

Next Generation Internet and the Global Digital Library

Ching-chih Chen

Professor

Graduate School of Library and Information Science

Simmons College, Boston, MA 02165, USA

It is indeed a great privilege to be with you all. I wish to thank Dr. Simon Lin for inviting me to take part in this exciting meeting. Some of you are aware of my international work in more than two dozen countries, and since 1987, I have devoted some of my time in bringing informational professionals from both developing and developed countries together by offering my own non-profit International Conference on New Information Technology, NIT with shoe-string budget. This started in 1987 in Bangkok, and have covered many geographical areas, including Singapore, Guadalajara (Mexico), Budapest, Hong Kong, Puerto Rico, Riga (Latvia), Alexandria (VA, USA), Pretoria, South Africa, and the latest one, the 10th NIT, in Hanoi, Vietnam in March 1998. Although during the last minutes, the meeting was postponed but the informal meetings went on as usual and we were very successful in fulfilling our goals. I am delighted that two colleagues from the Institute of Sino-nam in Hanoi, Vietnam are able to be with us today to enjoy the Academia Sinica's wonderful hospitality because of NIT. So you can see, that I have long been involved in the kind of activities very much sharing the same spirit of the PNC meeting.

Therefore, I am delighted to capture this wonderful moment to explore with you the exciting and changing time we are in, and how important it is for us to know what to do to prepare ourselves for the next century. If we can hold fast to our guiding principles on information access and provision, and know how to use the available technologies effectively, we can make this time of change a moment of dazzling opportunity for all information professionals.

In the last decade, the technological, social and economic changes have been particularly dramatic. Technologically, with the advent of microcomputers, optical discs and other mass storage media, telecommunication, digital imaging, computer graphic, multimedia, compression, etc... have dramatically changed the way we live, think, and communicate with each other, and certainly the way we use and view technologies.

In the most recent few years, the development in communications technology has been so dramatic that we are truly experiencing the incredible power of the open system, Internet. In fact, there is little historical precedent for the swift and dramatic growth of the Internet. In just a few short years, the Internet had evolved from a somewhat esoteric phenomenon to mainstream reality, and from what was a limited scientific communication network developed by the U.S. Government to facilitate cooperation among Federal researchers and the university research community, to a ubiquitous communication and information access tool across all levels of educational institutions, and a popular global open system, which has changed the way we work, we function, we learn, and we communicate.

With all these unprecedented changes, it should not be surprising that there is an increasing demand for better access to needed global information to enable us to have a bigger picture on the world in which we are living in a better global view on our environment, our history, our cultures, our economy, our science and technology, etc... Thus, information has become

the key to productivity, and there is a shift toward a knowledge-based learning-oriented "creative society." In this type of society, we are witnessing the following changes in emphasis:

- Societal values change from "acquiring" to "learning"
- Growing motivation of individuals for knowledge
- More people learn to use information creatively
- More demand for multimedia information
- More demand for global information

Knowing that many excellent papers at this conference are on the innovative use of this generation Internet for better information access and service provision, I have chosen to speak more on the Next Generation Internet and the changing role of libraries in this digital knowledge society.

As a member of President Clinton's Advisory Committee on High Performance Computing and Communications, Information Technology, and Next Generation Internet, I have been enormously privileged to work with a group of high-power IT leaders on the committee and have kept more up-to-date with the US government's new direction in this area of developments. Let me share with you first a short 5-minute video called "*Advanced Networking: Connecting to the Future*" (1) which was shown at the Netamorphosis, a demonstration on the potentials of the NGI in March 1998.

[Showing of *Advanced Networking: Connecting to the Future*]

Clearly we can begin to imagine what the future will hold for us for document delivery, for education, for home and schools, etc... We can begin to imagine the vision of the Next Generation Internet.

NEXT GENERATION INTERNET (NGI) INITIATIVE

The NGI Vision

In the 21st century, the Internet will provide a powerful and versatile environment for business, education, culture, and entertainment. "Sight, sound, and even touch will be integrated through powerful computers, displays, and networks. People will use this environment to work, study, bank, shop, entertain, and visit with each other. Whether at the office, at home, or traveling, the environment and its interface will be largely the same. Security, reliability, and privacy will be built in. Customers will be able to choose among different levels of service at varying price points. Benefits of this dramatically different environment will include a more agile economy, improved health care - particularly in rural areas, less stress on the ecosystem, easy access to life-long and distance learning, a greater choice of places to live and work, and more opportunities to participate in the community, the Nation, and the world" (2).

