

Zhaoneng Chen Hsinchun Chen Qihao Miao
Yuxi Fu Edward Fox Ee-peng Lim (Eds.)

LNCS 3334

Digital Libraries: International Collaboration and Cross-Fertilization

7th International Conference on Asian Digital Libraries, ICADL 2004
Shanghai, China, December 2004
Proceedings

 Springer

Global Memory Net: New Collaboration, New Activities and New Potentials

Ching-chih Chen

Professor

Graduate School of Library and Information Science
Simmons College, Boston, MA 02115, USA

Email: chen@simmons.edu

In technological terms, it has been a long time since my *PROJECT EMPEROR-I* -- a multimedia interactive videodisc project on the First Emperor of China's famous terracotta warriors and horses in 1984. At that time, *PROJECT EMPEROR-I* demonstrated that multimedia technology could change the way we seek, demand, and use information. Two decade later, fueled by enormous progress in science and technology, we have come a very long way from the use of interactive multimedia technology in the workstation environment to the global networked environment. We have moved from the use of hardcopy and analog resources to digital content, which users can search, retrieve and use instantly to meet their needs over the global network with no national boundaries. We have also moved from the offering of multimedia content of one specific subject topic to the digital content of all media formats on all related subject topics to the world instantly. We are truly living in a new period of unprecedented opportunities and challenges [1]! So, in this digital era, we have witnessed the exciting convergence of content, technology, and global collaboration in the development of digital libraries [2] with great potential for providing universal information access.

Thus, today's information seekers, regardless whether they are general public, school children, or those from research and higher education communities seek information for education, research, entertainment, or enrichment in very different ways from before. From the information resources point of views, the old model of "owning" a collection has given way to "sharing," and the new emphases have shifted from possessing large "physical libraries" to "virtual libraries" digitally distributed all over the world.

In the last two decades, I have experienced much of these transformations up-close and personal through my own R&D activities – from the creation of interactive videodisc and multimedia CD in the 80s and 90s to leading a current international digital library project, *Global Memory Net*, supported by the International Digital Library Program of the US National Science Foundation [2, 3].

GLOBAL MEMORY NET AND RECENT DEVELOPMENT

From PROJECT EMPEROR-I to Chinese Memory Net to Global Memory Net

In the early 80s, the by-product of *PROJECT EMPEROR-I's* is a set of interactive videodisc, called *The First Emperor of China*, content of which later was converted to a popular multimedia CD product of the same title in 1991 and published by the Voyager Company. The core image

collection of this product together with the extensive descriptive annotations (later known as metadata) of these resources has become the core collection of *Chinese Memory Net (CMNet)* which I proposed to NSF's International Digital Library Program (NSF/IDL) in 1999, and funded from 2000.

The NSF's supported *CMNet* since 2000 is intended to develop a model for international collaboration with various R&D activities in digital libraries. It hopes to accomplish "more" with "less," avoid duplication efforts, and capitalize R&D results from other major funded digital library R&D projects. *CMNet's* Chinese partners are Peking University, Shanghai Jiaotong University and Tsinghua University. Although we did not achieve one of the original goals in bringing the digital contents available in our partner institutions together, in the short four years, it has made progress in developing collaborative infrastructure for digital library development. Both *CMNet* and my *NIT 2001 conference* in Beijing played important role in fueling the development of digital libraries in China [4].

While building the digital library community and infrastructure, *CMNet* also started the labor-intensive R&D activity in content and metadata building. This activity has paid off because these invaluable image resources and metadata have formed attractive basis for a number of exciting and productive technology-oriented collaborative works with computer scientists, such as a few listed in the following with more complete reference provided in [5, 6]:

- Open Archive Initiative (OAI) research,
- Intelligent agent and text-based image retrieval [7, 8],
- Semantic sensitive content-based image retrieval [9],
- Digital video using the Infromedia technologies [10], and
- Machine learning for annotation [11].

