

# Digital Libraries and Universal Access in the 21<sup>st</sup> Century: Realities and Potential for US-China Collaboration<sup>\*</sup>

**Ching-chih Chen**

Professor

Graduate School of Library and Information Science

Simmons College

Boston, MA 02115

E-mail: [chen@simmons.edu](mailto:chen@simmons.edu)

**Abstract:** “Global Digital Library and Universal Information Access” was a keynote delivered by the author at the *First China-U.S. Conference on Global Information Access: Challenges and Opportunities*, held at the National Library of China, Beijing, China, August 21-23, 1996. Since then, digital libraries have flourished to make the “global digital library” more than a partial reality. This paper intends to update what has happened in the last 8 years with specific references to two of the major international projects which she has been heavily involved. She hopes to show how activities like these can truly provide enormous opportunities for US-China collaboration in light of China’s content-rich information environment.

The two projects are:

1. *The China-US Million Book Digital Library Project* for universal access. Currently over a dozen top-rated Chinese academic institutions are involved with heavy investment from both US National Science Foundation and industrial sources as well as the Ministry of Education of China.
2. *Chinese Memory Net (CMNet)* supported by the US National Science Foundation/ International Digital Library Program since 2000. *CMNet* has been expanded to *Global Memory Net (GMNet)* which has great potential for much more substantive collaboration.

While the memory of the *First China-U.S. Conference on Global Information Access: Challenges and Opportunities*, held at the National Library of China, Beijing, China, August 21-23, 1996 is still quite vivid in my mind, yet, in technological terms, 8 years is a long time! I remember that my keynote at that time was addressing “Global Digital Library and Universal Information Access” [1], and stressing the need to plan for global information infrastructure. Yet 8 years later, everyone is taken for granted with the use of the Internet and World Wide Web. We are witnessing that the information technological innovations has intertwined with interdisciplinary knowledge base, which is propelling the 21<sup>st</sup> century’s knowledge economy. Currently in the spring of 2005, much of what I advocated in 1996 and the use of multimedia and global network are not only the mainstream practices, but have been taken for granted. In fact, 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 mixed subject topics to the world instantly. We are truly living in a new period of unprecedented opportunities and challenges! So, in this digital era,

---

<sup>\*</sup> The portion of this paper related to *Global Memory Net* has been modified and constantly updated with new materials from different keynote and invited speeches on this topics in the last three years in different parts of the world. Repetition in describing what *Global Memory Net* is unavoidable.

we have witnessed the exciting convergence of content, technology, and global collaboration in the development of digital libraries [2, 3] 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, they want to find their needed information 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" which value their large number of volumes, to "virtual libraries" digitally distributed all over the world [4].

Given this kind of digital environment, we have reasons to be optimistic with US-China collaboration. I remember clearly at the end of the *First China-U.S. Conference on Global Information Access: Challenges and Opportunities*, there were a number of resolutions for bi-national collaboration among libraries. There were resolutions at the end of the *Second China-U.S. Conference* as well. I am sure that in the last 8 years, some of the resolutions were realized while the others still waiting to be accomplished.

For me, as an individual, my collaboration with China has come a long way and for more than quarter century. In preparation of this talk, I revisited the listings of my speeches offered during my first invited trip to China in 1979, shortly after President Nixon's visit. During the month-long visit, numerous speeches on "new information and networking" were given major library and information institutions in China, such as:

- Beijing -- The Institute of Scientific and Technical Information, National Library, Peking University, Tsing-Hua University, and Chinese Academy of Sciences;
- Xian -- Xian Jiao Tong University;
- Shanghai -- Institute of Atomic Energy, Shanghai;
- Foochow -- Fukien Library Association.

The trip ended with a most memorable talk, entitled "*New Trends in Technology Applications and Scientific Management - Potentials for Chinese Library & Information Development*," sponsored by the Chinese Library Association and the National Peking Library, held at auditorium of the History Museum, Tienmen Square, Beijing, June 7, 1979. Over 1000 people from Beijing and peripheral areas attended the meeting because it was rather rare at that time to have visitors from outside China. Figure 1 showed the



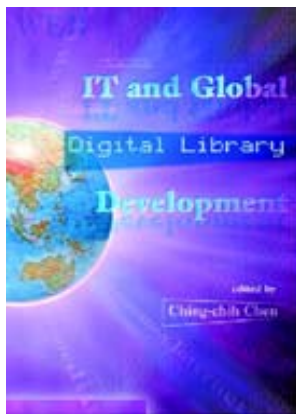
Figure 1. With Director Liu of NLC and his deputies at Tienman (1979)

picture taken with the Director of the National Library of China, Mr. Liu Ji-ping and his deputies, after the meeting at the Tienmen Square. This was the start of my 25-year collaboration with various institutions in China!

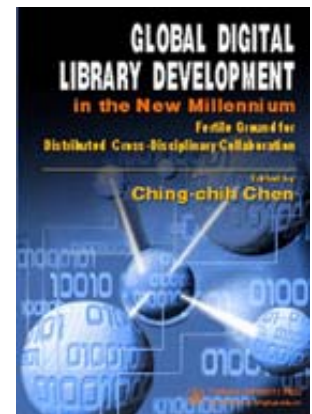
Then, since mid-1980s, I have experienced much of the transformations stated earlier up-close and personal through my own R&D activities – from the creation of interactive videodisc and multimedia CD in the 80s and 90s to organizing major international conferences promoting both global and US-China cooperation, from leading a current international digital library project, *Chinese Memory Net* and then *Global Memory Net*, supported by the International Digital Library Program of the US National Science Foundation [5], to being the co-PI of the *US-China Million Book Digital Library Project* [hereafter refer to as *US-China Million Project*]. Let me take this chance to share some of the highlights of these activities, and then address specifically the realities and potentials for US-China collaboration.

### ***NIT CONFERENCES ON DIGITAL LIBRARIES***

Of the twelve *NIT: International Conferences on New Information Technology* conferences organized by me since 1985, two -- NIT '99 in Taipei and NIT '2001 in Beijing were devoted fully to the digital library related topics. Both conferences contributed to global cooperation, including US-China, in advocating digital library for universal access. The *Proceedings* of both conferences were published as a full-length books as listed in the following. They document well both the calls and activities related to the topics of this paper, “Digital Libraries and Universal Access in the 21<sup>st</sup> Century”:



1. Chen, Ching-chih, ed. *IT and Global Digital Library Development*. Newton, MA: MicroUse Information, 1999.
2. Chen, Ching-chih, ed. *Global Digital Library Development in the New Millennium: Fertile Ground for Distributed Cross-Disciplinary Collaboration*. Beijing, China: Tsinghua University Press, 2001.



These two conferences were very early ones on the topics of digital libraries in the Asian region. Thus, they have led the way for the organization and offerings of many international conferences on digital libraries in the Pacific region, as well as in both mainland China and Taiwan. The conferences in China include the two *International Conference on Digital Libraries* organized by the National Library of China under the sponsorship of the Ministry of Culture of China in 2002 and 2004, the *International Forum on Digital Library and Project Negotiation* in Beijing in May 2002, the *International Asian Digital Library Conference* in Shanghai, with Shanghai Jiaotong University and Shanghai Library as the local conference organizer, in December 2004, etc. All these conferences have published their proceedings, which can offer much valuable information to those who are interested in topics presented and discussed in these conferences.

