

Journal of Zhejiang University SCIENCE

ISSN 1009-3095

<http://www.zju.edu.cn/jzus>E-mail: jzus@zju.edu.cn

New mode of universal access and Global Memory Net*

CHEN Ching-chih

*(Graduate School of Library and Information Science, Simmons College, Boston, MA 02115, USA)*E-mail: chen@simmons.edu

Received Sept. 25, 2005; revision accepted Sept. 26, 2005

Abstract: In this digital era, we have witness the exciting convergence of content, technology, and global collaboration in the development of digital libraries. The mode of universal access for information seeking and knowledge acquisition differs greatly from the traditional ways. 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. “Universal access” has taken on a very different meaning when one has the ability to share invaluable resources through the use of cutting edge technologies. The author has experienced much of these transformations through her own R&D activities—from the creation of interactive videodisc and multimedia CD on the First Emperor of China’s terracotta warriors and horses in the 1980s and 1990s to leading a current international digital library project, Global Memory Net (GMNet), supported by the US National Science Foundation. In presenting her vision for linking world digital resources together for universal access, she will share with the audience the latest development of Global Memory Net.

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

doi:10.1631/jzus.2005.A1206**Document code:** A**CLC number:** TP391

INTRODUCTION

During 1998~2002, I was privileged to serve as a member of the US President’s Information Technology Advisory Committee (PITAC), and all members were deeply involved in the drafting of several PITAC Reports to the President. The one related to digital libraries was the 2001 Report of the PITAC Digital Library Panel, entitled Digital Libraries: Universal Access to Human Knowledge (PITAC, 2001). It offers an ambitious vision:

“All citizens anywhere anytime can use any Internet-connected digital device to search all of

human knowledge. Via the Internet, they can access knowledge in digital collections created by traditional libraries, museums, archives, universities, government agencies, specialized organizations, and even individuals around the world. These new libraries offer digital versions of traditional library, museum, and archive holdings, including text, documents, video, sound, and images. ... Very-high-speed networks enable groups of digital library users to work collaboratively, communicate with each other about their findings, and use simulation environments, remote scientific instruments, and streaming audio and video. ... In this vision, no class-room, group, or person is ever isolated from the world’s greatest knowledge resources (PITAC, 2001).”

If we dissect this abbreviated vision statement, and compare the segments with the title and several themes of the ICUDL 2005 Meeting, we should be clear on the targets for which we should be aiming. It is clearly stated in the PITAC’s vision that one should

* Modified from an invited talk, entitled “New mode of universal access: Latest development of Global Memory Net,” delivered at the Distinguished Seminar Series of OCLC, Dublin, OH on September 30, 2005. <http://www.oclc.org/research/dss/>. Some of the conceptual contents are also modified from earlier keynote speeches, such as that presented at the LIDA Conference, Dubrovnik, Croatia, May 25, 2004, entitled “The promise of international digital library collaboration for innovative use of invaluable resources” (Chen, 2004).

be able to find any information he/she needs whenever and wherever needed. It is a vision so much easier said than done! It will require sustainable and long-term commitment of many. We should expect many obstacles on the long road to this “elusive” vision, as stated in the Report of the US National Science Foundation’s (NSF) Workshop on Research Directions for Digital Libraries, entitled *Knowledge Lost in Information* (Larsen and Watclar, 2004).

Despite of the challenges, in the US, we believe that we have made substantial advances in the technical area in terms of advancing capabilities, through an interagency program of integrated, interdisciplinary, project-oriented research initiatives in the last decade. Now, the convergence of content, technology, and global collaboration in the development of digital libraries should be a natural process. In this digital environment, from content or information resources point of view, no one institution—no matter how large it is, or one country—no matter how abundant in resources, can possibly have everything. Thus, the old model of “owning” a collection has to give way to “sharing,” and the new emphasis has to shift from possessing large “physical libraries” to “virtual libraries” digitally distributed all over the world. If we are talking about content building in “global” scale, then we must have global collaboration through global community building. Thus, our usual “user community building” has to be of the global scale (Chen, 2004; 2005).