The NGI Goals

To make this vision possible, President Clinton and Vice President Gore announced on October 10, 1996 the Next Generation Internet (NGI) Initiative, which will accelerate the

introduction of new networking services for our businesses, schools, and homes with the following specific goals (3):

NGI's first goal is to research, develop and experiment with advanced network technologies that will provide dependability, diversity in classes of service, security, and real-time capability for such applications as wide-area distributed computing, teleoperation and remote control of experimental facilities. These activities focus on network growth engineering, end-to-end QoS, and security.

Accompanying the development of advanced network technologies is NGI's second goal, development of the next generation network fabric. This effort will overcome today's speed bumps slowing end-to-end usable connectivity caused by incompatibilities in switches, routers, local area networks, and workstations. Two thrusts within this goal are planned: First, construction of a high-performance distributed laboratory consisting of the 100 NGI sites at universities, Federal research institutions, and other research partners at speeds in excess of 100 times today's Internet. This laboratory will be large enough to provide a full-system, proof-of-concept testbed for hardware, software, protocols, security, and network management required by the commercial NGI. Second, development of ultra-high speed switching and transmission technologies and end-to-end network connectivity at more than one gigabit per second. Such networks will be high-risk, pioneering networks limited to 10 NGI sites at speeds 1,000 times faster than today's Internet. The network fabric for NGI is vBNS for the moment.

These two goals—experimental research of advanced network technologies and development of the next-generation network fabric—will provide the basis for terabit per second networks operated by appropriate network management and control providing guaranteed end-to-end QoS.

Finally, Goal 3 to test the advanced capabilities of the first two goals. It will demonstrate a selected number of applications requiring these capabilities over the NGI network(s). Procedures have been established to ensure that selected applications provide robust, realistic, complete tests of technologies that can be extended and adapted to other applications. Initial applications are being chosen from the federally focused applications in appropriate technology classes, such as digital libraries, remote operation of medicine, and crisis management (2).

Virtual environments (VEs) will also be developed. These are applications of potentially wide-spread utility in medicine, design, collaboration, and training that demand high bandwidth, low latency communications, and have the potential to exploit vast amounts of computational and storage capabilities. (3).

This multi-agency NGI initiative—a solid partnership with industry, academia, government and the American public—provides the catalyst for the development of high-performance, secure, reliable networks of the future and holds great future promises.

Internet2

The NGI initiative is closely related to Internet2, a collaborative effort by more than 100 U.S. research universities to create and sustain a leading edge network capability enabling the creation of the broadband applications, and education. While the goals of the NGI and Internet2 are complementary and interdependent, they are clearly distinct. The NGI initiative is a Federal mission-driven R&D program, while Internet2 focuses on innovation in academic research and education applications. The NGI initiative will create an experimental, wide area, scalable testbed to develop mission-critical applications; Internet2 will meet end-to-end performance requirements by developing and deploying advanced network infrastructure. Much of the wide area testbed for Internet2 will be provided by the NGI Initiative. Both the NGI initiative and the Internet2 project will develop and test advanced network technologies not supported by today's Internet, primarily through NGI-funded research at Internet2 universities. Continued strong coordination and communication between the Federal and academic communities will be crucial for the success of both programs (2).

DIGITAL LIBRARY DEVELOPMENT

At the present time, we do not have a clear "theory" of digital libraries. The term, "digital library," has been used rather loosely. Although simple definitions have been given by many, they are generally not well defined. For example, Lesk called a "digital library" "a collection of information that is both digitized and organized" (4). This can have so many different interpretations ranging from a digital collection of digitized books, journals, reports, or photo collections to something more complex ones in network-based collaborative knowledge environment. Yet, they all share some common characteristics -- i.e., a digital library can be searched, accessed all over the world, and copied without error (4). In order to do this, a digital library must have digital CONTENT and this is a very labor-intensive activity. In addition, digital content must be properly stored, organized, so that it can be found or retrieved. This is why so many current "digital library" projects have been heavily involved in making the library collection digital or bringing together available digital files so that information can be accessed and shared via the network. These are "digital library" projects are mainly using this generation Internet.