## CHINA-US MILLION BOOK DIGITAL LIBRARY PROJECT

*Million Project* is the brainchild of Prof. Raj Reddy of the Carnegie Mellon University, and the project received funding from the US National Science Foundation with Dr. Reddy and Dr. Glorianna St. Claire as co-PIs. It has many components including *US-China Million Book Digital Library Project*, *US-Indo Million Book Digital Library Project*, etc. The history, vision, and objective of the *Million Project*, or *The Universal Library*, can be found at the *Universal Library*'s website, <http://www.ulib.org/html/index.html>. Figure 2 is the opening page of this site.

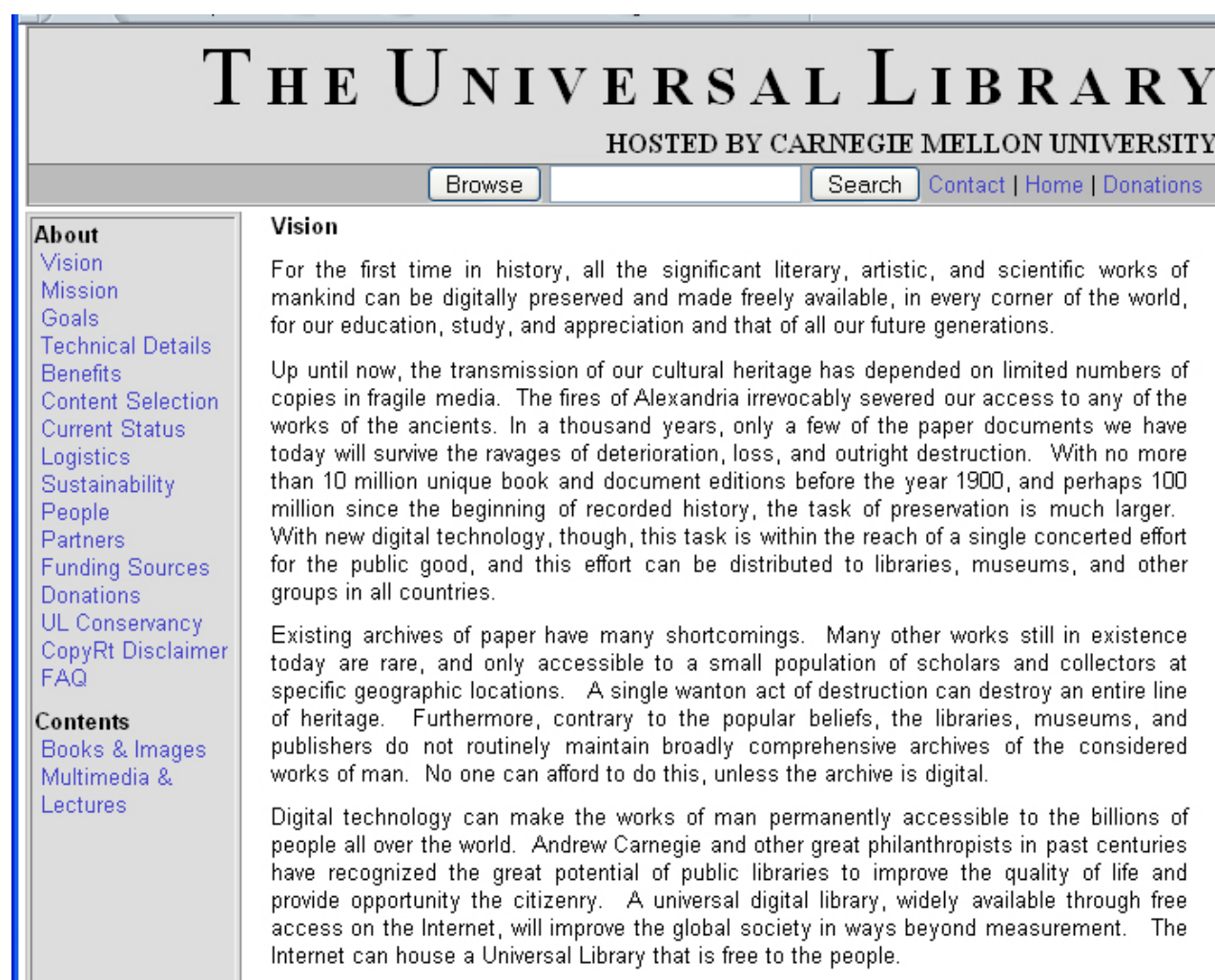


Figure 2. Home Page of the CMU's *Universal Library*

Dr. Reddy and I served on the US President's Information Technology Advisory Committee together during 1997 to 2002. Both of us share the same vision for universal access. He has advocated for "universal library", while I have advocated since 1993 "global digital library" [6]. For this reason, he asked me to serve as co-PI with him of the *US-China Million Book Digital Library Project*. As shown on the web page at <http://www.ulib.org.cn>, Figure 3, a "Memorandum of Understanding on the China-US Million Book Digital Library Project" was signed in Dec. 2000.





# The Universal Library

## China Site

Hosted by Graduated School of the Chinese Academy of Sciences

**Browse by Author:** [A](#)[B](#)[C](#)[D](#)[E](#)[F](#)[G](#)[H](#)[I](#)[J](#)[K](#)[L](#)[M](#)[N](#)[O](#)[P](#)[Q](#)[R](#)[S](#)[T](#)[U](#)[V](#)[W](#)[X](#)[Y](#)[Z](#)

**Browse by Title:** [A](#)[B](#)[C](#)[D](#)[E](#)[F](#)[G](#)[H](#)[I](#)[J](#)[K](#)[L](#)[M](#)[N](#)[O](#)[P](#)[Q](#)[R](#)[S](#)[T](#)[U](#)[V](#)[W](#)[X](#)[Y](#)[Z](#)

**Search:**

**About**

[Million Book Project](#)  
[Memorabilia](#)  
[Goals](#)  
[Download](#)  
[Partners](#)  
[Statistic](#)  
[CopyRt Disclaimer](#)

**Rare Collections**

Sikuquanshu

**Recent Update**

**About China-US Million Book Digital Library Project**

The human race would benefit by the creation and deployment of a Universal Digital Library with a vision to provide access to all human knowledge anytime, anywhere, including access, query, and print any book, magazine, newspaper, painting, image, music, video, or reference document. UDL's goal is to promote modern sciences and traditional wisdom to be utilized deeply, to enhance communication and comprehend between different cultures and to expedite human knowledge innovation and civilization progression. The first challenge for UDL is to organize one million books with text and images online and thus globally accessible.

To meet the challenge, China and US parties initiated the China-US Million Book Digital Library Project (short as Million Book Project). Dr. Raj Reddy, Herbert A. Simon University Professor of Computer Science and Robotics of Carnegie Mellon University, as Co-Principal Investigators of US side, Dr. Ching-Chih Chen, Professor and Former Associate Dean at Graduate School of Library and Information at Simmons College, Co-Principal Investigators of US side, Prof. Qin-Ping Zhao, Chief Officer of 211 Office of Ministry of Education of China, Professor of Computer Science of Beijing University of Aeronautics & Astronautics and Dr. Wen Gao, Deputy President and Professor of Computer Science, Graduate School of Chinese Academy of Sciences, as Principal Investigators of China side, signed "Memorandum of Understanding on the China-US Million Book Digital Library Project" in Dec. 2000.

In Aug. 2001, a Chinese delegacy with eight members led by Professor Yunhe Pan, president of Zhejiang University, visited US to promote Million Book Project. The delegacy visited CMU (Carnegie Mellon University), National Science Foundation and MIT (Massachusetts Institute of Technology). The steering committee responsible for project planning, policy establishing, resource providing and environment constructing. Four committee members from China side were President Yunhe Pan, Dr. Wen Gao (PI), Prof. Chi Huisheng (Vice President of Peking University), Prof. Hu Dongcheng (Vice President of Tsinghua University). Four committee members from China side were Dr. Raj Reddy (Co-PI), Dr. Ching-chih Chen (Co-PI), Dr. Victor Zue (Professor of MIT) and Dr. Gloriana St. Clair (Director of CMU Library).

