



Global Memory Net and the development of digital image information management system: Experience and practice

ZHANG Sheng-qiang¹, CHEN Ching-chih^{‡2}

⁽¹⁾Library, Sichuan University, Chengdu 610064, China)

⁽²⁾Graduate School of Library and Information Science, Simmons College, Boston, MA 02115, USA)

E-mail: cdzhangsq@hotmail.com; chen@simmons.edu

Received Aug. 5, 2005; revision accepted Sept. 10, 2005

Abstract: Global Memory Net (GMNet) is intended to be an effective gateway to the world cultural, historical, and heritage image collections from selected academic educational and research partners in the world. Much of these unique collections of great value to education and research are not currently accessible due to distance, form, and technical barriers. This project is to find new ways to enable users to access and exploit these significant research collections via global network. As GMNet is ending its first 5-year phase in October 2005, it has contributed substantially to the community building in digital library development by accommodating numerous collaborators and technical staff from various parts of the world to spend 3 to 5 months as a full-member of the GMNet team in Boston. They have come from different parts of China—such as Sichuan, Hainan, Shanghai and Xi'an; Croatia; and Hanoi, Vietnam. In addition to contribute to the overall system development and enhancement of system functionalities, they have brought valuable sample image collections of their own institutions/countries, and actually developed prototype collections as a part of GMNet. This paper describes the exciting and productive experience of the first of this visiting research group in developing the GMNet's Version 2.0 PHP-based system under Prof. Chen's overall supervision. It also describes both the system's technical level structure—user/Web-based application/data, and complex functionalities with multi-collection, multi-lingual, multi-modal searching capabilities; system management capabilities; as well as provisions for user uploads and retrieval for our own projects. This Version 2.0 system is built on the Linux/Apache/PHP/MySQL platform. What is described in this paper is an actual case which has formed a base for further new development by others in the research group. It demonstrates fully the value of the synergistic collaboration among global partners for universal digital library development. More information can be found in <http://www.memorynet.org/>.

Key words: Global Memory Net (GMNet), Collaboration, Digital image library system development, Management system, Cultural, Historical, Heritage collections

doi:10.1631/jzus.2005.A1216

Document code: A

CLC number: TP391

INTRODUCTION

Directed by Prof. Ching-chih Chen of Simmons College, Boston and supported by the US NSF/International Digital Library Program (IDL) since 2000, Global Memory Net (GMNet) is intended to be an effective gateway to the world cultural, historical, and heritage image collections from selected academic educational and research partners in the world. Much of these unique collections of great value to education

and research are not currently accessible due to distance, form, and technical barriers. This project is to find new ways to enable users to access and exploit these significant research collections via global network. Each collaborator of this complimentary and synergistic group possesses experience, knowledge, expertise, and capability in different but related research area(s). Each contributes either part of its superb culture and heritage collection, or cutting-edge techniques to facilitate the image retrieval.

Online users are provided with various search capabilities to enable them to retrieve the world cul-

[‡] Corresponding author

tural and heritage multimedia resources—images, sound, and videos. Although the initial GMNet's Version 1.0's search capabilities were developed by its collaborator, Prof. James Wang at Penn State University (PSU) using the collection content descriptions (metadata) and data files developed by Prof. Chen's research group and presented via html Web-based system developed at Simmons, the current systems (both Versions 2.0 and 3.0) are developed in house by the GMNet's own group on a Linux/Apache/MySQL platform and stored on an in-house server at Simmons together with the Web-based and image data files. The system is completed with its own traditional search and browse capabilities. Aside from the traditional search, it features the cutting-edge content-based image retrieval (CBIR) system, SIMPLiCity, originally developed by James Wang and Li Jia at Stanford, and then at PSU. The collaboration between PSU and GMNet is described in Chen and Wang (2002). As GMNet is ending its first 5-year phase in October 2005, it has contributed substantially to the community building in its digital library development by accommodating more than 7 visiting researchers, particularly with technical background, from various parts of the world to spend 3 to 5 months as a full-member of the GMNet team. The Chinese visiting researchers have come from different parts of China—such as Sichuan, Hainan, Shanghai, Xi'an, etc. In addition to contributing to overall system development and enhancement of the system functionalities, they have brought valuable sample image collections of their own institutions/countries, and actually developed prototype collections as a part of GMNet. By doing so, they have learned much from this participation as well as contribute to system enhancement and content development. It is an exciting experience in working together toward "universal access" of the world's valuable cultural, historical, and heritage resources.