The "digital library" applications solicited for the NGI initiative will not be concentrating in the baseline efforts for scanning and digitization. Instead, as articulated by the National Science Foundation (NSF)'s Human Centered Systems (HuCS) program, effort will be made to supports university-led research in the development of advanced methods for collecting, storing, and organizing information in digital form for network access; and advances in multi-agency supported basic research on multimodal interaction with computing systems, including speech, text, image, and multimedia advanced technology. As of today, NSF has already supported 6 major "digital library" projects and the FY 1998 NSF's HuCS R&D *Digital Library Initiative - Phase 2* areas include the continuing development of knowledge repositories and information agents that sort, analyze, and present massive amount of multimedia and multi-source information; collaboratories that provide access to knowledge repositories and facilitate knowledge sharing, group authorship, and control of remote instruments; systems that enable multi-modal human-system interaction including speech, touch, and gesture recognition and synthesis; and virtual reality environments and their use in scientific research, health care, manufacturing, and training (5).

Digital Library Initiative - Phase 2

Digital Library Initiative - Phase 2's research areas will address the many issues identified in the Santa Fe Planning Workshop on Distributed Knowledge Work Environments (6, 7):

- Human-centered research - this seeks to further understanding of the impacts and potential of digital libraries to enhance human activities in creating, seeking, and using information and to promote technical research designed to achieve this: The issues include:

- legal issues which arise with respect to copyright, intellectual property, privacy and confidentiality, personal and business equity, and security;
- the differences in culture, especially as reflected in the means for communication;
- generational gaps;
- developing multilingual, multiscriptural, and multicultural interfaces
- intelligent user interfaces
- Economic and social implications,
- Etc.

- Content and collection-centered research -- This will focus on better understanding of and advancing access to novel digital content and collections. The issues to be explored will include structuring, indexing, and metadata provision for collection; gathering, validating, and authorizing data; security and integrity of collection; preservation of collection; interoperability of content and collections; domain-specific information objects, etc.

- System-centered research -- This will focus on component technologies and integration to realize information environments that are dynamic and flexible; responsive at the level of individual, group, and institution; and capable of adapting large, amorphous, continually growing bodies of data to user-defined structure and scale. Many issues will include:

- scalability
- interoperability -- these issues include system, syntactic, semantic, linguistic, temporal interoperability, as well as integration of diversity of hardware and software.
- adaptability and durability -- ability to deal with multimedia representation, dynamic information structuring, information identification, legacy systems, deployment of metadata, standards, provenance, etc...
- Support for collaboration
- system evaluation and performance studies.

GLOBAL DIGITAL LIBRARY (GDL) AND UNIVERSAL INFORMATION ACCESS

In recent years, we have passed several "information jumps" -- from speech to writing to printing, and now to wire and wireless communications. Now, this gigantic open system -- the Internet -- has offered tremendous opportunities to library professionals, but it also has many inherent problems, such as traffic congestion, the lack of bandwidth, quality of services, etc... As the NGI develops, current Internet problems will have been addressed and what lies ahead is a real jump for us toward universal information access. This means that technologically, anyone, anywhere, could talk, write, confer with, or send multimedia -- textual, audio, and visual -- digital information to anyone else in any part of the world! So, the concept of the digital "Global Library," which I have been advocating in the last decade,

is not only conceptually sound but also technologically feasible now. Many years back when I was privileged to be a keynote speaker at the International Conference celebrating the 60th anniversary of the National Central Library in Taipei, I presented a very modest scheme (Figure 1) on the potential of a global digital library long before Internet came to be part of our lives, and now the time has come for us to realize this global digital library. Yet, we still need to pass many hurdles before we can reach this goal because there are still many barriers (8, 9). Some of the major issues have already been included in the previous section.