During the 2nd steering committee meeting hosted by China in Mar. 2002, the US delegacy with 12 members visited the Chinese university libraries involved in the project and the two technical centers in the Graduate School of Chinese Academy of Sciences and Zhejiang University. India became the 3rd partners of UDL in May. 2002. The 3rd steering committee meeting was hosted by India in Jan. 2003. The Chinese delegacy with 9 members led by Dr. Wen Gao and the US delegacy led by Dr. Raj Reddy visited India partners and the President of India.

**Collections**

- [CMU-SCS](#)
- National Publishing House, USA
- Satish Dhawan's Library

**TOP 20 Bittings Book**

- 1 [The higher civil service : An evaluation of Federa](#) 406
- 2 [European monetary unification : and its meaning for](#) 206
- 3 [A world of nations Problems of political moderniza](#) 181
- 4 [Our troubled hemisphere: Perspectives on United St](#) 156
- 5 [Latin American economic integration and U.S. polic](#) 140
- 6 [The assistant secretaries: Problems and processes](#) 137
- 7 [Do the poor want to work? A social-psychological](#) 136
- 8 [American multinationals and American interests](#) 134
- 9 [Force without war : U.S. armed forces as a politic](#) 115
- 10 [Why growth rates differ : postwa experience in ni](#) 115
- 11 [Spear and sceptor; army, police, and politics in tr](#) 115
- 12 [The Neglected Aspect of Foreign Affairs](#) 115

Figure 3. Website on the *China-US Million Book Digital Library Project*

Since extensive background information as well as the current resources is available from the Web site, discussion in this paper will be limited. It is important to state that the *Million Project's* Chinese partners include:

- 6 Phase I institutions
  - Chinese Academy of Sciences (Northern center),
  - Zhejiang University (Southern center),
  - Fudan University,
  - Nanking University,
  - Peking University, and
  - Tsinghua University.
- 8 Phase II institutions
  - Beijing Normal University,

- Chengdu University,
- Chinese Academy of Sciences Library,
- Jilin University,
- Shanghai Jiao-tong University,
- Wuhan University,
- Xian Jiao-tong University, and
- Zhong-san University.

As shown in Figure 4, these over a dozen major academic institutions are spread out in different parts of China. Each participating university-based Digital Library Center has agreed to choose and offer its special, unique, and valuable information resources - books, paintings, sculptures and cultural objects - for this digital project. Each participating center has been digitizing some of these resources under the sponsorship and supervision of the 211 of the planning Committee of the Ministry of Education of China.

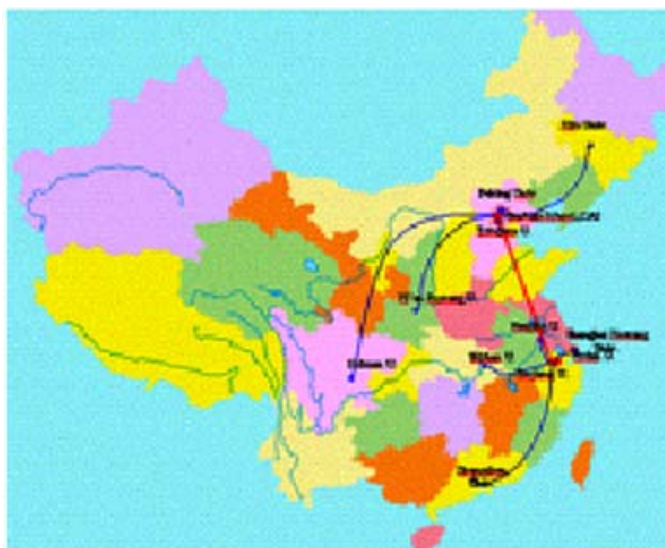


Figure 4. Locations of the Chinese *Million* institutions with regional centers (courtesy of China *Million* Project)

The Chinese partners has called themselves CADAL (China-America Digital Academic Libraries) and the activities of particularly Phase I institutions have been reported by Huang Tiejun and Gao Wen [7]. In 2004, Phase II institutions have come on board. I was privileged to have the opportunity to visit 5 of the 8 institutions in September 2004, and am pleased to report that during the first 9 months, encouraging results can already be seen with scanning centers in place and active digitization activities going on. Currently, each Million institution in China is diligently scanning their theses and dissertation collections, while concurrently also those local historical and cultural collections unique to each institution.

In addition to the Chinese collections, Carnegie Mellon University has arranged also to send materials in non-Chinese languages, mostly in English, to be scanned by the Scanning Center in Shenzhen, Guangdong. The results of these digital resources together with those available from other Million projects, like *US-India Million Book Digital Library Project*, are available for public use through the Internet at the Universal Library at CMU (<http://www.ulib.org/>) and the Internet Archive (<http://www.archive.org/>) in the US, as well as sites in China and India.

### **GLOBAL MEMORY NET**

*Million Project* described above is mainly text-based. The project, *Global Memory Net (GMNet)*, to be discussed and described next, is mainly multimedia with focus mainly on digital images at the moment, but will expand to include digital videos, music, voices, etc.

Although *GMNet* has been covered rather extensively in different parts of the world in recent months [2, 3, 4], because of the difference in audience, they will be presented again in order to stress the great potential of this project for US-China collaboration.

## From PROJECT EMPEROR-I to *Chinese Memory Net*

In the early 80s, the *PROJECT EMPEROR-I*'s by-product 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 [8]. This NEH funded project has collected thousands and thousands of invaluable images and multiple hours of videos of incredible value to scholars and general citizens. After the NEH funding was over, conscientious effort in building up more contents and more complete descriptive information (later known as metadata) of the image resources continued at a time when US National Science Foundation introduced and funded the First and Second Phases of Digital Libraries Initiatives (DL-I and DL-II). In 1999, when NSF first introduced its International Digital Library Program (NSF/IDLP), *Chinese Memory Net (CMNet)* was one of the first NSF/IDLP Projects [5].

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. Thus, extensive efforts were made to develop collaborative infrastructure with collaborators in:

- Beijing - Peking University and Tsinghua University;
- Shanghai - Shanghai Xiao-tong University;
- Taipei - National Tsinghua University, National Taiwan University, and Academia Sinica,
- US - Carnegie Mellon University (CMU), and Penn State University (PSU).

Originally *CMNet* hoped to bring collections of various distributed digital library systems on Chinese related topics together, with a potential home page something like Figure 5. This has proven to be both difficult and unfeasible. In the short four years, it has made progress in developing collaborative infrastructure for digital library development. For example, both *CMNet* and the *NIT 2001 conference* in Beijing, organized by me and mentioned earlier, have played important role in fueling the development of digital libraries in China and partially in Taiwan. For example, *CMNet* has helped our collaborators in both Mainland China and Taiwan to obtain funding supports from the respective governmental sources – counterpart of the US National Science Foundation, and thus has helped to initiate some of the significant digital library projects among our collaborators. Specifically, Tsinghua University's Architecture Digital Library was a good example.



Figure 5. Potential Home Page of Chinese Memory Net

Yet, it is fair to say that the over-all content development for *CMNet* beyond our own project effort related to the First Emperor of China's images has been slow and difficult.



