personal networks, knowledge and relationships of the human infrastructure take a long time to build and are critical to the usability of the resources. In particular, the advances in science and engineering are the fruit of the many years of cooperation in the national effort to unite computational and computer sciences [4].

THE DIGITAL LIBRARY MOVEMENT

Digital libraries (DLs) have been available for information collections and services for the representation, storage, retrieval and preservation of multimedia information. The people involved in these activities include:

1. The leaders and professionals of existing libraries (including archives and museums);
2. The computer and information technologies research and development communities that pursue digital libraries as an advanced application of distributed computing systems; and
3. A wide range of knowledge-based communities and institutions (information users).

The best research and development in DLs occurs in projects that align mutual self-interest and create joint learning between these three above groups. The cyberinfrastructure movement has revolutionary implications for all the three groups. The growing ubiquity of digital infrastructure means that the bulk of new information is "born digital." Government and private funding have also funded retrospective conversion to digital representation of important collections of physical-analog materials including text, sound, images, motion pictures and even 3-dimensional objects. The core mission of libraries is access to information via physical access, intellectual access and long-term (preservation) access. The DL research has been broad and multidisciplinary including topics such as distributed systems architecture; information federation, access and retrieval of multi-media objects; information economics, agent technologies, metadata structures and their
automated creation, human-factors, evaluation of use, and multilingual issues. The primary entrance to the library is becoming a web portal and thus libraries are paying more and more attention to the quality of their web portals and issues of human-computer interaction.

In India, many institutions / organizations are involved in DL research and development. These initiatives helped to establish even consortium approach among the libraries. Indian National Digital Library in Engineering Science and Technology (INDEST) consortium provides subscription to e-resources for 38 institutions like Indian Institute of Technology (IITs) and Indian Institute of Management (IIMs), Indian Institute of Science and other Government institutions. The E-journals consortium of Council of Scientific and Industrial Research (CSIR) laboratories helps the researchers, scientists working in these laboratories for accessing the e-journals. National Institute of Science Communication and Information Resources (NISCAIR) has been identified as a focal point for e-journals consortium. CSIR has 39 laboratories with different disciplines and specialization in each of their field with own specialized databases. The Grid project of CSIR for accessing these databases by all laboratories will have cyberinfrastructure facilities.

OTHER DL INFRASTRUCTURE INITIATIVES

Vidyanyidhi (Meaning 'Treasure of Knowledge' in Sanskrit) is India's premier DL initiative to facilitate the creation, archiving and accessing of doctoral theses. It is an information infrastructure, a portal of resources, tools and facilities for doctoral research in India. It is envisioned to evolve as a national repository and a consortium for e-theses through participation and partnership with universities, academic institutions and other stakeholders. It enhances access to Indian theses and enlarges the reach and audience for Indian doctoral research works.

Centre for Development of Advanced Computing (C-DAC) developed high performance-computing cluster PARAM Padma for Digital Library Applications and initiative to store and retrieve digitized contents (images, audio and video). They have handled various natures of DL projects like Sarasvati Mahal Library website, Academy of Sanskrit Research, Indian Institute of Astrophysics and Indian Art Preservation Research Project.

Dr Abdul Kalam, President of India, emphasized the importance of web portals and DLs in India, while addressing the inauguration of International Conference on Digital Libraries- 2004 (ICDL-2004). DLs should be user friendly and give equitable access to explicit information, irrespective of place, educational or economic status. These should be integrated to have access of all the digitized material. He stressed the need for the creation of cyberinfrastructure to integrate even Tele-education, Government file storage, Land record storage and voters’ list digital book, which should be considered as a national mission.

Therefore, new academic programs should educate the appropriate professionals to create and manage such environments from a socio-technical perspective – DL architects and digital librarians. Some of the traditional library schools have been transformed into much broader entities and become part of a new information schools movement. But there is still much to be done in the area of national and international library cooperation in federating collections and services that are fundamental to anytime and anyplace information access. There exists functional disaggregation like institutions other than libraries may assume library functions and libraries may assume some of the traditional functions of publishers or bookstores. One of the example is HighWire Press, (highwire.stanford.edu) is a division of the Stanford University Libraries, which produces the online versions of high-impact, peer-reviewed journals and other scholarly content.

GOALS AND VISIONS

The goal of the cyberinfrastructure of libraries is to advance human thought. They are envisaged for more applications, capabilities, and efficiency; reuse and multiple-use of designs; capture of commonality; spread of best practice; archiving interoperability; provision of tools and services; shared facilities and assistance and expertise. The
vision is to build more ubiquitous, comprehensive digital environments that become interactive and functionally complete for research communities in terms of people, data, information, tools and instruments and that operate at unprecedented levels of computational, storage, and data transfer capacity [5].