This paper's first author, Shengqiang Zhang, is the Head of Library Automation of Sichuan University Library. He is the first of this visiting research group to arrive in Boston on April 11th, 2005, spending almost 5 months with GMNet under the support of the United Board of Christian Higher Education Fellowship and NSF. This paper describes his exciting and productive experience in developing the GMNet's Version 2.0 PHP-based system under

Prof. Chen's overall supervision. This Version 2.0 not only can fully leverage the features of SIMPLiCity's CBIR functionalities of Version 1.0, but also greatly expand the systems' traditional search capabilities, and link with broad-based search environments such as OCLC's World Cat, Google Scholar, and Million Project. It has fully utilized the DC-compliant metadata developed for each image collection by GMNet research staff; explored the use and retrieval of other multimedia files, such as sound and video; and developed the bilingual retrieval capabilities. Since late July 2005, Version 2.0 and Version 1.0 have been fully integrated into Version 3.0 (see Fig.1 for the latest version's Home Page of GMNet). The new design and integration of Version 3.0 was based on Version 2.0 and was mainly carried out by another visiting researcher from University J. J. Strossmayer in Osijek, Croatia—Boris Badurina.

SYSTEM DEVELOPMENT

Guided by GMNet's vision and conceptual framework as set by Prof. Chen (Chen, 2005a; 2005b; 2004a; 2004b; 2004c; 2004d), GMNet is a multi-purpose image knowledge base and portal to meet multiple needs of multiple users in the world, who are interested in cultural, historical, and heritage contents. For this reason, the multiple kinds of technologies, ranging from basic and cutting edge ones, are utilized to enable dynamic retrieval of the valuable resources as well as dynamic management of various system components which include collections, archives, information about "us", policies, users, evaluation, as well as news, as shown in Fig.1. Of all these, the most important component clearly is "Collections" because content is of overriding significance of any digital libraries, and technology is only the tool. In order to build more content, global collaboration is essential. For this reason, GMNet is truly a project where we experience the real convergence: (1) Content; (2) Technology; (3) Global collaboration.

From the system development point of view, it is a very challenging task to develop such a system. With countless discussions with Prof. Chen, the development and constant modifications has been a real and rather complex zig-zag process with a system structure which can be summarized in Fig.2.

This is a three-layer structure: data, application, and user:

(1) The data layer is the baseline layer, which stores all types of data, including all the image resource data, metadata describing these resources, archival data, as well as management related data on users, policies, and collaborators. These are stored in the server using MySQL.

(2) The Web service and applications layer houses all the Web-based applications developed us-

ing Linux/Apache/php, as well as the cutting-edge content-based image search programs, SIMPLicity/Wave Zoom, developed by Profs. James Wang and Li Jia of PSU.

(3) The user layer: the users can access all images and related resources of GMNet by using the developed applications via any Web browser.