On the other hand, our diligence and painstaking effort in creating metadata on the emperor images really paid off. The invaluable image and video resources as well as metadata have formed attractive basis for a number of exciting and productive technology-oriented collaborative works with computer scientists, who really need relevant real-life data to work with. Some of the collaborative research activities are listed in the following with more elaboration on some of the activities in later part of this paper:

- Open Archive Initiative (OAI) research – with the collaborators in China since each is using quite different metadata.
- Intelligent agent and text-based image retrieval – collaborate with Prof. V. W. Soo of the National Tsinghua University in Hsinchu, Taiwan [9, 10],
- Semantic sensitive content-based image retrieval – collaborate with Prof. James Z. Wang of Penn State University [11, 12]
- Digital video using the Informedia technologies – collaborate with Prof. Howard Wactlar of Carnegie Mellon University [13], and
- Machine learning for annotation - collaborate with Profs. James Z. Wang and Jia Li of Penn State University [14].

### From *Chinese Memory Net* to *Global Memory Net (GMNet)*

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 our activities of *CMNet* with the core contents related to the images and video related to the First Emperor of China. In the first two years of *CMNet* (2000-2002), we made considerable progress in the use of cutting-edge technologies in the organization and retrieval of multimedia contents, specifically the digital images. The success in the technical application area has attracted considerable interest and thus resulted in collaborative activities with several major institutions in different countries other than China. This made the expansion of the scope of *CMNet* to *GMNet* since 2002 a natural necessity. For example, *Project Restore* is an exciting collaboration between University of Florence and *GMNet*. It involved several thousands of images of significant artifacts in Italy which were badly damaged over time or by water, heat, etc. and restored with the incredible nano-particle chemical technology of the University of Florence [15]. Figure 6 is an excellent example. For images like this, they don't

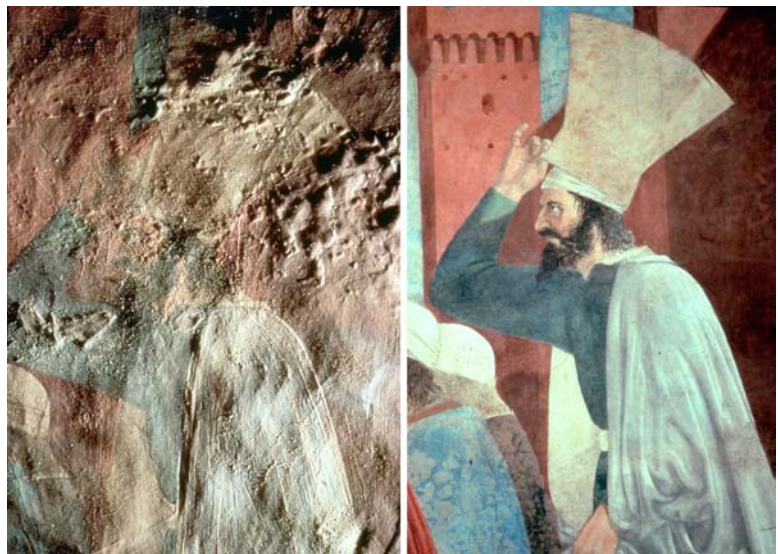


Fig. 6. Damaged artifact showing pre- & post-restoration images



belong to *CMNet*. They have to be properly included under “Italian Memory.”

With the expansion of *CMNet* to *GMNet* in 2002, we can now cover the ‘memory’ of any part of the globe [2, 3, 4] in addition to those Chinese memories. Also, *GMNet* is now having a more accurate vision by truly providing capabilities to bringing all distributed digital library systems together rather than the earlier objectives “to bring collections of various distributed digital library systems on Chinese related topics together. In other words, even we do not have the actual collections; we can point to the collections once a relevant image is retrieved.

Figure 7 is a tentative *GMNet* homepage. It shows clearly that there is a space holder for all countries in the world although this tentative homepage has listed only a few continents and countries under each in the Geographical category.



Figure 7. Tentative Home Page of the *Global Memory Net*

By expanding *CMNet* to *GMNet*, this expedites the digital library collaborative development and frees the R&D activity from unnecessary logistical delays and inflexibilities. Since there are over 200 countries in the world, there are endless opportunities for digital collection development, digital partnership, and collaborative research activities. There is currently a long list of topics, such as those listed in the following, and the list is growing longer quickly:

- China - Chinese painting, many historical unique collections, architecture, historical site, historical figures, etc.
- Cambodian - Ancient temples, etc.
- Japan – Temples etc.
- India - Architecture, palaces, temples, goddess, etc...
- Thailand – Palaces, etc.
- Vietnam - Historical development of the former Saigon
- Italy - Historical artifacts, art objects
- Europe - Cathedrals, Castles, etc.
- World - Global musical instrument
- World digital collections, national libraries, etc.

Figure 8 offers a quick visual look of a small portion of these topics.



Figure 8. Selective topics of *Global Memory Net*

In addition, the current direction also include the possibilities of *GMNet* serving as a functional multimedia gateway or portal to world invaluable “memory” resources available in all types of resource organizations - libraries, museums, archives, academic institutions, etc. This offers incredible opportunities for easy universal access of world’s treasures [16, 17]. *GMNet* offers users the world – and not just “China” -- instantly! [18]

## ***HOW GLOBAL MEMORY NET OFFERS THE WORLD!***

The name of *Global Memory Net* clearly articulates both the potential coverage and scope of this project [2, 3, 4]. It is global coverage. Valuable information can be accessed via the “Geography” category. Plan is being made to provide world map to permit users to access to any country or area by clicking on the proper location of the map. They can also be found via specific project, like the Emperor Project, Project Restore, etc.

Although currently *GMNet* concentrates its efforts in the cultural, historical, and heritage types of “memory,” this is more because of the project starts with the large number of Emperor images. Similar methods and techniques can be used to initiate global scientific or medical memories as well. In this regards, we look to our content collaborators to develop based on their interests and needs.

Clearly *GMNet* supplements well *US-China Million Project* described earlier. *Million Project* is still currently mainly text-based, while *GMNet* starts with images, and are moving to digital videos, music, and other multimedia formats. The only textual information is related to annotations and descriptive information included in metadata, as well as actual reference materials which the retrieved images will be linked to. Although every single element of the metadata can be retrieved, but the “cutting-edge” way of retrieving images is not through text-based retrieval. For this reason, the following discussions will offer mainly examples related to the cutting-edge content-based image retrieval of digital images, and with only limited mention of the digital video potential. Since we are talking about US-China cooperation, I shall make an effort to choose images related to China in this presentation.

For images of the First Emperor of China’s terracotta warriors and horses and those collections with substantially large number of images, *GMNet* is a comprehensive image digital library on those subjects. For many world’s cultural and heritage contents with only small number of images, *GMNet* serves as an effective digital portal which offers the world instantly to the information seekers, and then once the user selects the desired images retrieved, he/she can be referred to the relevant site directly for more information.

### ***Image Retrieval***

It is impossible to describe all the features of *GMNet* in a short introduction. I shall present them briefly here with new examples since descriptive information is available in previous keynote and invited speeches [2, 3, 4, 19]. In the simplest way, one can just imagine taking a visual tour of a selected cultural, heritage, and historical topic all while sitting at one’s computer. This soon to be available *GMNet* on the Internet will provide image retrieval capabilities with considerable textual supports in a way not possible before. For example, from the page like that shown in Figure 7, if one selects the Emperor collection, 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 by conducting either the traditional search using Google protocol if predefined specifics of the images are known, or by the cutting-edge semantically sensitive content-based image retrieval. On the other hands, if one likes to search other topics, one can choose that by country or by project. Since searches for Emperor images have been presented extensively before, this paper will provide examples on topics other than Emperor.