Once it is possible to develop a multimedia digital library in one subject disciplinary or for one geographical area, it is upward scalable to include more subject topics and bigger geographical areas. This was the case with the expansion of the scope of *CMNet* to *GMNet since 2002*. *GMNet* developed out of the *CMNet* project which concentrates on images and video related to China's ancient culture. In the last two years, more collaboration with several major institutions in different countries has become a reality and thus *CMNet* is changed to *Global Memory Net*. It is being expanded to cover the 'memory' of other parts of the globe [3].

As shown in Figure 1, the tentative *GMNet* homepage, *GMNet* literally has space holder for all countries in the world although this tentative homepage has listed tentatively only a few continents and countries under each at this moment.

The Scope of *Global Memory Net*

The name of *Global Memory Net* clearly articulates both the potential coverage and scope of this project [1, 2, 3]:

1. **Global** – It means global coverage.
2. **Memory** – "Memory" refers to all types of treasures, thus *GMNet* has the structure to cover all kinds of invaluable memories related to culture, heritage, history, art, music, science,

technology, medicine, etc. However, at this initial stage, and with initial entry of the extensive visual memory related the First Emperor of China's terracotta warriors and horses, *GMNet* is focused on the world significant cultural, historical, and heritage materials. Once this focus is well underway, *GMNet* will expand to cover other kinds of "memories."