With this system structure, various functions such as the following are developed with the relationships shown in Fig.3:

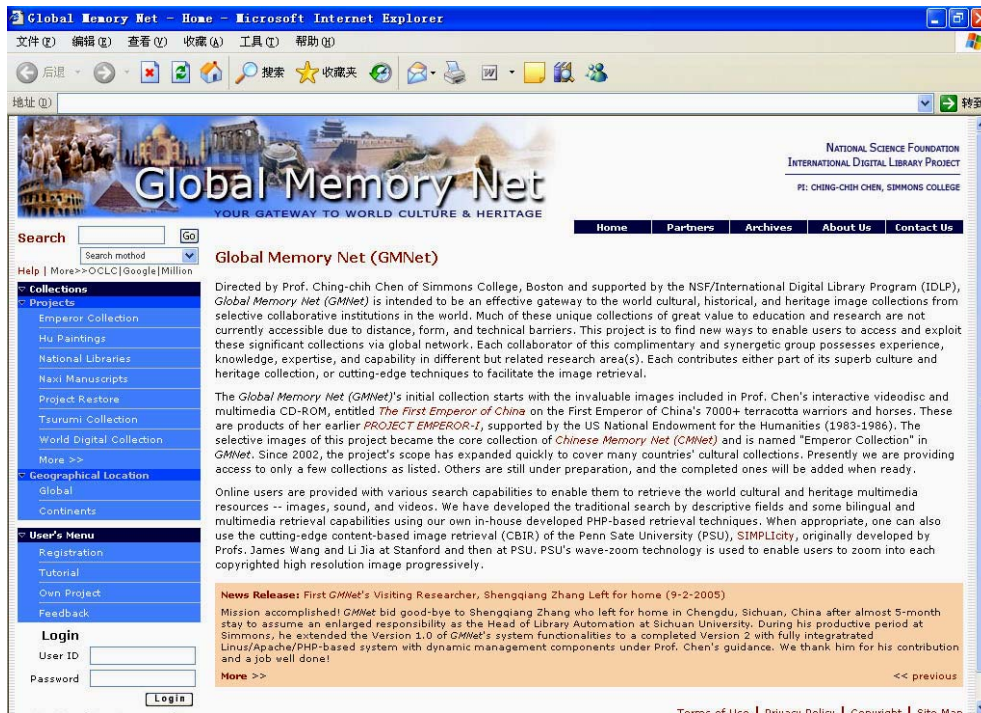


Fig.1 Home page of Global Memory Net's Version 3.0

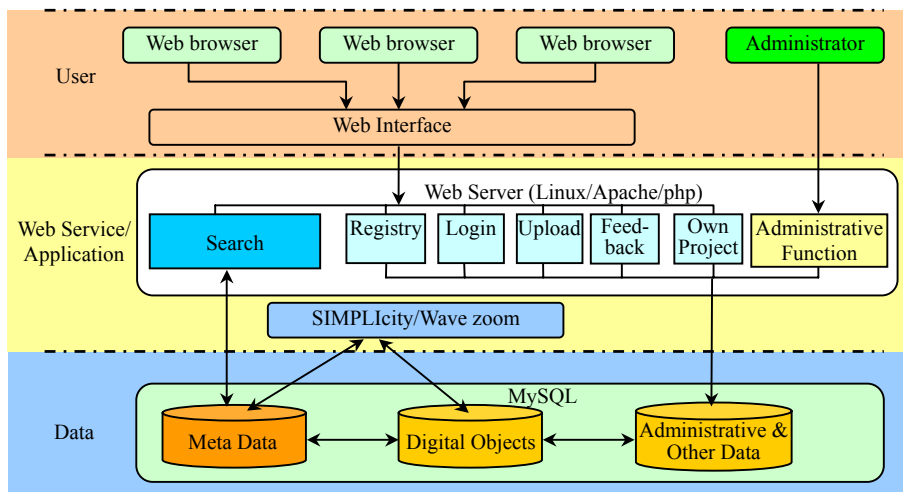


Fig.2 GMNet's system structure

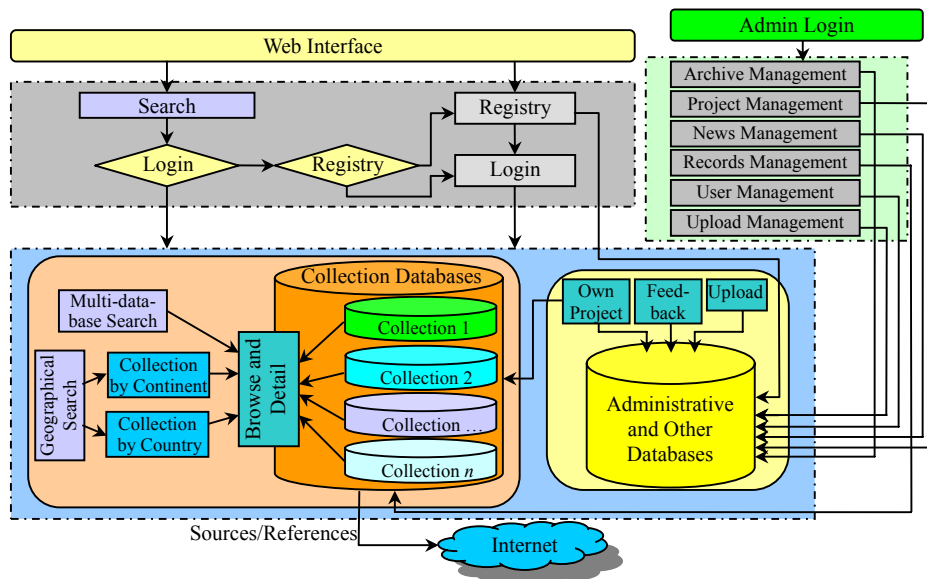


Fig.3 GMNet's system functional chart

- (1) Both single and multiple collection image search/retrieval using both traditional metadata-based or cutting-edge content-based methods.
- (2) System management.
- (3) User registration, evaluation, upload and own project development.

Although all components of GMNet are important and have unique functions, search and retrieval functions related to “Collections” are the most ones. As shown in Fig.4, individual image collections regardless whether they are GMNet’s own such as the Emperor Collection or those of the collaborators, such as Naxi Manuscripts of the Library of Congress’ Asian Division or Project Restore with University of Florence, or those 2000+ digital collected