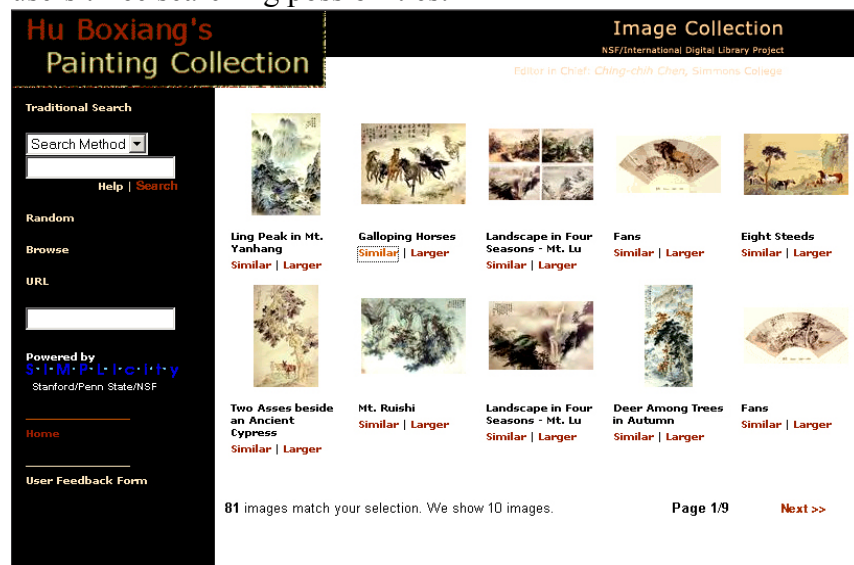
- Traditional Image Search

When one knows what he/she is searching for, one can search literally every field of the metadata as mentioned already, such as creator, title, location, time period, description, keyword, reference source, etc by using the Google syntax. In this approach, keyword search is likely to be the most popular one. Thus, if “keyword” search is selected, and the search terms are types by using the Google syntax with as “+” indicates the “required” term. Almost instantly from the thousands of images in the image base, the search will present search results showing the first 10 images located meeting the search requirement first. In this type of searches, precise retrieval of available images is made.

- Semantically Sensitive Content-Based Image Retrieval

However, in most cases, one generally does not have any idea on what kind of images are available in *GMNet* except that it is international in coverage. Just like in a library, we need to provide the user an opportunity to browse the stack, and find what they need and want. Currently, most image databases do not offer the users the chance to browse. In the case of *GMNet*, one can use 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 funding [11, 12]. This allows users to browse, retrieve, enjoy, and learn in just seconds through multiple thousands of digital images as described in the following:

**Examples 1: Hu Bo-xiang’s Painting Collection** as shown on the right cover, the search screen of Hu’s painting will be displayed, as shown in Figure 9. One notices immediately that two ways of image retrieval methods are provided on the left panel -- “Traditional Search” first permitting the requester to search any desired terms under every metadata field. It then followed by three buttons which provide the users three searching possibilities:



- Random – by clicking on this, images in the image base will show up on the right panel randomly as shown;
- Browse – by clicking on this, users will be able to browse images 10 or 15 at a time from page to page until they spot the desired image.

Figure 9. Random showing of images from Hu’s Collection  
(Courtesy of Hu’s daughter, Prof. H. S. Hu)



- URL – the user will be able to ask the system to find images that are similar to the one located on a given URL address on the Web.

Until recently, most archival images were not available in digital form. Now we have together in one place a large quantity of invaluable digital materials from multiple countries. One can ask the system to bring out image icons randomly, or to browse the images by displayed icons page to page until one locates the image of interest. For example, when the icons of the mages of the Hu's painting collection are displayed randomly in Figure 9, one spots a “galloping horse” image on the second to the left of the first row of particular interest. In this case, one can ask the system to provide all images “similar” to the one chosen by simply clicking “Similar” without typing any word, *GMNet* will display in seconds all the images similar to the one selected (Figure 10).



Figure 10. Four images with “horses” are shown

This opens up all possibilities for all related maps which are totally unknown to the user prior to the showing (see Figure 10).

Once the related images are displayed, one will be able to find instantly more textual descriptive information as well as reference sources and in some case, full-text descriptions on a chosen image by clicking “Info” as shown in Figure 11. If the chosen image needs to be enlarged, then click on “larger,” and multiple levels of zooming will be possible to show the desired details of the image. Concurrently, dynamic digital water mark



Fig. 11. Requested descriptive information of the image

will be instantly generated at any zooming level to offer the “ownership” information of the image (see description on the next example).

## Example 2: Library Congress’s Naxi Collection

This is an exciting development! The richness and uniqueness of global collections at the Library of Congress requires no further description. Naxi Manuscript Collection is the only unique one of its kind in the world, and it is owned by Library of Congress and housed in its Asian Collection. The collection is accessible via the Web (Figure 12), and one can search information in its traditional way by searching keyword, subject, title, etc. as shown in Figure 13.

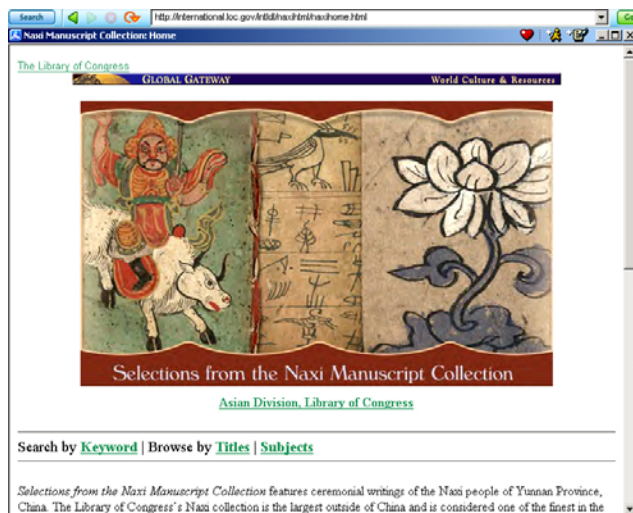


Figure 12. LC’s Naxi Home Page

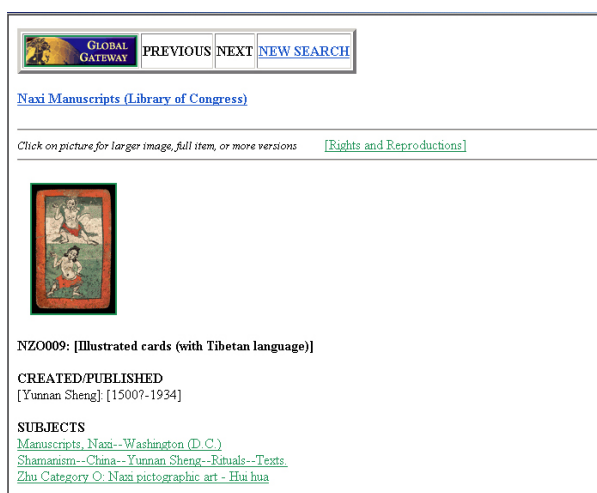


Figure 13. Image retrieved by keyword searching

With the enthusiastic support of the Head of its Asian Division, Dr. Hwa-wei Lee, *GMNet* has the privilege to include the unique and beautiful images of the Naxi manuscripts’ collection of the Library of Congress. Although one can access the Naxi collection as shown in Figures 12 and 13, our approach provides an unusually easily access to this unique image collection not possible before. Instead of showing the retrieved image one at a time, we provided the users a glimpse of all images available (see Figure 14) [22], and when one locates one of interest (the



Figure 14. All types of Naxi images are randomly shown

upper left one, “illustrated card with Tibetan language”, similar images can be requested by a simply click on “Similar,” and all images are displaced at once quickly (Figure 15).