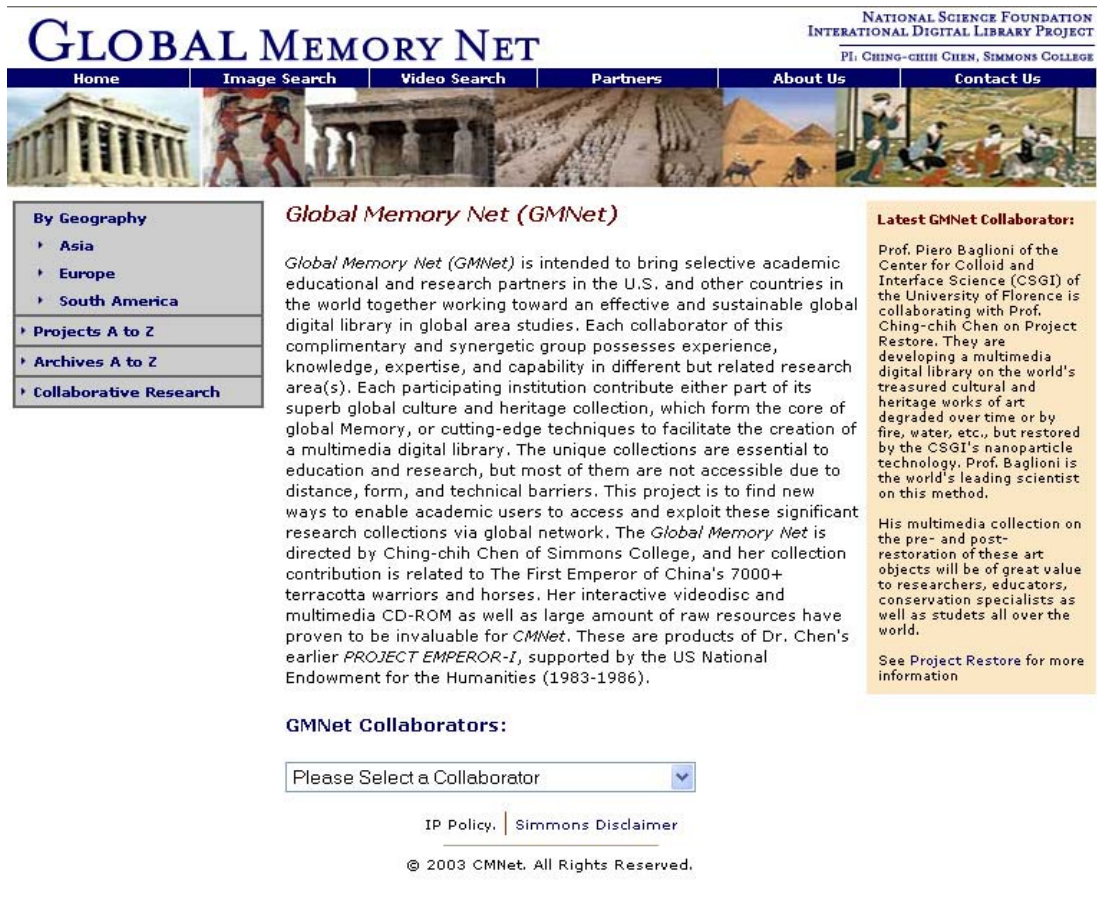


Figure 1. Home Page of the Global Memory Net

3. **Memory** – “Memory” refers to all types of treasures, thus *GMNet* has the structure to cover all kinds of invaluable memories related to culture, heritage, history, art, music, science, technology, medicine, etc. However, at this initial stage, and with initial entry of the extensive visual memory related the First Emperor of China's terracotta warriors and horses, *GMNet* is focused on the world significant cultural, historical, and heritage materials. Once this focus is well underway, *GMNet* will expand to cover other kinds of “memories.”
4. **Net** – This means that *GMNet* hopes to network all significant global resources together. Instead of encouraging the development of small and fragmented digital libraries, it hopes to be a networked portal to offer needed resources instantly with the simple click of the mouse.

Global Memory Net Offers the World Instantly

For the First Emperor of China's content, *GMNet* is a comprehensive image digital library on that subject. For other world's cultural and heritage contents, *GMNet* is an effective digital portal which offers the world instantly to the information seekers.

It is impossible to describe all the features of in such a short introduction. In the simplest way, consider *GMNet* an easy to use digital portal utilizing the cutting edge image retrieval technology to enable one to take a visual tour of any country’s culture, heritage, history, and world contributions, all while sitting at one’s computer. This soon to be available *GMNet* will provide, in addition to the traditional search by image retrieval capabilities with considerable textual supports in a way not possible before. For example, from the page like that shown in Figure 1, one can go to China and then Emperor Image Base quickly. Then one will be able to retrieve invaluable images related to the First Emperor of China, for example, by conducting the traditional search using the Google protocol if predefined specifics of the images are known. In this case, one can search literally every field of the metadata, such as creator, title, location, time period, description, keyword, reference source, etc. In this approach, keyword search is likely to be the most popular one.

However, in most cases, one does not have any idea on what kind of images are available in *GMNet*. Just like in a library, we need to provide the user an opportunity to browse the stack, and find what they need and want. In this case, in *GMNet*, we powered our images’ random retrieval with the cutting edge content-based image retrieval technique, SIMPLiCity, developed at the Stanford University under NSF’s DL-I phase, and then at the Penn State University under NSF/ITR [Ref. 9 provides more references]. This allows users to browse, retrieve, enjoy, and learn in just seconds through multiple thousands of digital images accurately and effectively.

For example, when the icons of the images of the Emperor collection are displayed randomly in Figure 2, one spots the image related to “Han silk” of interest. In this case, one can ask the system to provide “SIMILAR” images by clicking “Similar” without typing any word, *GMNet* will display in seconds all the images in the collection similar to the one selected. This opens up all possibilities for all related maps which are totally unknown to the user (see Figure 3).

