From educational multimedia to "global digital library" development: Convergence of technology, content, and international collaboration

Ching-chih Chen
Graduate School of Library and Information Science
Simmons College
300 The Fenway
Boston, MA 02115
chen@simmons.edu

Abstract: Multimedia technology has shown its promises for educational applications at all levels of academic institutions as well as business settings for sometime, long before the Internet and World Wide Web have become commonplace. Then, in recent years, particularly the last half decade, we have witnessed how the Internet has changed our world in every conceivable aspect -- everything from how we communicate, learn, conduct research, deal with information, do business, receive health care, and dealing with our governments at all levels. Thus, desktop education has been extended to networked-based learning and education.

As to information provision, academic institutions have long been advocating the importance of their libraries and library collections. The old model is "the larger the collection and the closer it is to the users the better." But now in the Internet era, this powerful global information network has provided the kind of infrastructure and environment natural for international collaboration if only the desired information resources are "digital." Thus, the new emphases have shifted from "owning" to "sharing," from "physical libraries" to "virtual libraries," from one central library with huge collections to the federated distributed digital library systems all over the world… With the exciting convergence of technology, content, and global collaboration, we can visualize the coming of a kind of "global digital library" unattainable in non-networked environment.

In this paper, the author will share some of her own experiences in all these areas and transformations -- from the creation of the award-winning interactive videodisc and multimedia CD-ROM, called *The First Emperor of China*, in mid- and late-1980s, to her current international digital library project, called *Chinese Memory Net*, supported by the US National Science Foundation, as well as the latest China-US Million Book Digital Library Project.

Introduction

Long before the Internet and World Wide Web (WWW), multimedia technology has already been effectively utilized to develop powerful educational applications at many levels of educational institutions as well as in business settings. As the Internet and WWW proliferate to such an extend, they have truly changed the way we communicate, teach, learn, conduct research, do business, deal with government, etc. The fast- and broadband global network provides a kind of environment natural for globally collaborative learning and research. Thus, much of the existing multimedia applications have also been transformed mostly for network-based learning.

With the proliferation of digital information resources on the web as well as the growing popularity in creating digital information, as well as in converting print and non-print resources to the "digital" ones, we have sensed the fast growing interests in developing "digital libraries," "virtual libraries," and the like everywhere.

In the following few pages, this author was asked to briefly discuss these developments and transformations as they relate to her own experiences. Her projects will be used to demonstrate this exciting time when technology, content, and international collaboration are truly converging to enable us to provide access to needed human knowledge in such a way far beyond our imagination even just a couple years ago.

Vision for Universal Information Access

Prior to this digital age, the model for better information provision and services is to create larger collections at its own sites. This is why academic institutions value the large number of books and journals available in their own

libraries. But now in the Internet era, this powerful global information network has provided the kind of infrastructure and environment natural for international collaboration if only the desired information resources are "digital." Thus, the new emphases have shifted from "owning" to "sharing," from "physical libraries" to "virtual libraries," from one central library with huge collections to the federated distributed digital library systems all over the world…

This is the time we can begin to entertain a bold vision for universally accessible collections of human knowledge as advanced by the President's Information Technology Advisory Committee's (PITAC's) Panel on Digital Libraries (February 2001):

> **"***All citizens anywhere anytime can use any Internet-connected digital device to search all of human knowledge. Via the Internet, they can access knowledge in digital collections created by traditional libraries, museums, archives, universities, government agencies, specialized organizations, and even individuals around the world. These new libraries offer digital versions of traditional library, museum, and archive holdings, including text, documents, video sound, and images. But they also provide powerful new technological capabilities that enable users to refine their inquiries, analyze the results, and change the form of the information to interact with it…*
>
> *"Very-high-speed networks enable groups of digital library users to work collaboratively, communicate with each other about their findings, and use simulation environments, remote scientific instruments, and streaming audio and video. No matter where the digital information resides physically, sophisticated search software can find it and present it to the user. In this vision, no classroom, group, or person is ever isolated from the world's greatest knowledge resources."*

The Convergence of Technology, Content, and Collaboration

"Anytime, anywhere, on any device" stands strong this year as many groups' most-quoted slogan, including those of the digital libraries and networks, as well as those in commerce. Sure, this ambitious and crafty near-alliteration can only become a phrase of reality with the continued developments in pervasive computing. These include not only very-high-speed networks, but also the wireless ones. In fact, when we say "any internet-connected digital device" and "anywhere," we are or will be definitely pointing to the wireless device. Wireless LANs are maturing, offering faster speeds and interoperability with a multitude of devices. From the technology side, it seems easy to see what we need and where we are going.