COMPONENTS OF INTERDISCIPLINARY DIGITAL LIBRARY RESEARCH

In 2002, I co-chaired the DELOS-NSF Working Group in Digital Imagery for Significant Cultural and Historical Materials (Chen and Kiernan, 2002; Chen *et al.*, 2005). We presented a conceptual model of interdisciplinary digital library research programs with a triangular relationship among people, content, and technology with the center area as “Applications and Use” as shown in Fig.1 (Larsen and Watclar, 2004).

Here, the scope and parameters set for “users” and “use” are:

(1) Users: ALL citizens of the world regardless age, ethnic group, education, social status, religion, etc.

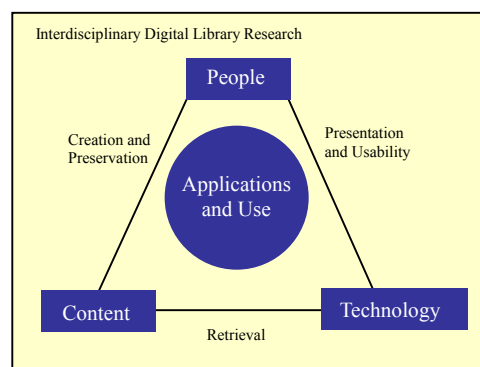


Fig.1 Conceptual framework of interdisciplinary digital library R&D

(2) Uses: For whatever they need.

Contents—All subject areas, all types of formats, all types of digital collections created by all types of organization;

Geographical areas—All parts of the world.

Content is the vast array of significant materials of subject areas throughout the world. Technologies are the enabling research and development in all related technical areas such as information retrieval for multimedia contents, multilingual retrieval, image processing, artificial intelligence, visualization, data mining, etc. In other words, an interdisciplinary digital library research team will develop empowering technologies to enhance the way people can create, disseminate and access the content. Research into application development, presentation and usability will help to focus our R&D work with specific application in mind for the purpose of increasing universal access.

NEW MODE OF UNIVERSAL ACCESS AND GLOBAL MEMORY NET

Effective and collaborative application development in digital libraries will contribute to the eventual realization of the PITAC vision as articulated earlier. Thus, applications and use will be the centerfold of this paper.

To address this, I would like to start by approaching information seeking and use of information from a more conceptual angle. When talking about information seeking to satisfy the users’ information needs, organizations like libraries have traditionally follow a linear reference model:

Find out what one wants? [through Author, Title, Keyword, Terms, Subject, etc.]

→ Find the book(s) or publication(s) containing the information. Use the keyword(s) etc. to search library card catalog;

→ Find the book(s) or publication(s) containing the information.

Many libraries approach the use of digital resources in the same manner. In other words, taking Web search as an example, one determines the appropriate term(s) [keyword(s), subject(s), etc.], searches the Internet, goes the located Web site(s), and finds the information in that navigational path. Likewise, many digital libraries are developed with this similar retrieval route in mind using metadata as the sole sources of “descriptive” information.

This traditional approach has been very effective for finding needed information for many years and will continue to be, provided one knows what he/she is looking for! Yet, in the current digital environment with the rich and diversified resources available, we should expand our capabilities far beyond the traditional ones. We should realize that “digital libraries offer unparalleled access to information for a far broader range of users than prior physical and organizational arrangement” (Chen, 2004). In order to benefit from this digital environment, we need to find out whether we have a scalable, interoperable infrastructure that is able to bridge context, culture, and language, and enable us to gather, organize, utilize and share the rich information resources effectively. We need to change the traditional linear reference model to a new one on use and usability, in which information resources actually “talk” to each other. It is a model which stimulates the users’ thinking and learning; redefines user experiences, and gears to a much wider-range and broader-based user groups. In this direction, “universal access” is taking on a very different meaning, while data, information, knowledge, users, information services and applications all have much more expanded therefore different definitions than before (Chen, 2004).

Role of Global Memory Net in this new paradigm

Now I shall shift gear to share my own experience during this period of dynamic technological transformations since 1983. The presentation will be quite visual, but for the benefit of the readers, I shall

provide some summary descriptions.