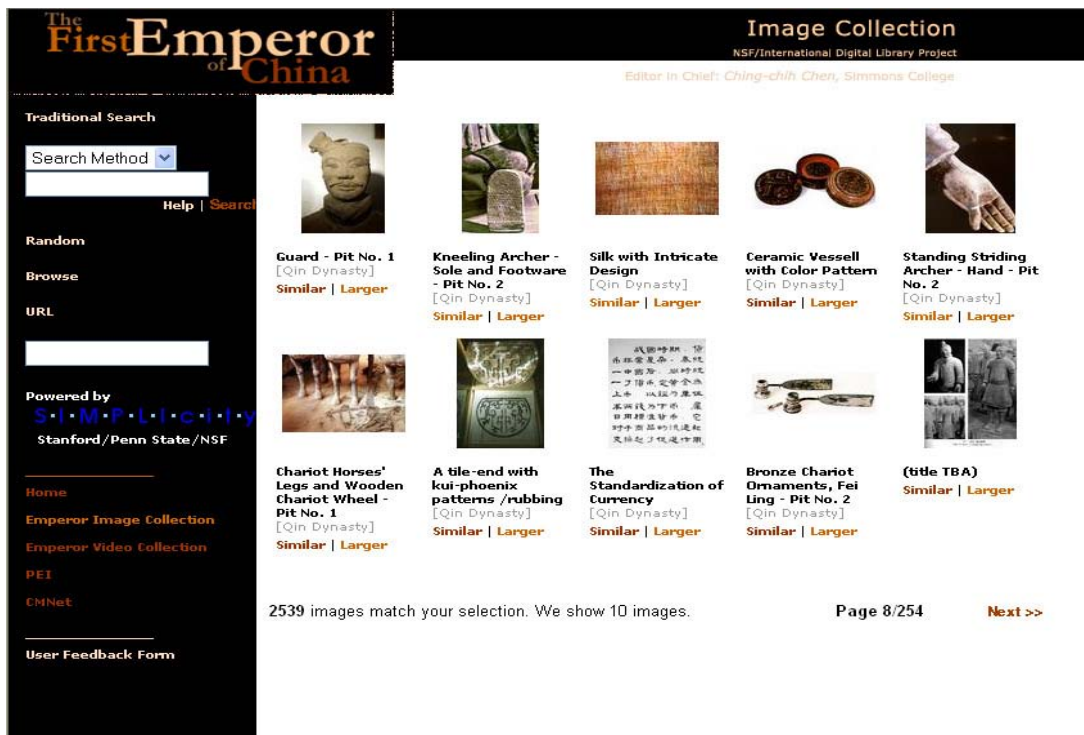


Figure 2. Random images for user’s browsing and selection

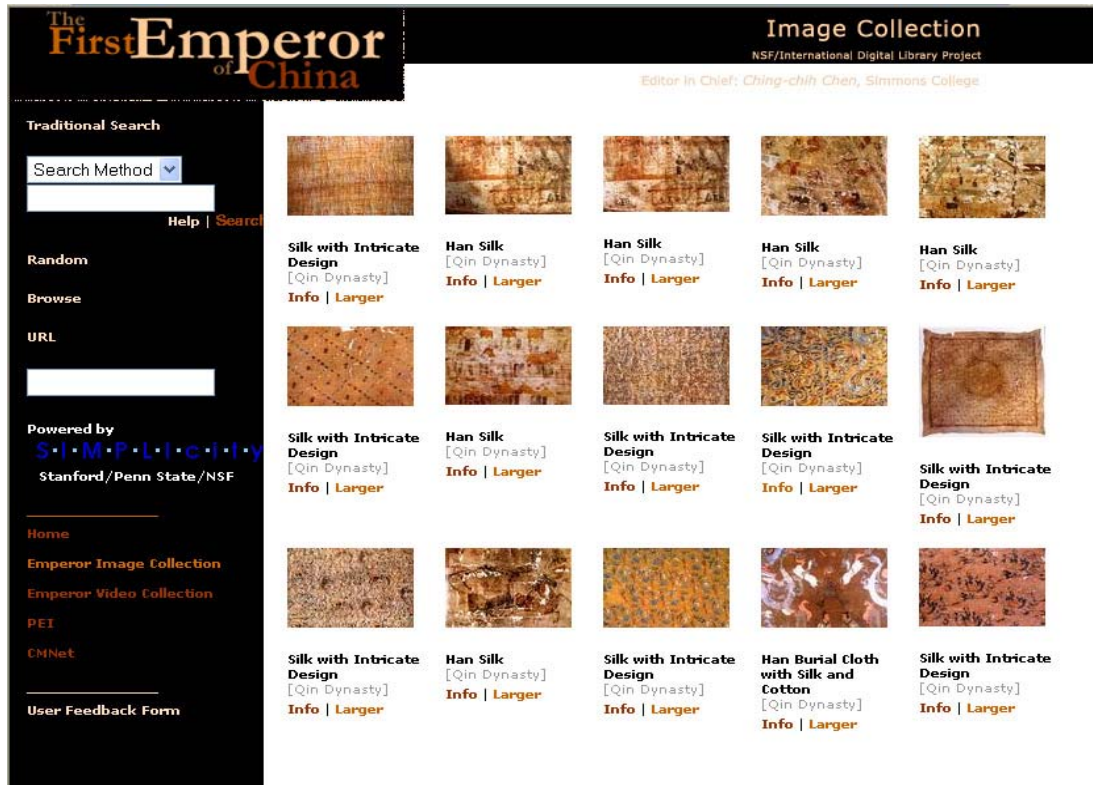


Figure 3. Similar images on the chosen “Han silk” are displayed

Once these massive numbers of images are displayed, one would be able to enlarge a chosen image by clicking on “larger”, and multiple levels of zooming will be possible and dynamic digital watermark will be instantly generated to offer the “ownership” information of the image as shown in Figure 4. One will be able to find more textual descriptive information as well as reference sources and in some cases, full-text original source on a chosen image instantly by clicking “Info” Fig. 5).

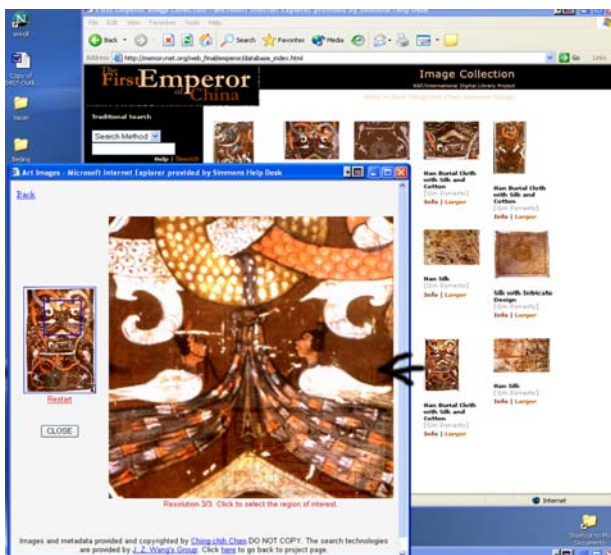