But, the real challenge is not relating to "technology." When we are referring to "all of human knowledge," we are surely pointing to the enormously huge "digital contents," much more so than any current data centers, or digital libraries can possibly begin to offer. While technology is maturing rather quickly, the bigger challenge is the creation of the large quantity of high-quality digital contents. Take libraries as examples, this is why we are seeing all kind of digital library initiatives ranging from the digitization of specific collections of invaluable resources to elaborate regional, state, national, or international collaborate efforts devoted to creating digital contents, developing innovative technological tools, and introducing exciting new ways of information retrieval and service provision. Similar activities are going on in the museum and archival worlds globally.

In order to create high-quality digital contents, we need to carefully select the potential candidates for digitization. Then, we need to have proper technological tools to design, manage, preserve, and retrieve the digital contents. Yet, no one institution can possibly possess all the valuable contents of a given topic, thus, collaboration becomes essential. While in old days, and even now in its traditional ways, libraries learn to cooperate via "interlibrary loan" to share their resources. In this current digital age with its high-speed networks, it becomes a reality for institutions or individuals to collaborate regardless the distance between them. In fact, one can collaborate just as easily with collaborators in the other side of the globe as those next door to his/her office.

With this kind of exciting convergence of technology, content, and global collaboration, we can visualize the coming of a kind of "global digital library" positively unattainable in non-networked environment.

From Multimedia to Collaborative Global Digital Library Development

In the last 15 years, this author has engaged herself in a few major federally supported research and development projects. The experience gained through this exciting journey from multimedia to collaborative global

digital library development will be briefly described in the following while more detailed information on the current *Chinese Memory Net (CMNet)* can be found in Chen (2001):

Educational Multimedia and *The First Emperor of China*

In 1984, when the Internet was still unknown to most people, and CD-ROM was not yet introduced to libraries, because of the high cost, with the help of industry such as SONY, the Library of Congress was probably the only library in the US could have videodisc technology applications. Recognizing then the potentials of the interactive videodisc technologies, I embarked on a major project, PROJECT EMPEROR-I, supported by the US National Endowment for the Humanities. This project demonstrated that this interactive videodisc technology could indeed help to overcome the barriers of time and distance by enabling educators, researchers, students, as well as the general public to gain multimedia access to The First Emperor of China's fascinating 7,000 terracotta figures of warriors and horses. In early 1990s, the project effort was expanded to use cutting-edge multimedia technology. As a result, the Voyager Company published two award-winning products (interactive videodisc and multimedia CD-ROM) in 1991 and 1993 respectively (http://voyager.learntech.com).

With numerous user-friendly multimedia courseware that provide flexibly links to needed texts, audios, still images, and videos (sample demonstration will be made during presentation, see also Figure 1 for the Table of Contents screen, and Figure 2 for sample screens); as well as interviews, the products have been widely used in various educational settings as well as by general public. As shown in Figure 1, the features include:

- Textual commentary from experts in the field of Chinese history and archaeology
- Bilingual English/Chinese soundtrack
- Dazzling photography of hundreds of life-size terra cotta warriors and horses
- Original film footage of the excavation in QuickTime video
- An "Image Index" with essays linked to photographs
- A thumbnail index to all photographs
- Reference tools, including maps, a timeline, an extensive bibliography, and a glossary with Chinese characters and audio of Chinese pronunciation

The products also received many honors and awards. Although well received by their users, they are created for use with desktop computer facilities only, and cannot be shared over the network. The interactive videodisc is an analog product, while the multimedia CD is not web-compatible.