ELEMENTS OF CYBERINFRASTRUCTURE

Building, operating and using advanced cyberinfrastructure must be done in a systemic context that exploits mutual self-interest and synergy among computer and information, and social science research communities who see it as an object of research, and other ("domain science") research communities who see it as a platform in service of research. The need for highly coordinated, large, and long-term investment in

1. Fundamental research to advance cyberinfrastructure;
2. Development activities to create and evolve the building blocks of advanced operational cyberinfrastructure;
3. Institutions with people and facilities to provide operational support and services; and
4. High-impact applications of advanced cyberinfrastructure in all areas of science and engineering research and allied education.

Cyberinfrastructure encompasses the bottom two layers supporting these activities. First, it provides computational and communication resources, software and services that are shared by the digital science community. Second, it provides repositories of shared data and software that can be appended, accessed and utilized in the course of those applications. “Resources and services” comprise both information technology and human resources [6].

The cyberinfrastructure should include grids of computational centers; comprehensive libraries of digital objects including programs and literature; multidisciplinary, well-curated federated collections of scientific data; thousands of online databases and vast sensor arrays; convenient software toolkits for resource discovery, modeling, and interactive visualization; and the ability to collaborate with physically distributed teams of people using all of these capabilities. This vision requires enduring institutions with highly competent professionals to create and procure robust software, leading-edge hardware, specialized instruments, knowledge management facilities, and appropriate training. The resources may include Archival repositories, Digital Libraries, Computational Resources and Sensor Networks. Furthermore, all these coordinated endeavors support for specific activities like to benefit education, general science awareness and policymaking. There is a need for coordinated participation by academia, private industry, non-government agencies and laboratories, and state, regional and national centers [7].

BUILDING A SUCCESSFUL CYBERINFRASTRUCTURE

Cyberinfrastructure requires an unprecedented coordination of hardware, software, and human infrastructure to meet the needs of science and society. It provides once-in-a-generation opportunity to leverage the extraordinary momentum of the science and technology community to build a global information infrastructure. To achieve this potential, the following principles should be adopted:

- Cyberinfrastructure must incentivize for real cooperation among its participants and for serious software infrastructure development.
- It should provide a full spectrum of resources to be of maximal benefit.
- It must include disciplinary scientists, social scientists, computer scientists, and technologists.
- The scope, goals, structure and budget of the cyberinfrastructure program must be aligned.
- This program must account for the human infrastructure required for development, deployment and operational success.

The following Figure illustrates the cyberinfrastructure structure:
The foundation of the infrastructure consists of the libraries, archives and museums that preserve information; the bibliographies, finding aids, citation systems, and concordances that make that information retrievable; the journals and university presses that distribute the information; and the editors, librarians, archivists, and curators who link the operation of this structure to the scholars who use it. Above the layer are software programs, services, instruments, data, information, knowledge, and social practices applicable to specific projects, disciplines, and communities of practice. Between these two layers is the cyberinfrastructure layer of enabling hardware, algorithms, software, communications, institutions and personnel. This layer should provide an effective and efficient platform for the empowerment of specific communities of researchers and for the users [8].

Grid and collaboratory environments built on cyberinfrastructure can enable people to work routinely with colleagues at distant institutions, and with junior scientists and students as genuine peers, despite differences in age, experience, race, or physical ability. These new environments can contribute to science and engineering education by providing interesting resources, exciting experiences and expert mentoring to students, faculty and teachers anywhere. The new tools, resources, human capacity building and organizational structures emerging from these activities will also eventually have even broader beneficial impact on the future of education at all levels and likely on all types of educational institutions. A vast opportunity exists for creating new research environments based upon cyberinfrastructure, but there are also significant risks and costs if not done quickly and at a sufficient level of investment [9].

**SOME OF THE EXAMPLES AT INTERNATIONAL LEVEL**

a) The National Science Foundation (NSF) program has more than a decade of hands-on experience, developing and using computational and data management infrastructure to enable new generations of scientific advances. It has the potential to change the direction and evolution of science and technology. The program’s design, goals, mission and structure have a fundamental impact on our scientific culture.

b) InfoVis Cyberinfrastructure provides access to a comprehensive set of software packages easing the exploration, modification, comparison and extension of data mining and information visualization algorithms. Diverse software packages were bundled into learning modules. Access to a large-scale data repository, extensive computer resources and a growing set of references are provided as well [10].

c) The National Virtual Observatory (NVO) - George Djorgovski presented his experiences and impressions in a discussion entitled “Virtual Observatory, Cyber-Science, and the Rebirth of Libraries” about cyberinfrastructure. He described the NVO as “a complete, dynamic, distributed, open research environment for the new astronomy with massive and complex data sets.” [11].