Figure 4. Chosen image enlarged with digital watermark

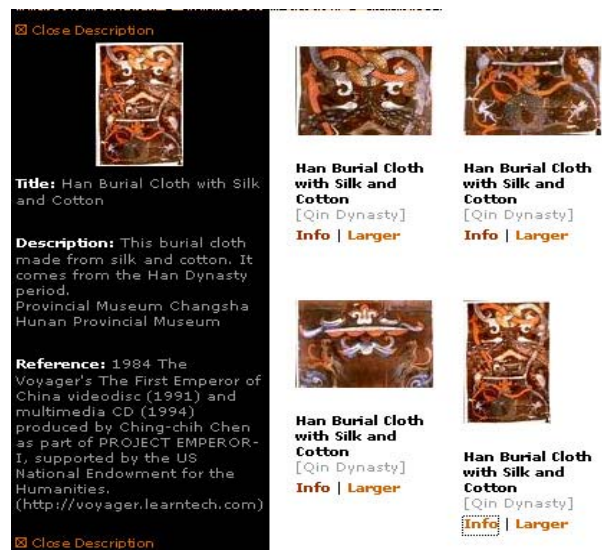


Figure 5. Metadata information is on the left column

Global Memory Net – New Collaboration

The move to *Global Memory Net* has enabled us to expand our collaborative community and subjects greatly, far more than we could imagine during the first three years of *CMNet*. Starting from our exciting collaboration with University of Florence on Project Restore [12] with the exciting pre- and post- restoration images of the invaluable Italian art objects (see Figure 6) with nano-particle chemistry technology, other collaborative collections have mushroomed to include many countries and subjects. We shall name a few examples in the following table and Figure 7:



Figure 6. Project Restore's pre- and post-restoration images from University of Florence

- China - Chinese painting, many historical unique collections, architecture, etc.
- Cambodian - Ancient temples, etc.
- India - Architecture, palaces, temples, goddess, etc...
- Vietnam - Historical development of the former Saigon
- Italy - Historical artifacts
- Europe - Cathedrals, Castles, etc.
- World - Global musical instrument, national libraries, etc.



Figure 7. Some sample image collections

In addition to these, the collaboration with the Asian Division of the Library of Congress is an exciting development. Currently we have included the unique Naxi manuscripts' images of the Library of Congress in *GMNet* (see Figure 8). In addition, we are also exploring the collaboration with Unesco's *Memory of the World* Programme. We have identified over 1000 digital collections in the world, and it is possible for us to retrieve all web sites of digital collections of similar color and design of an organization instantly, such as those of the Unesco's *Memory of the World* Programme as shown in Figure 9. Once the website is selected, information on the site can be located instantly, and the user can be linked to the site instantly. Currently, Unesco has 91 digital collections from 45 countries of this nature, thus, our digital portal has certainly boosted the accessibility and value of these collections instantly.



Figure 8. Images of the unique NAXI manuscripts' can be retrieved instantly



Figure 9. Images of Unesco's *Memory of the World* sites can be retrieved and linked instantly (in yellow)

Global Memory Net – New R&D Activities

As mentioned earlier, although *GMNet* has concentrated thus far on digital image cultural and heritage collections thus far, we have already collaborated with Carnegie Mellon University in exploring the sophisticated digital video retrieval capabilities using the world renown Informedia Technologies [2,3,10], see Figure 10. The Chinese University of Hong Kong has extended the Informedia Technology capability to handle Chinese language, and we are exploring the possibility of using that.