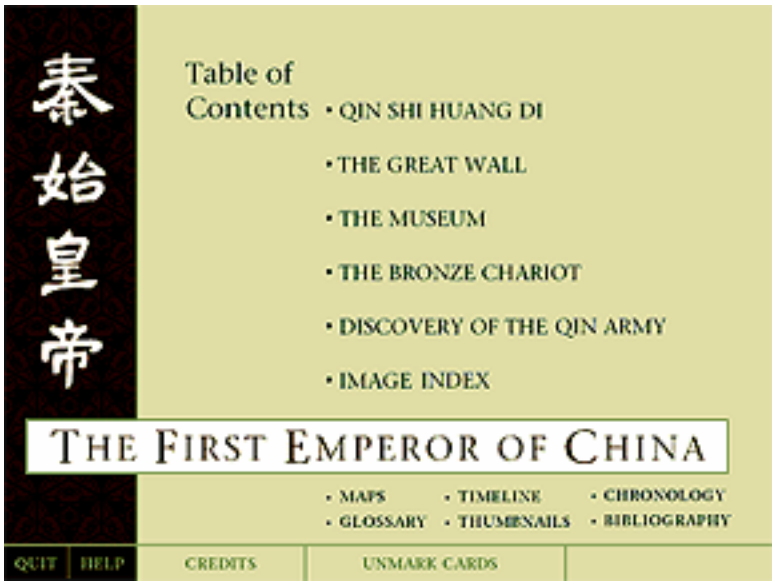

Figure 1: "Table of Contents" screen of *The First Emperor of China*'s multimedia CD

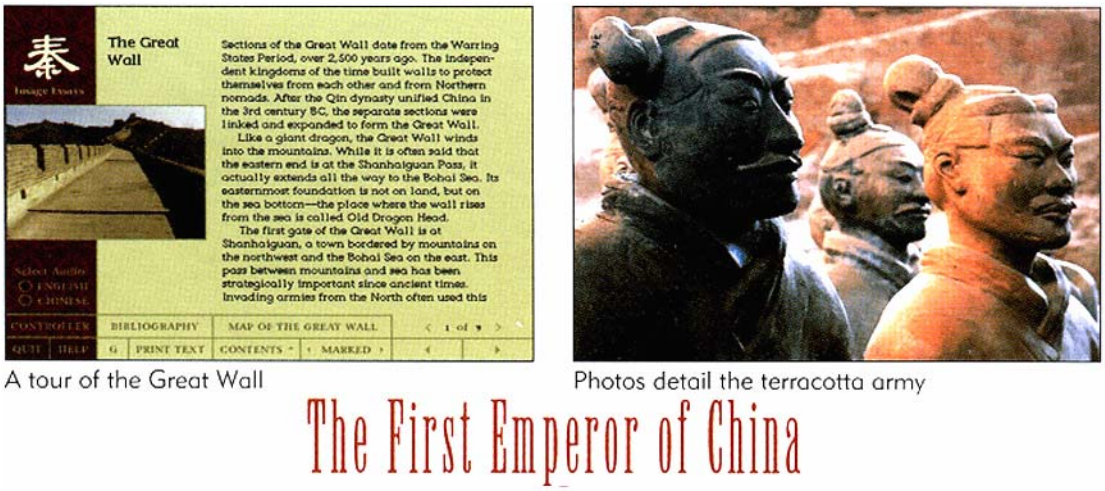

Figure 2: Sample screens from *The First Emperor of China*'s multimedia CD

With the advent of communications and information technologies in the early 1990s, and the anticipation of the coming of the Internet, it became clear to me that in order to promote better information sharing in the coming "digital" age, there is a real need to work toward a "digital knowledge base." Again, *The First Emperor of China was* a perfect example used to prepare several feature articles including that in the 1994 issue of IBM's *Multimedia Today* (Figure 3) (Chen, 1994) to advocate such a concept.

Figure 3: Images of the first two-pages of the featured article on "digital knowledge base"

The power of current technology supports sophisticated systems for connecting people to all kinds of old and new forms of knowledge. It also links between the expanding physical and intellectual universes. Thus, looking at a traditional library, for example, throughout history, its mission is to facilitate the free flow of information. While this mission remains unchanged, yet the ways in which information and knowledge workers (including librarians) used to fulfill this mission is constantly changing due to the advent of technology. In other words, the tools they use to create, collect, transmit and preserve recorded information sources, and the ways they organize, store, retrieve and use all types of available information have changed. The coming of digital knowledge age has profound impact on the kind of information services available for people and institutions.