from around the world, can be searched by either traditional method—by the metadata fields, or by the cutting edge content-based image retrieval. In addition, collectively all images can be searched by using the in-house developed multi-database searching.

Since GMNet is international in scope, geographical search can increase the ease of use by users considerably. Thus functionalities to search collective databases by “global” and “continents” are provided:

- (1) Global: Here a user can select a continent or geographical area, and locate all the collections related to that continent or area.
- (2) Continents: Here the 230+ countries and geographical areas of the world are listed by continent, and a user can be referred to all collections related to a

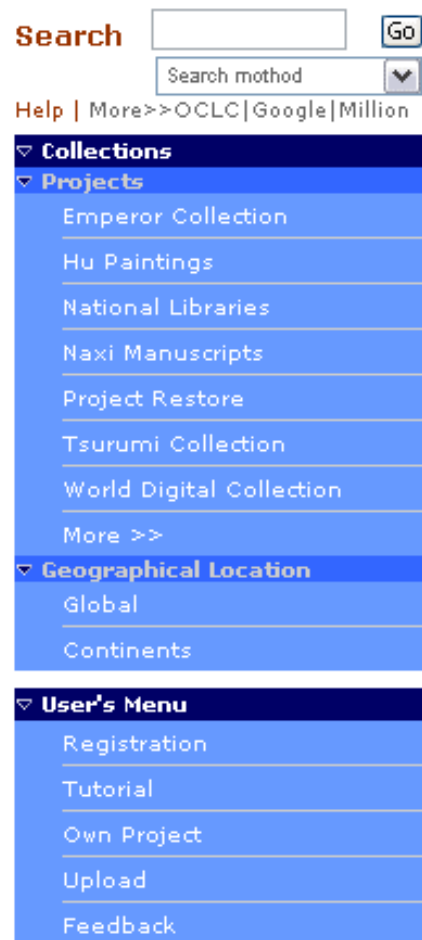


Fig.4 Left navigation bar showing dynamic collection listings and search components

country of choice. In addition to the provision of metadata related to the chosen collection, one will be linked instantly to the available webpage (URL) of that collection.