Figure 15. All similar images related to “Illustrated cards with Tibetan language” are shown

This truly opens up all possibilities for all related images which are totally unknown to the user. Once these massive numbers of images are displayed, one would be able to enlarge a chosen image – say the middle of the first row - by clicking on “larger,” and multiple levels of zooming will be possible and dynamic digital water mark will be instantly generated to offer the “ownership” information of the image as shown in Figures 16 and 17.



Figure 16. Chosen image is enlarged with digital water mark

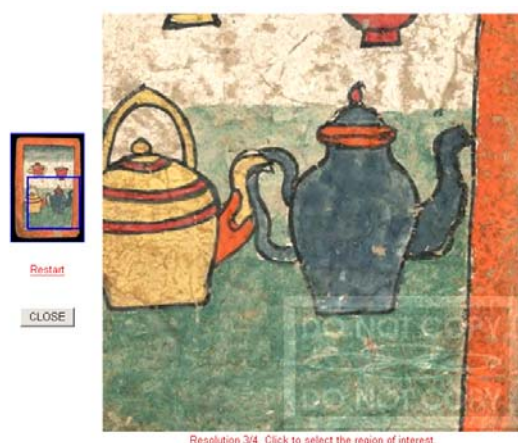


Figure 17. Portion of the image is enlarged more



One will be able to find more textual descriptive information as well as reference sources (see Figure 18) and in some cases, full-text original source on a chosen image instantly by hyperlinking.

### Example 3: World Digital Collection and UNESCO's *Memory of the World*

We have currently identified over 1400 digital collections in the world. We are able through both traditional search and content-based retrieval techniques to single out all 90+ digital collections from 45 countries registered under *Memory of the World* (Figure 19). Once a user has identified the desired collection, information on the site can be located and linked instantly. For example, the first image on the far left of the first row of Figure 19 is titled "Records of the Qing's Grand Secretariat. It is one of the two listed under China for the *Memory of the World*. We can link to that website instantly. Thus, our digital portal has certainly boosted the accessibility and value of these collections instantly.

Currently, we are exploring closer and more substantive collaboration with the Unesco's *Memory of the World* Programme.

**The LC's NAXI Collection**

**Title:** [Illustrated cards (with Tibetan language)]

**Description:** ""Selections from the Naxi Manuscript Collection features ceremonial writings of the Naxi people of Yunnan Province, China. The Library of Congress's Naxi collection is the largest outside of China and is considered one of the finest in the world. The Naxi use a unique pictographic writing system that is similar to the ancient Egyptian and Mayan writing systems. It is the only living pictographic language in the world today. This online presentation features 185 manuscripts, a 39? -foot funerary scroll and an annotated catalog of the entire collection."" - from Library of Congress Asian Division's Naxi Home Page, <http://memory.loc.gov/intldl/naxihtml/naxihome.html>. ""The Library's collection of 3,342 Naxi manuscripts is the largest collection outside of China and is unrivaled in quality, quantity, and variety among Naxi collections in Europe, the People's Republic of China, and Taiwan. The Library's collection contains many unique examples of the only living pictographic language in the world today. Naxi pictographs differ from Chinese characters in that they appear more like Egyptian or Mayan hieroglyphs, with many recognizable figures of animals and objects. Because of the uniqueness of the language, studying the manuscripts can be quite difficult for scholars."" 525 1500?-1934

**Reference:**  
<http://memory.loc.gov/intldl/naxihtml/naxihome.html>

[Close Description](#)

Figure 18. Descriptive information on the Naxi image

**The World DigiMem Collection**  
 NSF/International Digital Library Project  
 Editor in Chief: Ching-chih Chen, Simmons College

**Traditional Search**  
 Title:   
 Memory of the world  
[Help](#) | [Search](#)

[Random](#)  
[Browse](#)  
 URL:   
 Powered by [S.I.M.P.L.E.R.](#)  
 Stanford/Penn State/NSF  
[Home](#)  
[User Feedback Form](#)

**Image Collection**

 Memory of the World: Records of the Qing's Grand Secretariat [1999] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Warsaw Ghetto Archives (Emanuel Ringelblum Archives) [1999] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Anthology of Great Buddhist Priests? Zen Teachings [2001] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Illuminated Codices from the Library of the Bratislava Chapter House [1997] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Basagic Collection of Islamic Manuscripts [1997] <a href="#">Info</a>   <a href="#">Larger</a>
 Memory of the World: Holy Koran Mushaf of Othman [1997] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Codices from the Daxaca Valley [1997] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Kandilli Observatory and Earthquake Research Institute Manuscripts [2001] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Dainu Skapis - Cabinet of Folksongs [2001] <a href="#">Info</a>   <a href="#">Larger</a>	 Memory of the World: Kalman Tihanyi's 1926 Patent Application "Radioskop" [2001] <a href="#">Info</a>   <a href="#">Larger</a>

Figure 19. Screen images of some of the 90+ Memory of the World Projects

As one of the oldest countries in the world, China's 5000-year cultural, historical and heritage resources are truly rich and abundant. Millions and millions of cultural resources have been passed on from one generation to the other. Yet, most of these rich resources remain unknown to the world and certainly difficult to introduce and expose them to those outside China. *GMNet* provides an effective avenue to do just that since images of these invaluable resources can now be accessed easily by interested peo-



ple throughout the world. It is a perfect information and knowledge delivery channel. It is important to note again that the dynamic digital water marks will automatically appear when these images are shown in any sizes larger than the thumb nails. This not only will protect the intellectual properties of the creators and/or owners of the artifacts, but also will discourage any illegal copying of the images.

Thus, the potential for US-China collaboration in this area is truly great! While we are always open to any possibilities for collaboration with interested institutions, it is also encouraging to note that our work seems to have no end in sight. Countries like China, India, Greece, Egypt, and Italy have also many open “living museums” which permits us to start our work even without any collaborator. For example, when I visited Chengdu in September 2004, I was able to gather images on Sanxingdui (Figure 20) and Dujiangyan. This is true wherever I go in the world, like Dubrovnik, Hanoi, Florence, Athenes, Bangkok, Niles, Jerusalem etc. *GMNet* can start many topics before the content collaborators are on board.

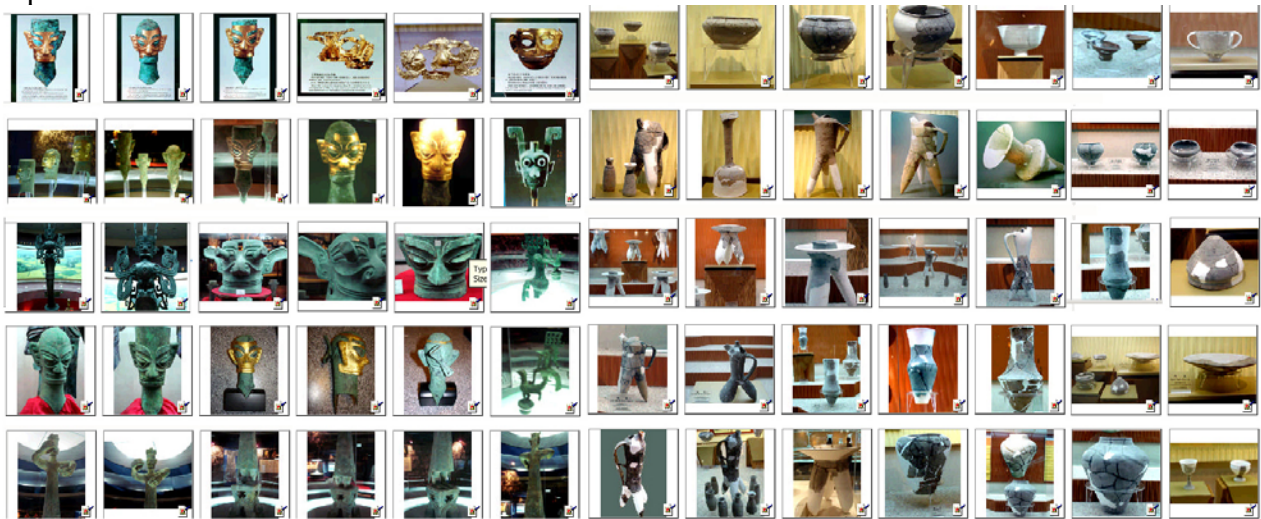


Figure 20. Images from Sanxingdui in Chengdu dated back 4800 years ago!