Chinese Memory Net (CMNet): A Model for Collaborative Global Digital Library Development

Chinese Memory Net (CMNet): US-Sino Collaborative Research toward A Global Digital Library *in Chinese Studies* is one of the very first collaborative research projects, supported by the NSF's International Digital Library Program (IDLP), which announced its first awards in May 2000. *CMNet* hopes to create a global collaborative model for distributed cross-disciplinary digital library development in different parts of the world. Although "Chinese studies" is the selected focus, yet the definition is deliberately broad to include all Chinese related subject fields, thus the "Chinese Memory." The projects' R&D results are expected to be beneficial to other "studies" as well, and expandable to include all types of collaboration in different parts of the world (Chen, 2001).

Why CMNet?

With the growing interest in and need for education and research in academic institutions world wide, there is an urgent need to develop an effective collaborative worldwide system for sharing rich information resources. Currently most of the needed information is not available in digital form. While efforts in converting them to digital collections are increasingly obvious in various parts of the world, the activities are haphazard and are being introduced with insufficient concern for interoperable methods, standards and technologies. Furthermore, almost all efforts in many countries are not coordinated internationally. Thus, a pressing question would be:

> *How to develop effective international cooperative research that can help avoid duplication of effort, prevent the development of fragmented digital systems, and encourage productive interchange of essential knowledge and scholarly data around the world?*

CMNet's Research Approaches and Activities

CMNet has first brought selective academic educational and research partners in China -- Peking University, Tsinghua University, and Shanghai Jiaotong University; Taiwan -- National Taiwan University, National Tsinghua University, and the Academia Sinica; and the USA together working toward an effective and sustainable global digital library in Chinese Studies. The collaborators and affiliates of this complimentary and synergetic group possess experiences, knowledge, expertise, and capabilities in different but related research areas with their research activities funded by their respective governmental agencies. Several participating institutions will bring to the research some superb collections of Chinese culture and heritage, which form the current core of *CMNet*'s Chinese memory. As the collaboration grows, the content and scope related to the Chinese memory will grow accordingly so will the research activities and techniques involved. These collections are unique and essential to education and research, but currently are not accessible or usable via global networks because of distance, form, and technical barriers. This research project is to find new ways to enable academic users to access and exploit these significant research collections via global networks. The R&D activities will also contribute to enhance the capabilities and effectiveness of the collaborative digital library development (Chen, 2001).

The diversified Research areas include the following:

- Metadata techniques and standards for Chinese information
- Digital image systems
- Intelligent HCI techniques – Include multi-lingual, multi-modal search and retrieval; translation engine; etc.
- Agent-based information gathering and service for digital library
- Digital library architecture – Include interoperability & scalability study on Internet environment; management framework and organization schema for distributed, heterogeneous, and mass information repositories; etc.
- Digitizing techniques for Chinese inform. – Include pattern analysis & recognition for ancient books; video/audio structuring process; interactive 3D display techniques; etc…
- Prototyping systems for Chinese cultural and heritage information
- Intellectual property rights protection -- Safeguarding digital library contents and managing electronic copyright
- Linguistic technology and resources for English-Chinese bilingual information system
- Lexicon-based knowledge linking -- Approaches towards a WordNet

CMNet'S Multimedia Prototype System at Home Base

The need for creating digital contents has been discussed earlier. But, the difficulties are many. It is important to note that copyright and intellectual property issues have been major barriers for large digital content building. For this reason, the contents of *The First Emperor of China*'s videodiscs (a set of 4 sides with each containing 56,000 analog images and 30-minute digital sound track) and multimedia CD (with 6-hour interactive courseware with links to over 500 digital images and 30-minute digital video as well as text files) are invaluable and must be capitalized. In addition, the database consisting of rather detailed information on the most significant 2000 some images of the videodiscs proves to be also invaluable. Each record consists of fields that can be easily translated to some of the core elements of current-day metadata standards, such as those of the Dublin Core.

While time and space preclude more detailed discussion (see Chen, 2001 for more information), the following figures will show some of the features of the latest CMNet's *The First Emperor of China* Collection's prototype image system consisting of 250 images of very high resolution, ranging from 10 to 75 MB each in its compressed TIFF format. This visual collection utilizes Insight® of Luna Imaging, Inc. This Exhibition enables distributable and scalable online access to images from anywhere. It provides a dynamic and interactive user environment where one can view, compare and organize images, or even present them to others. Figure 4 is the Group Window of *The First Emperor of China* Collection. All functions for finding, displaying and using images begin from this window.