With the possibilities of linking distributed digital collections from most of the over 200 countries in the world, interoperability and ease of system use are of paramount importance in system design. In addition, the abilities to permit the system administrators to browse, search, and edit various image databases, news items, as well as user profiles are significant, thus Version 2.0's system development has included these functionalities as an essential component of the management system. Finally, with the enormously rich digital image resources of GMNet as well as those 2000+ global digital collections linked by GMNet, Version 2.0 provides an easy "user" component to enable users to both use and contribute to the collection.

CONCLUSION

The GMNet's system and content development is a continuing process. More system functionalities, multilingual and multi-modal capabilities, as well as different types of digital contents are added and increased at a fast pace, fully taking advantage of the collaborators' strength at various different levels. What it is described in this paper is an actual case of learning and development from one visiting researcher. The knowledge gained in this short-period R&D has been exciting and productive. It not only enhances the GMNet's system functionalities, but also enables further collaboration and development when the researcher returns to his own institution. The work completed has also formed a base for further new development by others in the research group. More information can be found in <http://www.memorynet.org/>.

ACKNOWLEDGMENT

The Global Memory Net is supported by the US National Science Foundation/International Digital Library Program (Grant No. NSF/CISE/IIS-9905833). We would like to thank the United Board of Christian Higher Education in Asia and the US National Science Foundation for the support of the activities of Shengqiang Zhang with GMNet.

References

- Chen, C.C., 2004a. Global Memory Net: New Collaboration, New Activities and New Potentials [Keynote Speech]. *In: Digital Libraries: International Collaboration and Cross-Fertilization. Proceedings of the 7th International Conference on Asian Digital Libraries (ICADL 2004)*, Shanghai, China. Springer-Verlag, Berlin, Germany, p.73-83.
- Chen, C.C., 2004b. Global Memory Net Offers the World Instantly: Potentials for Universal Access to Invaluable Chinese Contents [Keynote Speech]. *Proceedings of Digital Library—Advance the Efficiency of Knowledge Utilization*, Beijing, China. Scientific and Technical Documents Publishing House, Beijing, China, p.43-67.
- Chen, C.C., 2004c. The Promise of International Digital Library Collaboration for Innovative Use of Invaluable Resources [Keynote Speech]. *In: Human Information Behaviour & Competences for Digital Libraries. Proceedings of the Libraries in the Digital Age*, Dubrovnik and Mljet Island, Croatia. University J. J. Strossmayer in Osijek, Osijek, Croatia, p.7-15.
- Chen, C.C., 2004d. Global Memory Net: Potential and Challenges for Archiving and Sharing Cultural and Heritage Resources [Keynote Speech]. *International Conference on Digital Libraries*, Delhi, India. TERI (The Energy and Resources Institute), Delhi, India, 1:3-10.
- Chen, C.C., 2005a. New mode of universal access and Global Memory Net. *Journal of Zhejiang University SCIENCE*, 6A(11):1206-1215.
- Chen, C.C., 2005b. Digital Libraries and Universal Access in the 21st Century: Realities and Potential for US China Collaboration [Invited Speech]. *Proceedings of The 3rd China-US Library Conference*, Shanghai, China. National Library of China, Beijing, China, p.138-167.
- Chen, C.C., Wang, J.Z., 2002. Large-scale Emperor Digital Library and Semantics-sensitive Region-based Retrieval. *The Proceedings of Digital Library—IT Opportunities and Challenges in the New Millennium*, Beijing, China. Beijing Library Press, Beijing, China, p.454-462.