In the last two decades, I have experienced these transformations up-close and personal. I have created the interactive videodisc called “First Emperor of China”, in the early 1980s; then with the coming of digital technology, I have converted the contents of the analog videodiscs and authored the digital multimedia CD of the same title in the early 1990s. When my institution was not big enough to be involved in any digital library activities of the first phase of the Digital Libraries Initiative (DL-I) of the US National Science Foundation (NSF), I was busy in building up contents, and developed more complete descriptive information (later known as metadata) of these resources. In 1999, when NSF first introduced its International Digital Library Program (NSF/IDLP), I proposed the Chinese Memory Net (CMNet), which became one of the first NSF/IDLP Projects. Since 2002, CMNet has expanded its scope to Global Memory Net.

Although, as the name suggests, Global Memory Net (GMNet) can and should accommodate all types of “memories” therefore all subject areas, but in order to focus our developmental work, for now, GMNet is focusing in topics related to culture, history and heritage; and hopes to be an effective gateway to the world cultural, historical, and heritage image collections from academic educational and research partners in the world. Much of these unique collections of great value to education and research as well as to general public are not currently accessible due to distance, form, and technical barriers. GMNet 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 effective retrieval of multimedia resources. More background information on the conceptual framework of GMNet as well as the system structure and development of the recent version of GMNet can be found in Zhang and Chen (2005), and shall not be repeated again.

In this paper, we shall use the Home Page of GMNet (Fig.2) as a starting point to show how the invaluable images and other multimedia resource of

the world's culture, history and heritage are organized, dynamically managed, and intuitively retrieved and delivered to the end users for their uses via both traditional and cutting-edge technologies.

This Home Page of GMNet (Fig.2) as well as the system's structural charts in Zhang and Chen (2005) shows clearly that there are several major components. These include:

(1) Collections. This is the single most important component because content is the most important part of any digital libraries. The results of the images collections will be displayed at the center of the search screen;

(2) User. In addition to user's registration and login, this component permits users to upload their resources as well as to use the project resources to create their own projects;

(3) Archives. Archival information about GMNet and related projects;

(4) About Us. Information about us and our partners/collaborators;

(5) News. Worthy news items about GMNet;

(6) Policies. Such as copyright, privacy policy, etc.

The components on About Us, Archives, News, Policies, and User are generally dynamically managed and rather obvious to all readers. Thus, this paper will concentrate on how GMNet is conceptualized and developed, in light of the new mode of universal access, to enable every potential user to benefit from the invaluable world's digital image and multimedia resources.

GLOBAL MEMORY NET OFFERS THE WORLD INSTANTLY

It is impossible to describe all the features 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. Here one is able to conduct "traditional" searches by choosing any field or all fields of the metadata—creator, title, location, time period, description, keyword, description, reference source, etc.—using the Google syntax. To do this effectively,



Fig.2 Home Page of Global Memory Net's Version 3.0

one must have some knowledge of the chosen collection, and know some of the predefined specifics of the contents of a chosen image collection. For this traditional approach, keyword search is generally the most popular one, although one can also search any other or all fields. Fig.3 shows, for example, when one enters the Emperor Collection, and knows precisely that the “tile ends” are of specific interest, he/she can simply type “+tile, +end” (+ denoting that the word must be present) and select any one of the fields like keyword, or all fields, and search. Relevant results will be shown instantly.

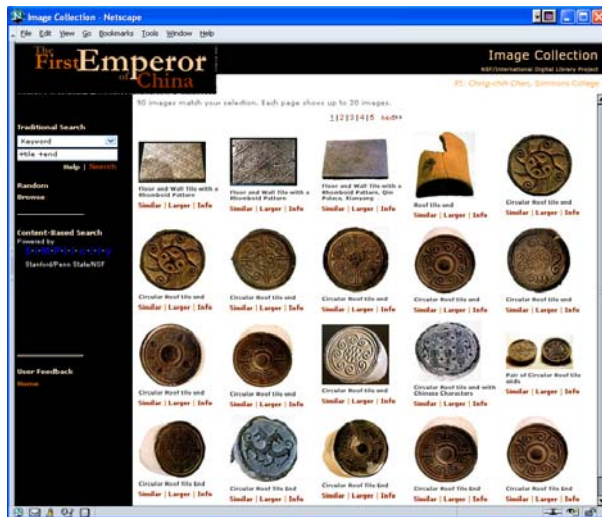


Fig.3 One conduct traditional search using “+tile +end” as search terms. Close to 100 images were found, with first 20 showing on the first screen