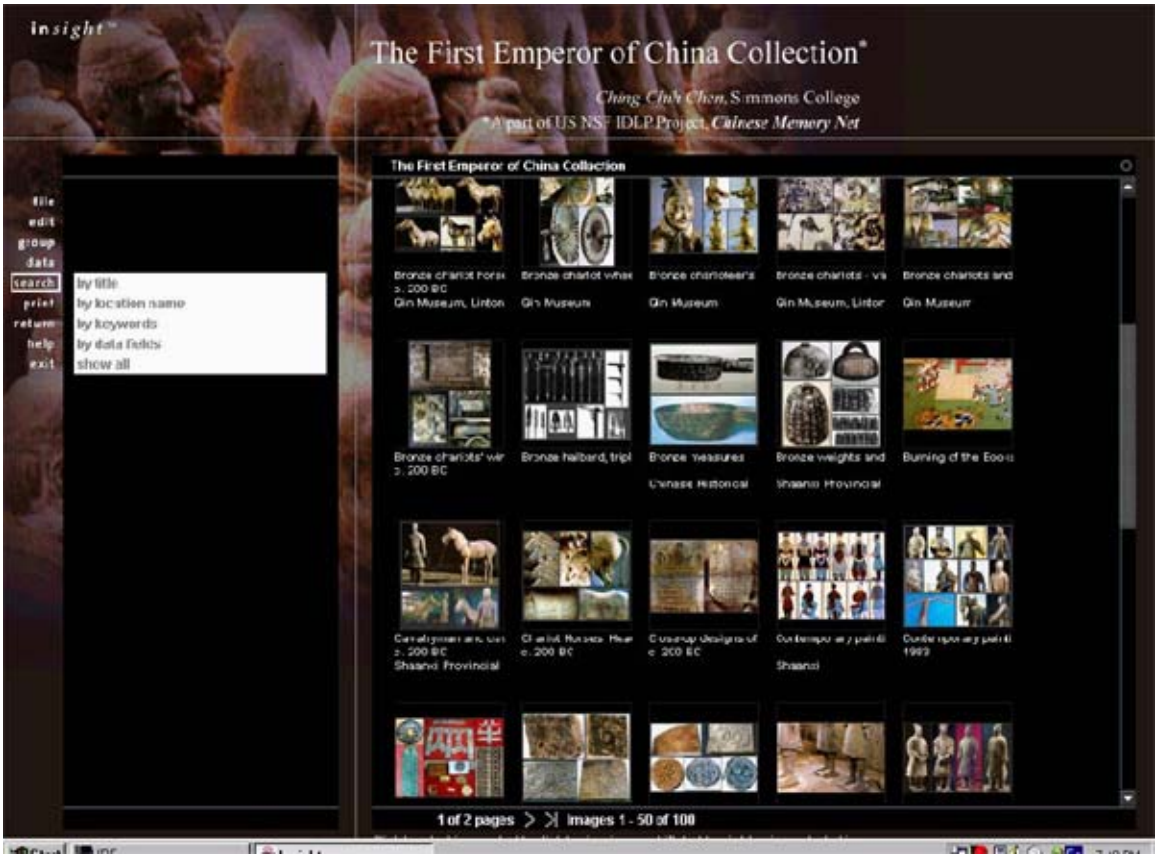

Figure 4: The Group Window of *The First Emperor of China* Collection (Visual Exhibition)

When a given image is selected -- in this case, the standing terracotta warriors -- the image's descriptive data is retrieved, shown on the upper-right portion of Figure 5 and the chosen images can be retrieved one by one for comparison purposed. In Figure 5, not only we can see the simulated original multi-metalic colors of the full-size figures, but also we can select any portion of it for enlargement for further comparison and verification purposes. Since each of these images are of very high resolution, it can be zoomed many times for very fine details.