d) The Digital Library for Earth System Education (DLESE) is a distributed community effort involving educators, students and scientists working together to improve the quality, quantity and efficiency of teaching and learning about the Earth system at all levels [12].

e) American Council of Learned Societies (ACLS) established a national commission, to investigate the cyberinfrastructure needs of scholars in the humanities and social sciences. It allows scholars to focus their intellectual and scholarly energies on...
the issues that engage them and be effective users on new media and new technologies [13].

f) The Australian Government is providing $542 million over 2004-05 to 2010-11 to continue to provide researchers with access to major infrastructure, link infrastructure funding more directly to Australia’s National Research Priorities and foster greater research collaboration and the collaborative use of infrastructure.

ACTIVITIES FROM THE CYBERINFRASTRUCTURE

Research libraries need to explore new models of cooperation to create common federated collections and functional cooperation with other organizations in their day-to-day activities for:

- Online access to complete literature coverage of the credentialed, archival literature of the relevant to various research communities;
- Stewardship and curation services for enormous collections of scientific data;
- Providing stewardship and access to instructional material and the intermediate products (pre-publication products) of research activities;
- Leadership in digitizing, organizing and curating unique special collections belonging to the library and providing access to them;
- Providing continuous and open forms of scholarly communication.
- Current awareness information services and community specific information portals (“virtual special libraries”) for individuals, projects and/or research organizations [14].

USES OF CYBERINFRASTRUCTURE

It helps the college-level educators

- To locate potentially-useful large digital databases, highlighting differing strengths, features, ease of use and potential educational applications
- To download, contribute and comment on examples of how cyberinfrastructure can be applied in a variety of educational settings and activities (e.g., lectures, problem sets, in-class exercises, lab exercises, student research projects, etc.)

It enables the undergraduate and graduate students to locate and navigate selected online databases, as well as evaluate, screen, and manipulate various types of data available from these sites, many of which were not originally designed with student needs and abilities in mind.

Finally, this motivates the database developers

- To consider and plan for educational applications and requirements as they expand current databases and create the cyberinfrastructure of the future.
- Provides the underpinnings for a revolution in research and education, rich, diverse datasets combined with analysis tools, arrayed in easily accessible formats.
- Facilitate a paradigm shift in documenting and preserving observational and experimental data independent of inference and interpretation.
- Enable and enhance process-oriented research by providing a rapid and flexible means of extracting and analyzing data required for the development of new, testable, end-member hypotheses, and
- Revolutionize Library science education by the early integration of students into scientific inquiry via direct access to primary data and providing analysis tools [15].

ISSUES INVOLVED IN CYBERINFRASTRUCTURE

A vast opportunity exists for creating new research environments based upon cyberinfrastructure, but
there are also real dangers of disappointing results and wasted investment for a variety of reasons including under funding in amount and duration, lack of understanding of technological futures, excessively redundant activities between science fields or between science fields and industry, lack of appreciation of social/cultural barriers, lack of appropriate organizational structures, inadequate related educational activities and increased technological balkanizations rather than interoperability among multiple disciplines. But human resources are critical to making cyberinfrastructure and applications work, keeping them working and providing user support.

FUTURE

Advances in computational technology continue to transform scientific and engineering research, practice and allied education. Research libraries for the 21st century should be in the era of information abundance and complexity, play major roles / functionality on - data discovery services, data provider federators, primary and / or derived data archivers, domain expertise, quality control, relationship with web portals and search engines and requires new type of scientific management and organization structures.

CONCLUSION

The cyberinfrastructure presents a series of research challenges. The complexity of managing the massive scale of heterogeneous computation, communication and storage resources needed to achieve this future networked environment. An advanced cyberinfrastructure promises to revolutionize the conduct of science and engineering research and education in the 21st century. To achieve this, databases and analysis tools must be combined in an information system developed at disciplinary levels and made available to the national scientific community. To reach critical mass, an advanced cyberinfrastructure activity would require interagency partnerships as well as collaboration between the physical and life sciences, computer science and the social sciences. The opportunity is evidenced by both progress from the developments in information technology and the mushrooming of cyberinfrastructure projects for specific fields, initiated by scientists in those fields. It is not seen as an end in and of itself, but, rather, as being important in that it can enable new research in science and engineering. According to Kim, “a billion dollars invested in cyberinfrastructure may well result in ten trillion dollars in economic growth.”

ACKNOWLEDGEMENT

The author is grateful to the Director, NISCAIR and Head; ETTG for their encouragement and support for submitting this article.

REFERENCES

11. DIANE(Goldenberg-Hart). op. cit 5
12. www.diese.org
14. ibid.
15. ATKINS (Daniel). op. cit. 2