However, in most cases, one does not have any idea on what kind of images are available in a given image collection. This is a very important point! Thus, just like in a library, we need to provide the users an opportunity to browse the stack, and find what they need and want. In this case, in GMNet, we provide our users the ability to “browse” what are included in a collection in the order of image organization, or to request the random display of images included in the collection by clicking on the “Random” button. This “random” feature is a significant one because this will provide the user a fast exposure on the scope of coverage as well as some of the specific keywords for pursuing traditional searches as stated above. Yet, what is even better is that once an image strikes the interest of a user, then images of similar color and shape can be instantly displayed by clicking on

“Similar” button. This latter feature utilizes the cutting edge content-based image retrieval technique, SIMPLcity, developed by Prof. J. Wang of Penn State University (see Zhang and Chen for more discussions on this) (Chen and Wang, 2002). The combinations of the above mentioned capabilities permit the users to browse, retrieve, enjoy, and learn about a chosen subject in just seconds through multiple thousands of digital images accurately and effectively.

Take the Naxi Collection of the Asian Division of the Library of Congress (LC) as another example. For most cases, one does not even know what “Naxi” is (a minority ethnic group in China’s Yunan Province), then how can one be expected to know how to retrieve by keyword and browsed by title or subject as required in the LC’s site as shown in Fig.4.



Fig.4 LC’s Naxi Collection can be searched by keyword

Fig.5 shows how GMNet has enhanced both the retrievability and usability of the LC’s unique collection on the Naxi culture in the world with dramatically different approaches. When a user does not know what Naxi is all about but has entered the Naxi Manuscripts from Fig.2, “Random” search can provide enormous amount of information to a user in a great hurry. For example, Fig.5 shows that the images randomly displayed on the screen. From that display, one learns immediately that there are many keywords which can be used for searching the LC site (Fig.4), such as “ceremonial”, “sacrifices”, “pictographs”, “animals”, “divination”, “deity”, “fortune-telling”, “Tibetan language”, etc.

Now, having a much better idea on the types of images included in the Naxi collection, if one selects the pretty picture (“Illustrated card with Tibetan language”), the user has three choices for asking more

information at the click of any of the three choices (Fig.6):

(1) Similar—for all the images of similar color and shape. This opens up all possibilities for all related images which are totally unknown to the user earlier;

(2) Larger—for enlarging the images at multiple levels of zooming depending on the resolution of the image. Once the image is larger than the icon, digital watermark will be shown dynamically to protect the copyright and intellectual property of the image creator, and in this case, the image owner is the Library of Congress;

(3) Info—for textual descriptive information showing all the available metadata fields. In the case of Naxi, it will have the Library of Congress’ Naxi site indicated in the “Source”, which can be linked instantly. When full-text original source available in PDF document, sound, and video are available, they can be linked instantly as well.