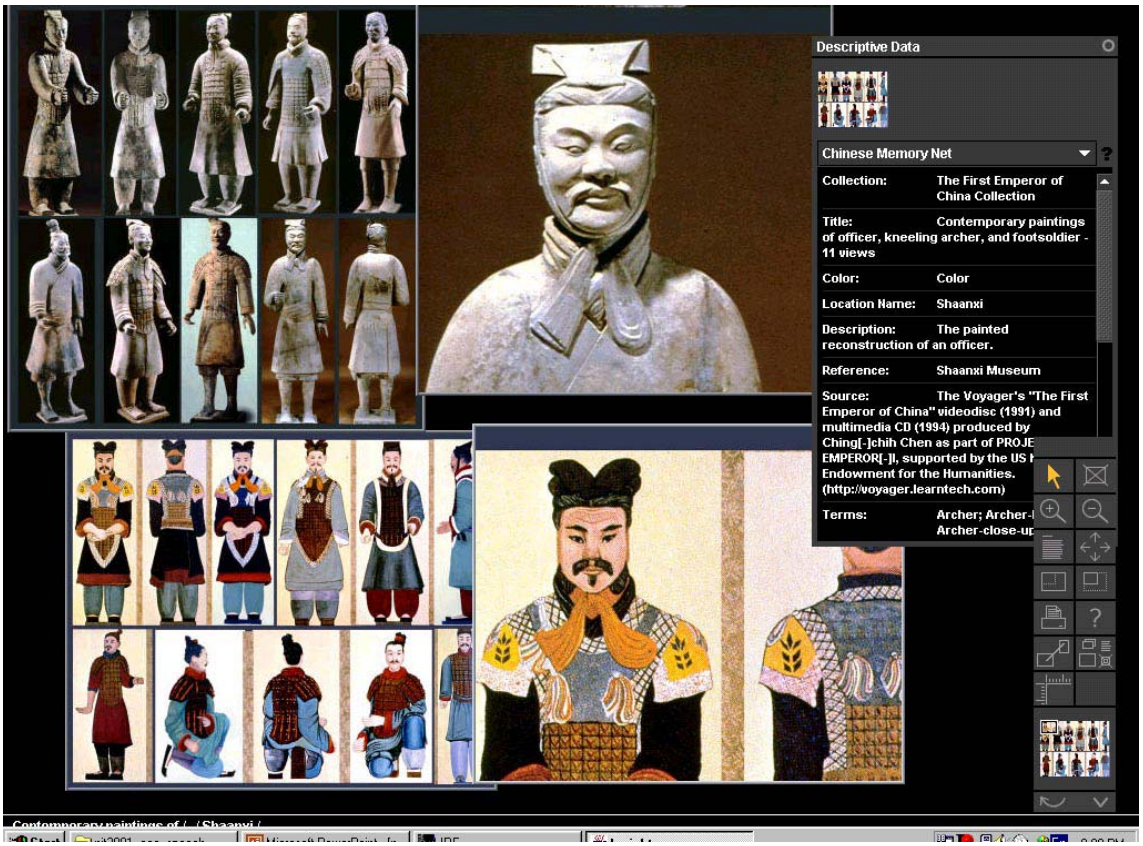

Figure 5: High-resolution images of the standing warriors retrieved from *The First Emperor of China* Collection (Visual Exhibition) for closer study in the Image Workspace

In addition to high-resolution images, relevant and related digital videos can also be retrieved over the Internet, and hyperlinks to relevant text and non-text resources can also be made. This clearly demonstrates the incredible potential for scholarly study and research. With this type of information available globally via Internet, we can truly see how "information access" has been elevated to a totally different and much higher level.

The China-US Million Book Digital Library Project

Since digital content is the key to global digital library development, it is exciting to discuss briefly the newly announced China-US Million Book Digital Library Project, which is precisely to meet the challenge to place large-scale valuable digital contents online and making them accessible globally via the Internet.

The Chinese part of this ambitious project is to be led by Dr. Gao Wen, Professor of Computer Science, Vice President of University of Science and Technology of China, Deputy President of Graduate School of Chinese Academy of Sciences; Chief Scientist of National HI-Tech Development Projector Intelligent Computing System Subject. The US part is to be led by Dr. Raj Reddy, University Professor of Computer Science and Robotics and Former Dean of the School of Computer Science, and Co-Chairman of the US President's Information Technology Advisory Committee (PITAC) and Dr. Ching-chih Chen, Professor of the Graduate School of Library and Information, Simmons College, Boston; Member of PITAC; and Principal Investigator and Project Director of the US National Science Foundation International Digital Library Program's *Chinese Memory Net (CMNet).*