For more information on *GMNet* before the website is available for public use, visit [www.memorynet.org](http://www.memorynet.org)

## ***FUTURE DEVELOPMENT***

In addition to continue the building of a great variety of image collections and global partnership, future development will move more aggressively to the areas of digital video, sound and audios.

Carnegie Mellon University’s well-known Informedia Project is one of the six original NSF/DLI-1 projects. It has continued its further development in digital video related technologies and tools ever since 1995. Collaboration between Informedia and *CMNet* has enhanced perspectives from cultural and historical video documentaries. Its multi-lingual (English and Chinese) has also posed challenges in its speech recognition related research [13]. When the Informedia technology is ready for web-based use, *GMNet* will be ready to use it. Figure 21 shows some of the screens generated from the latest collaboration. Upper left shows that when “emperor” is searches, 60 video segments with that word have been identified and can be retrieved as shown in the left middle screen, these segments can be visualized in timeline as shown in the lower left screen.

Map is shown in the upper right screen, and when one of the video is chosen, the video will play in the upper right of the lower right screen, and below that, the actual text will also be displayed with the word “emperor” highlighted in red. The running bar between the video segment and the textual annotation shows the red line(s) where the word “emperor” will appear when the video playing reaches the indicated area(s).

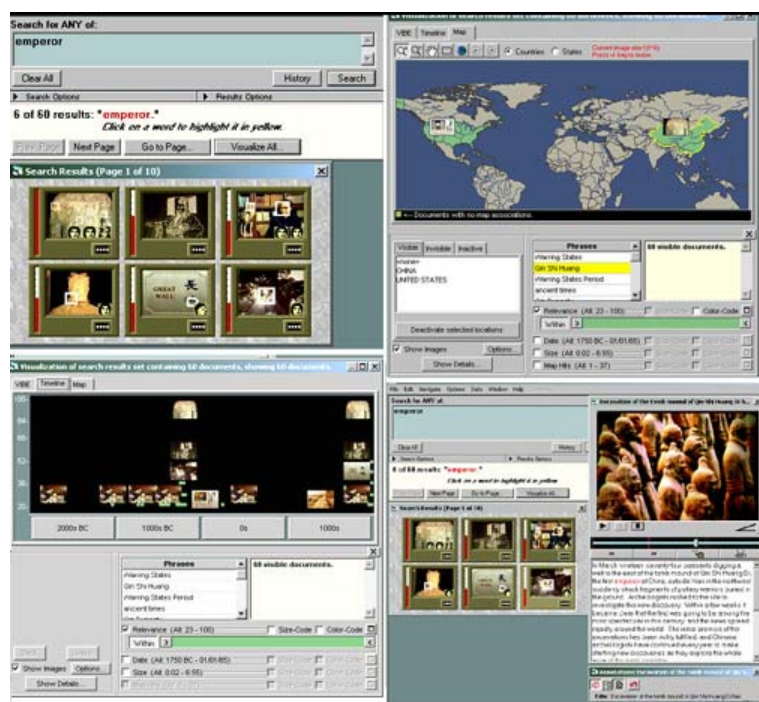


Figure 21. CMU's Informedia and Emperor's digital video

In addition to digital videos, our research will also explore the potential use of sound and music. One of the perfect starting points will be with the world's musical instruments. Figure 22 shows

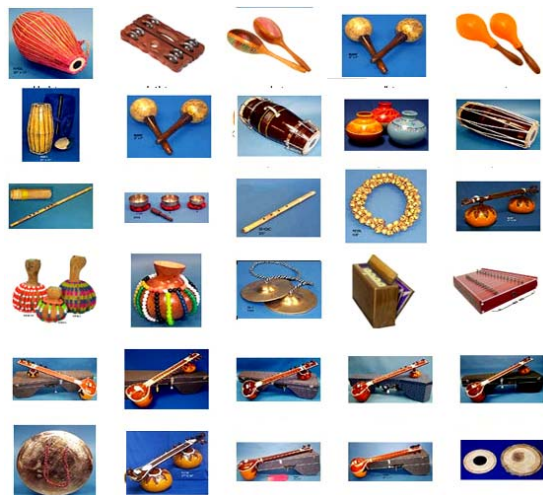


Figure 22. Some of the world musical instruments that such an image base is being constructed. It is our hope that the instruments will also be linked to music and sound when available. Another possible area would be with the language learning and writing. In all these areas, there are great possibilities for US-China collaboration!

One final mention of an exciting activity would have to be my other NSF/IDLP [NSF/IIS-Special Projects (IIS-0333036)] 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 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. This research will also capitalize the existing rich descriptive annotation for research purposes. In addition to Ontology-based image retrieval, the project will deal with machine-learning-based and content-based image retrievals, as well as the difficult object-based partial image searches. We also hope to extend research to include intellectual property (IP) protection technique.

## CONCLUSION

During 1997-2002, I was privileged to serve on the US President's Information Technology Advisory Committee (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."* [21]

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.

In considering international digital library research and development, it is important for us to revisit the conceptual model presented by the *DELOS/NSF Working Group on Digital Imagery for Significant Historical, Cultural and Heritage Materials*, of which I am a US co-Chair (Figure 23) [22]. From this model, it is clear that *GMNet* is developing substantial multimedia contents – currently mostly images -- both in house as well as linking them in distributed systems together through the use of the global network. The retrieval of these contents is using both the existing as well as cutting edge technologies. They are made available for use by general public as well as scholars and researchers through via the Web. This paper clearly addresses mainly the "content" aspects with mentions to the technologies utilized.

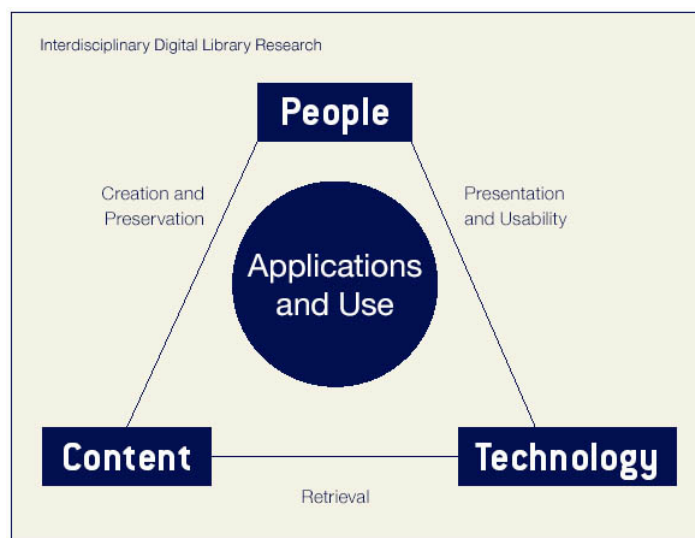


Figure 23. Conceptual model for digital library R&D

As to "contents," 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, 4]. 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.