Figure 1. The early GDL conceptual scheme

THE GLOBAL DIGITAL LIBRARY (GDL) PROTOTYPE

Currently with the easy and widespread use of the Internet and WWW, there are thousands and thousands of websites available in the cyberspace. In this big mess of digital ocean intermingled with multiple kinds of information -- commercial, entertainment, educational, and others -- repositories of digital libraries with knowledge-based contents do exist. Yet, these repositories are certainly of very small minority, and thus not only precise and quick retrieval of them is not always easy since the current web-based retrieval capability is far from ideal but they are also largely lost in this big digital ocean. Even when it is retrievable, each site differs from the other in terms of contents, depth of knowledge presented, style of presentation and user interface. As a result, users are generally confused and perplexed. Thus, one of the key challenges is to develop a digital library structure which can provide a coherent, consistent view of as many of these repositories of the similar nature as possible, and yet, it is still possible to retain the diversity of each site which can be responsive to unique constituencies. The current GDL prototype is such an attempt, which links the repository information of various types of educational institutions -- such as national libraries, national archives, major museums, networks, and research/academic libraries -- together in one single global digital library system with a coherent and consistent interface. In a way, these digital knowledge bases have been filtered out from the big digital ocean, and be placed in a system for easy information access.

Figures 2 and 3 shows how Library of Congress' digital libraries can be easily accessed by entering the GDL system and click on the Library of Congress graphic icon (one of the three national libraries of the U.S.) Similarly, a user can access the digital libraries of the major academic and research institutions like that of the Academia Sinica by choosing Taiwan from Asia, and then click on the Academia Sinica's graphic icon, as shown in Figures 4 and 5.

Figure 2. The GDL Page of the American National Libraries

Figure 3. Accessing the Contents Available at the US Library of Congress

Figure 4. Accessing the Digital Libraries of the Academia Sinica

Browsing in the GDL can show that the depth of contents for each major library in the world differs greatly from strictly directional type of homepage information to the substantial digital library collections such as those of the Library of Congress of the U.S.A. (many countries' national libraries do not have even a web site). Thus, the need for major libraries to build digital knowledge-based contents is obvious.

Once the digital contents are available, the global information will practically be at an individual's fingertips, the educational process will be freed to focus on how to access and organize this information, rather than on the acquisition of "facts." Information technology will allow students and teachers, as well as domain experts, to author and publish their own educational materials. Multimedia technologies will yield applications with a whole range of new and more accessible functions, likely to have major impacts on education as well as entertainment. The emerging Information Infrastructure will deliver education content to school, to work, and to the home, for self-directed study, collaboration, and easy access to specific expertise. In conjunction with radical rethinking of the associated curricula, information technology can enable "just-in-time learning," providing critical information when the learner needs it (10).

The global digital library (GDL) can involve in most of these activities which happen in the Internet today and the NGI. This means that the GDL could be anything. The CONS and PROS of the Internet/Cyber-World will also have impact on the GDL as well, thus many researchable areas can be immediately identified. Many research areas identified earlier for Digital Library Initiative - Phase II do apply, and some of the most immediate ones can be:

- Data Transport: The basic function and all services will rely on it at the beginning. In this area, the research can involve:
 - QoS: Scalability model and end-to-end bandwidth and latency guarantees on the Internet.
 - High Speed for multimedia information
- Security and intellectual property issues
- Roaming - How to extend the current GDL network beyond its own infrastructure, and how to offer the GDL users the ability to log in with other DL data/multimedia-bases served by roamers. How to design an authorization that takes place in the background so that users require no additional software.
- Billing system for those sites which will require fees for services.

Clearly the GDL can facilitate the transforming potential of "information access" and "information creation," can eliminate the barriers between users and information providers, can promote broad-based access, and can address critical public policy issues -- remember that the educational technology is considerably more than connectivity.

CONCLUSION

While it is certain that we will have the powerful information superhighway, we are far from having our information resources available in digital form so that they can be linked together by utilizing the available technologies. Being available in digital format will be the first requirement. Once the information sources are available in digital form, the central concern to the endusers will have to be the authority and quality of content in digital libraries. Take the national libraries for example, the GDL prototype experiment has revealed vividly that

although there are substantial numbers of national libraries which have homepages available, most do not have knowledge-based contents. Even few content-based information are available from national museums and archives. Thus, this suggests that there is a desperate need for quality content building.

Currently we have used the words *cruising* and *surfing* to describe our behavior on the Web, few have invoked the words *learning* or *engaging* when we browse or use the Web, as pointed out by Niclolas Negrosponte. But, how can we possibly learn and engage when the type of digital information available is not of high quality or not knowledge-based? Also, even they are available, but if it is not accessible, how can one engage?