In addition to digital videos, our research will also

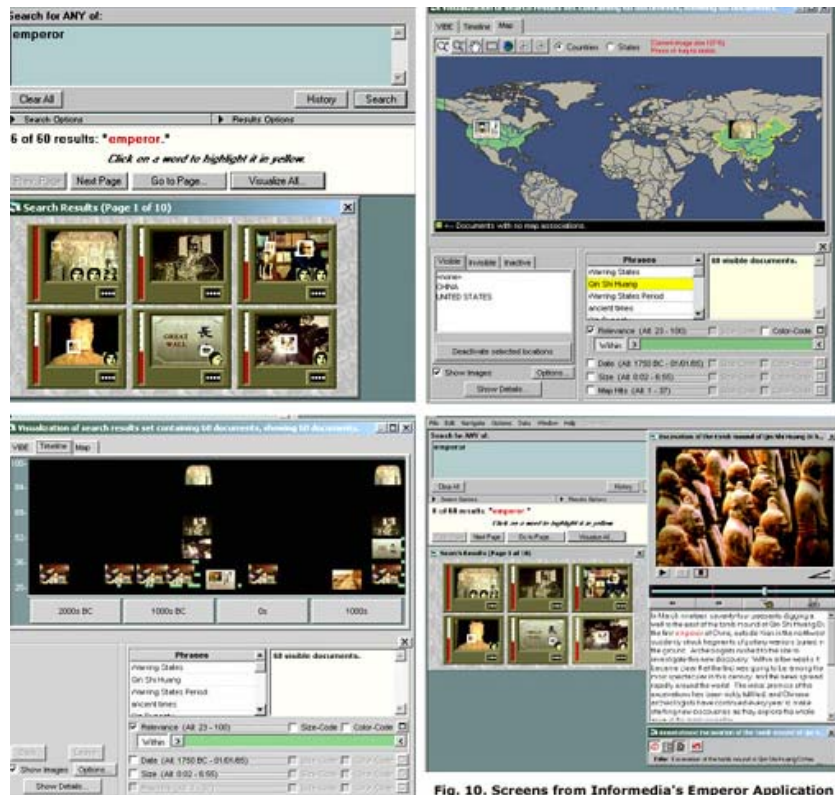


Fig. 10. Screens from Informedia's Emperor Application

incorporate sound and music. One of the perfect starting points will be with the world's musical instruments. Another possible area would be with the language learning and writing.

One final mention of an exciting new activities would have to be our newly funded NSF/IDL [NSF/IIS-Special Projects (IIS)] 2-year project from 2004-2006, entitled "*International Collaboration to Advance User-oriented Technologies for Managing and Distributing Images in Digital Libraries*" with James Z. Wang of Penn State University and Jianbo Shi of University of Pennsylvania as my co-PIs. This project will develop user-oriented image management of distribution technologies for digital libraries. An interdisciplinary team of computer and information scientists from US, China, and Taiwan will investigate efficient ways to search digital collections of images using an integrated approach. The team will use real-world digital library datasets to develop user-oriented technologies suitable for practical deployment. Notably, the research will utilize an existing collection consisting of a large quantity of images associated with The First Emperor of China's terracotta warriors and horses of all types of resolution and with enormous cultural significance as well as the existing rich descriptive information. In addition to Ontology-based image retrieval, machine learning based content-based image retrieval including, we will explore the difficult object-based partial image searches. We also hope to extend the intellectual property (IP) protection techniques.

CONCLUSION

During 1998-2002, I was privileged to serve on the US President's Information Technology Advisory Committee's (PITAC). Our PITAC's Digital Library Panel's Report, *Digital Libraries: Universal Access to Human Knowledge*, has a vision for digital libraries:

"All citizens anywhere anytime can use any Internet-connected digital device to search all of human knowledge. ... In this vision, no class-room, group, or person is ever isolated from the world's greatest knowledge resources." [13]

This is a vision easily said than done! There are many obstacles on the road, thus we are a long way from approaching this "elusive" vision.

From "sharing" and "accessing" points of view, we must first have much more "quality" digital contents, we must collaborate internationally in content building because no one can have everything, then we must have the technology to cope with these contents, and the infrastructure to deliver, access and retrieve them [2,3]. This is what *Global Memory Net* is inspired to do specifically in content building and method development areas. The new collaboration and new R&D activities have expanded our research horizon, and have offered us great opportunities for digital library community building, for making digital collections alive and accessible, and for contemplating much more practical R&D agenda in areas of metadata standards, interoperability, scalability, retrievability of difficult multimedia contents, and usability of these resources for knowledge creation.

ACKNOWLEDGMENTS

Chinese Memory Net and *Global Memory Net* is supported by the US NSF/IDLIP under grant no. IIS-9905833. PROJECT EMPEROR-I was supported by the Humanities in Libraries Program of the US National Endowment for the Humanities. Examples given on the use of SIMPLicity is a collaboration with James Z. Wang of the Penn State University whose work is supported by the NSF/ITR program under grant no. IIS-0219271.