It is gratifying that in the short couple of years, *Global Memory Net* has demonstrated how international collaboration and community building in promoting large-scale content building, coupled with new technological tool and method development, can indeed offer users the world in a way not possible before. The potential for delivering and marketing invaluable world multimedia resources as well as for US-China cooperation should also be clear. The best is yet to come!

In the last four years, we have learned a number of lessons regarding “cooperation.” Real cooperation means more than just superficial willingness and rhetoric. It requires real commitment and willingness to iron out both logistical and technical difficulties. In the “give” some and “take” some environment, one will find that sharing is really a win-win situation for all! *Global Memory Net* has shown the potential for collaboration and is ready to do more! We welcome more US-China collaboration!

### **ACKNOWLEDGMENTS**

*Chinese Memory Net* and *Global Memory Net* are supported by the US NSF/CISE/IIS/IDL Program 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 in collaboration with James Z. Wang of the Penn State University whose work is supported by the NSF/CISE/ITR program under Grant no. IIS-0219271. The new image management work is funded by the US NSF/ CISE/IIS/IDL under grant no. IIS-0333036.

### **REFERENCES**

1. Chen, Ching-chih. “Global Digital Library and Universal Information Access,” in *Proceedings of the First China-U.S. Conference on Global Information Access: Challenges and Opportunities*, held at the National Library of China, Beijing, China, August 21-23, 1996.
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. pp. 43-56 (English version) and pp. 57-67 (Chinese version).
4. 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.



5. Chen, Ching-chih, "Digital library development in the US and global reach: Roles of federal government agencies and PITAC," *Proceedings of the Digital Library - IT Opportunities and Challenges in the New Millennium: An International Conference*, Beijing, July 9-11, 2002. Beijing, China: National Library of China, 2002. pp. 17-27 (Chinese version). pp. 28-41 (English version).
6. Chen, Ching-chih, "Technological potentials for the Global Library: Realities and challenges," *Proceedings of the International Conference on National Libraries: Towards the 21<sup>st</sup> Century*, Taipei, April 20-24, 1993. Taipei, Taiwan: NCL. pp. 859-868.
7. Huang Tiejun and Gao Wen, "China-US Million Book Digital Library Project: Progression and it's digital book specification," in *Proceedings of CCDL: Digital Library – Advance the Efficiency of Knowledge Utility*, Beijing, September 5-8, 2004. Beijing: National Library of China, 2004. pp. 240-244 (English version), and pp. 245-249 (Chinese version).
8. Chen, Ching-chih. *The First Emperor of China*. Santa Monica, CA: The Voyager Company, 1991. This is a videodisc product and is the popular version of this title [A Cindy Award videodisc product]. Chen, Ching-chih. *The First Emperor of China: A multimedia digital CD*. [Mac and Windows dual version]. New York: The Voyager Co., September 1994. *MacUser* (October 1994) selected this as one of the "50 Best CDs of 1994".
9. 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.
10. 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. *Proceedings of the ACM/IEEE JCDL Conference, 2003*. 12 pages.
11. Chen, Ching-chih and James Z. Wang, "Large-scale emperor digital library and semantics-sensitive region-based retrieval," *Proceedings of the Digital Library - IT Opportunities and Challenges in the New Millennium: International Conference*, Beijing, July 9-11, 2002. pp. 454-462.
12. 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," *Proceedings of the ACM Multimedia, Workshop on Multimedia Information Retrieval*, Juan Les Pins, France, December 2002. 6 pp.
13. 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.

14. Wang, James Z., Kurt Grieb, Ya Zhang, Ching-chih Chen, Yixin Chen, and Jia Li, "Machine annotation and retrieval for digital imagery of historical materials." Submitted to *International Journal on Digital Libraries' Special Issue on Multimedia Centents*, 2004.
15. Baglioni, Piero, Rodorico Giorgi, and Ching-chih Chen, "Nanoparticle technology saves cultural relics: Potential for a multimedia digital library," presented at the *DELOS/NSF Workshop on Multimedia in Digital Libraries*, Chania, Crete, Greece, June 2-3, 2003. [http://memorynet.org/pdf/baglioni\\_crete.pdf/](http://memorynet.org/pdf/baglioni_crete.pdf/)
16. Chen, Ching-chih, "Past forward - Digital media for cultural heritage: The case of the Global Memory Net," Invited lecture given at the *10<sup>th</sup> Annual Lecture of Informatics*, sponsored by Informatics, Bangalore, India, February 29, 2004.
17. Chen, Ching-chih, "Sharing the abundant human knowledge when old cultures meet in cyberspace: *Global Memory Net* for international collaboration," *U&I Special Issue* (Bangalore, India), pp. 4-5 (March-April 2004).
18. Chen, Ching-chih, "*Global Memory Net* offers users the world instantly," *Library Times International*, 21 (1): 1-4 (July 2004).
19. Chen, Ching-chih, "*Global Memory Net*: New collaboration, new activities and new potentials" in *Proceedings of the International Asian Conference on Digital Libraries*. Shanghai, China: IACDL 2004.
20. Chen, Ching-chih, "*Global Memory Net*: Digital portal for global resource sharing and closing digital divide," Poster presentation at the *70th IFLA General Conference and Council*, August 22th - 27th 2004, Buenos Aires, Argentina.
21. 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](http://www.hpcc.gov/pubs/pitac/pitac-dk-9feb01.pdf)
22. Chen, Ching-chih and Kevin Kiernan, eds. *Report of the DELOS-NSF Working Group on Digital Imagery for Significant Cultural and Historical Materials*. December 2002. Also, Ching-chih Chen, Howard Wactlar, James Z. Wang and Kevin Kiernan, "Digital Imagery for Significant Cultural and Historical Materials - An Emerging Research Field Bridging People, Culture, and Technologies," *International Journal on Digital Libraries*, 2004, to appear.

## Biographical Information of the Author



*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 produced the award winning interactive videodisc and multimedia CD entitled, *The First Emperor of China*. 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. The Proceedings of *NIT 2001*, held at Tsinghua University, Beijing, was published as *Global Digital Library Development in the New Millennium: Fertile Ground for Distributed Cross-Disciplinary Collaboration* by Tsinghua University Press in 2001.

Since 1993, she has been advocating the global digital library concept by linking libraries, museums and archives all over the world together, and this *Global Digital Library Initiative* has helped the development of digital libraries in numerous countries. Since 2000, she has led a NSF/International Digital Library Project, *Chinese Memory Net (CMNet)*. She is also co-PI with Prof. Raj Reddy of the *China-US Million Book Digital Library Project*. She is a member of the Advisory Committee of DELOS (European Digital Library Network) and co-Chaired the *DELOS/NSF Working Group on Digital Imagery for Significant Historical, Cultural and Heritage Materials*. She has been advocating the need for international consortium in making cultural and heritage digital contents accessible to users. To this end, *Chinese Memory Net*, serving as a model for archiving, content building in specifically image and video areas, as well as international collaboration, has grown now to be *Global Memory Net*, with collaborators from different part 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 1997-2002, she served as a member of the *US President's Information Technology Advisory Committee*.

A sought after international speaker, in 2004 alone, she was a keynote speaker at the *International Conference on Digital Libraries* in Delhi, India; the *Libraries in the Digital Age (LIDA 2004)*: International Conference, Dubrovnik and Mljet, Croatia; the *International Conference on Digital Libraries*, Beijing, China; the International Asian Digital Library Conference, Shanghai, China; and the Invited Annual Lecturer of the *Annual Lecture in Informatics in Bangalore*, India. She also delivered invited speeches in Mysore, India; Orlando, Florida; Xian, Guanzhou, and Haikou, China; and Yokohoma, Japan.