The Web is a digital landmark. It can create a new and more accessible subworld, like a window-shopping experience. But, the Internet is now like a city - people can go places, and can visit communities. But when they arrive at a place and try to make things happen, they often end up frustrated. According to Negrosponte, at the turn of the millennium, we will find a billion users on the Net, i.e., roaming programs, not people.

Being at these cross roads, in addition to speculation on the libraries in the next millennium, we must make sure that we can develop in this seemingly exciting networked environment, a vision for our global library's future, and define its role in facing a new frontier. It is important for us to visualize that not only all types of libraries in our country would be connected to the super-network, but globally all libraries would be part of the network as well. In anticipating the growing demand to use the Net and the Web for more suitable purposes: communicating, learning, experiencing, the GDL prototype has been created to provide users with a window-shopping experience on the world's rich information resources. In other words, the one-stop-shop place for people to access the world's rich information resources. But we must have the real thing so that the library Net and Web users will not end up frustrated! This is a real challenge!

Despite of potential difficulties, barriers, and challenges, one this is sure that the technologies and the infrastructure are in place now for us to experiment an universal global library. With this, the role of academic librarians for the 21st century will unquestionably change substantially. For the first time ever, lack of proper technology is no longer an obstacle. But, technology is not the end in itself rather the means to an end. We should not suffer from the loss of direction caused by preoccupation with ever-changing technology. As the academic librarians speculate on their work in the next millennium, what we must do is to make sure we can develop a vision for our library's future in this Internet-enabled society, and define its role in facing a new frontier. But, the greater challenge for us is to understand the evolutionary path and the transitional strategies that will be necessary to guide us from where we are today to any Next Generation Internet-related vision of tomorrow (11).

"What makes this point in human history so interesting is that we are all -- people, organizations, and governments -- struggling with the vast changes in our processes and procedures that the Internet is bringing about, and will continue to bring about. The challenge is not to predict the end point of the change -- namely, what the Internet-enabled future will be -- but rather to better understand what we must do today and tomorrow, no matter what the Internet becomes or brings about... What people and institutions can do to cope with the changes the Internet will bring, however it emerges" (11). This is a big order!

REFERENCES

1. "Advanced Networking: Connecting to the Future". A 5.45-minute video produced by the University Corporation for Advanced Internet Development (an Internet2 group) and the U.S. National Coordination Office for Computing, Information, and Communications for the *Netamorphosis - Demonstration of the Next Generation Internet* - on March 11-13, 1998 in Washington DC.
2. U.S. National Coordination office for High Performance Computing and Communications. *Technologies for the 21st Century: Computing, Information, and Communications (Supplement to the President's FY 1998 Budget)*. Washington DC: NCO, 1998 pp.3-4.
3. U.S. National Coordination office for High Performance Computing and Communications. *America in the Age of Information: A Forum [A Report to the Committee on Information and Communications on the July 1995 Workshop]*. Washington DC: NCO, 1995. p.30.
4. Lesk, Michael. *Practical Digital Libraries: Books, Bytes, and Bucks*. San Francisco, CA: Morgan Kaufmann Publishers, 1997.
5. U.S. National Coordination office for High Performance Computing and Communications. *Technologies for the 21st Century: Computing, Information, and Communications (Supplement to the President's FY 1998 Budget)*. Washington DC: NCO, 1998 pp. 4-5.
6. *Digital Libraries: Report of the Santa Fe Planning Workshop on Distributed Knowledge Work Environments*, March 9-11, 1997, Santa Fe, New Mexico. Ann Arbor: University of Michigan.
7. U.S. National Science Foundation. *Digital Libraries Initiative - Phase 2. Announcement*. 1998.
8. Chen, Ching-chih, "Global digital library initiative: Prototype development and needs," *Microcomputers for Information Management*, 13 (2): 133-164.
9. Lynch, Clifford, and Garcia-Molina, Hector, "Interoperability, scaling, and the digital libraries research agenda," *Microcomputers for Information Management*, 13 (2): 85-132 (1996).
10. U.S. National Coordination office for High Performance Computing and Communications. *America in the Age of Information: A Forum [A Report to the Committee on Information and Communications on the July 1995 Workshop]*. Washington DC: NCO, 1995. pp. 36-40.
11. Institute for Information Studies. *The Emerging Internet*. Queenstown, MD: IIS, 1998. p. xvi, xix.