REFERENCE

1. Chen, Ching-chih, "Global Memory Net Offers Users the World Instantly," *Library Times International*, 21 (1): 1-4 (July 2004)
2. Chen, Ching-chih, "The promise of international digital library collaboration for innovative use of invaluable resources," in *Human Information Behaviour & Competences for Digital Libraries* (Keynote in *Proceedings of the Libraries in the Digital Age*, May 25-29, 2004, Dubrovnik and Mljet, Croatia. pp. 7-15.
3. Chen, Ching-chih, "Global Memory Net offers the world instantly: Potentials for universal access to invaluable contents," in *Proceedings of CCDL: Digital Library – Advance the Efficiency of Knowledge Utility*, Beijing, September 5-8, 2004. Beijing: National Library of China, 2004.
4. Chen, Ching-chih, ed. *Global Digital Library Development in the New Millennium: Fertile Ground for Distributed Cross-Disciplinary Collaboration*. Beijing: Tsinghua University Press, May 2001.
5. Chen, Ching-chih, "Global Memory Net: Potential and challenges for archiving and sharing cultural and heritage resources," *Proceedings of the ICDL (International Conference on Digital Libraries) 2004*, Delhi, India, February 25-27, 2004. Delhi, India: TERI, 2004. pp. 3-10.
6. Chen, Ching-chih, "Past forward - Digital media for cultural heritage: The case of the Global Memory Net," Invited lecture given at the 10th *Annual Lecture of Informatics*, sponsored by Informatics, Bangalore, India, February 29, 2004.
7. Soo, V. W., C. Y. Lee, C. C. Lin, S. L. Chen, and Ching-chih Chen, "Automated semantic annotation and retrieval based on sharable ontology and case-based learning techniques," full-length paper accepted for present at the *ACM/IEEE Joint Conference of Digital Libraries*, Houston, TX, May 29, 2003. Published in *Proceedings of the ACM/IEEE JCDL Conference*, 2003. 12 pages.
8. Soo, V. W., C. Y. Lee, C. C. Yeh, and Ching-chih Chen, "Using sharable ontology to retrieve historical images," *ACM/IEEE JCDL Proceedings, Portland, OR, July 15-18, 2002*. pp. 197-198.

9. Wang, James Z., Jia Li and Ching-chih Chen, "Interdisciplinary research to advance digital imagery indexing and retrieval technologies for Asian art and cultural heritage," *Proc. ACM Multimedia, Workshop on Multimedia Information Retrieval*, Juan Les Pins, France, December 2002. 6 pp.
10. Wactlar, Howard D. and Ching-chih Chen, "Enhanced perspectives for historical and cultural documentaries using Informedia technologies," *Proceedings of the ACM/IEEE Joint Conference of Digital Libraries, Portland, OR, July 15-18, 2002*. pp. 338 – 339.
11. Wang, James Z., Kurt Grieb, and Ching-chih Chen, "Machine annotation and retrieval for Digital Imagery of Historical Materials," Submitted to *Journal of Digital Libraries: Special Issue on Multimedia Contents*.
12. Baglioni, Piero, Rodorico Giorgi, and Ching-chih Chen, "Nanoparticle technology saves cultural relics: Potential for a multimedia digital library," in *Online Proceedings of DELOS/NSF Workshop on Multimedia Contents in Digital Libraries*, Crete, Greece, June 2-3, 2003.
13. US. President's Information Technology Advisory Committee (PITAC). Digital Library Panel. *Digital Libraries: Universal Access to Human Knowledge*. February 2001.
www.hpcc.gov/pubs/pitac/pitac-dk-9feb01.pdf



Dr. Ching-chih Chen is Professor of the Graduate School of Library and Information Science, Simmons College, Boston, USA. A sought-after international consultant and speaker in over 40 countries, she is an author and editor of over 35 books and more than 180 scholarly journal articles. She was the Chief Conference Organizer of a series of 12 *International conferences on New Information Technology (NIT)* from 1986-2001 in different parts of the world.

A Fellow of the American Association for the Advancement of Science, she has received many awards and honors, including the Best Information Science Teacher Award of the American Society for Information Science, the Library and Information Technology Association's LITA/Library Hi Tech Award, the LITA/Gaylord Award for the Advancement in Library and Information Technology, and many others. During 1998-2002, she served as a member of the US President's Information Technology Advisory Committee.