According to the Chinese Ministry of Education's News Release:

In China, 10 or more primary universities and institutes, including Peking University, Tsinghua University, and the Graduate School of Chinese Academy of Sciences in China will be organized to participate in this Million Book Project, led by the 211 Office of the Ministry of Education. Each

participating university-based Digital Library Center has agreed to take on the needed research and development tasks in developing high-end computer systems to meet the needs of the Million Books Project. More significantly, each agrees to choose and offer its special, unique, and valuable information resources - books, paintings, sculptures and cultural objects - for this digital project. Each participating center will agree to digitize these resources under the sponsorship and supervision of the 211 of the planning Committee of the Ministry of Education of China.

In the US, under the leadership of Dr. Reddy of Carnegie Mellon University, and Dr. Chen of Simmons College, other interested universities and companies will be involved to participate in this project. They will offer financial and technical support including hardware such as computers, scanning equipment, and software, as well as necessary technical training. Both sides will join hands in conducting research and development work in creating tools and contents, as well as addressing many forthcoming major technical issues.

This project is just beginning, but the potential is great. Technologically we can visualize the federated collection of high-quality research and source materials to be shared well beyond China and US. The synergies that emerge from the interactions between this program and CMNet, and among these and other projects are indeed great. This project will show that digital services can provide unique resources that otherwise might not be available to educators and researchers.

Conclusion

Over the years, higher education communities around the world have been pioneers in providing online access to their libraries and other database services. In recent years, these efforts have influenced national libraries, museums, and archives, many of whom start to make more and more of their resources available online via the Internet. Thus, a growing number of initiatives have been devoted to the digitization and preservation of national treasures. As a result, a wealth of online digital services covers the spectrum of academic interest. In addition to reference databases and digital journals and books, we have at our fingertips digital archives consisting of texts, photographs, images, maps, films, videos, interviews, etc… Truly, technology has changed the way we learn, teach, do research, access information, and provide information services. In marching toward the virtual worlds of information and knowledge, a whole new set of challenges is waiting for us as well. While standards that support interoperability will be essential, much tougher issues are also obvious, such as copyright, intellectual property, and security on one end, and multilingual and multimodal systems on the other, to name just a few.

According to the US Census 2000, "281.4 million people, 143.4 million women, 138.1 million men, 35.3 million Hispanics, 33.9 million African American, 10.1 million Asian Americans, 1.4 million families" used the Internet. The numbers for now are undoubtedly larger. As technology grows, so do the numbers of Internet users. As educators, we are obviously concerned on what these users are using the Internet for. This clearly will lead to the need for quality "contents." Examples offered in terms of building high-quality contents and making them available for the users via global collaboration in this paper should be helpful.

From the technology side, while we have marveled at the incredible technological advancement of this past decade, particularly in computing and communications, let me conclude by repeating what Bill Gates said on May 23, 2001 at a meeting of about 140 corporate leaders at Microsoft's 5th CEO Summit, "the spectacular bursting of the dot-com bubble is not the end of the technology boom, only the end of the beginning," and "the next decade is the big one. This is the decade when your involvement with computing will be pervasive… The next few years will see software and the Internet tie together everything from PCs to handheld computers to new devices like the tablet PC, a folder-sized portable computer with a large-screen that can be written on."

Clearly, the exciting future is yet to come. We can see that the vision for universal information access will be realized sooner than we think!

References

Chen, Ching-chih (1994). Multimedia and the First Emperor of China: Moving toward a digital knowledge base. *Multimedia Today*, 2 (2), 68-71.

Chen, Ching-chih (2001). Chinese Memory Net (CMNet): A Model for Collaborative Global Digital Library Development. *Global Digital Library Development in the New Millennium: Fertile Ground for Distributed Cross-Disciplinary Collaboration*, edited by Ching-chih Chen. Beijing: Tsinghua University Press, May 2001. 21-32. Based on the Opening Speech presented at the *NIT 2001: 12th International Conference on New Information Technology*, Beijing, China, May 29-31, 2001.

President's Information Technology Advisory Committee. Panel on Digital Libraries. (2001). *Digital Libraries: Universal Access to Human Knowledge. Report to the President.* Arlington, VA: National Coordination Office for Information Technology Research and Development. http://www.itrd.gov.