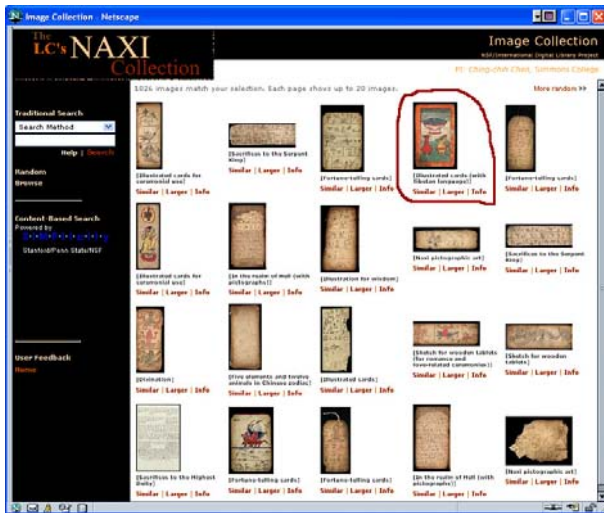


Fig.5 Random images for user’s browsing and selection

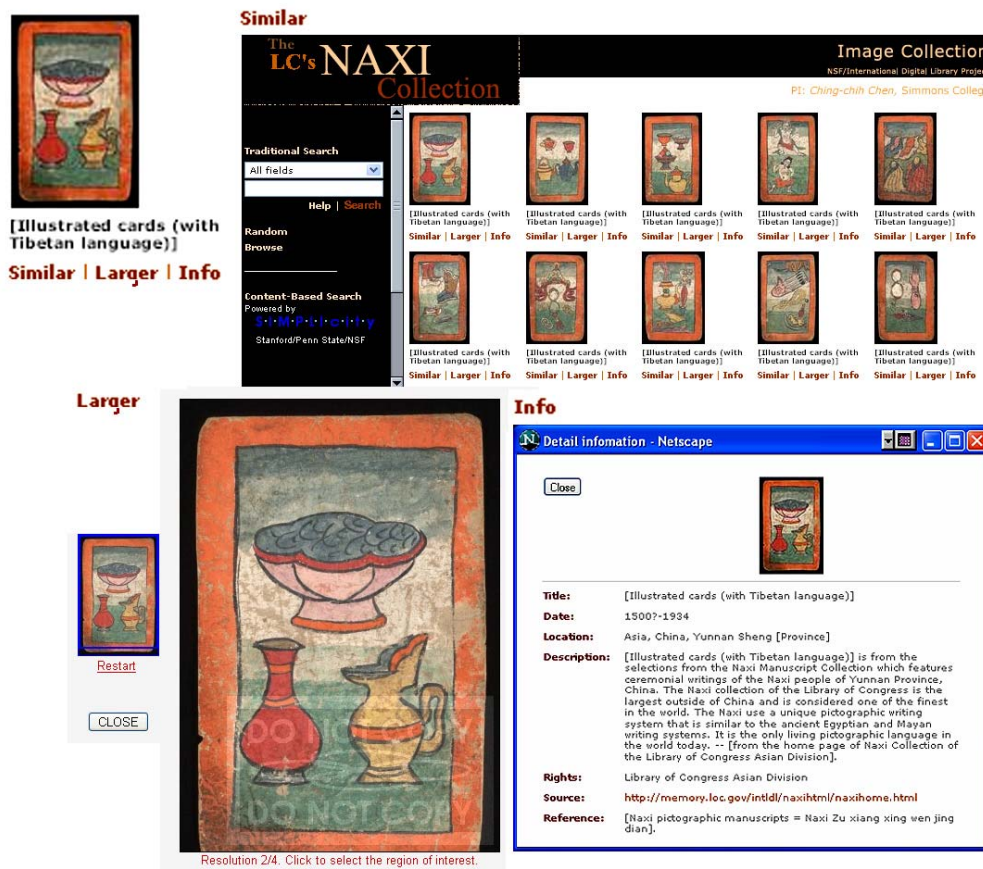


Fig.6 Three choices for a user to gain more information on the chosen image

Linking images to bibliographical resources

Thus, with just a few clicks, a user's knowledge on a subject can have a quantum jump from total ignorance to knowing rather substantially about the subject. At this point, should the user wishes to find more information, such as books, journals and additional Web resources on Naxi or any other related topics, he/she can be linked instantly to these bibliographical and Web sources, such as OCLC's World Cat¹, Universal Library, Internet Archive and Google (Fig.7). With these instant links, user can truly and seamlessly access the world invaluable resources in all formats at their fingertips. This is also where the users can benefit from projects like the Millions Books when both the quantity and quality of the digital resources grow!

Linking to world digital resources

Since GMNet is intended to be an effective gateway to the world cultural, historical, and heritage

¹ OCLC (Online Computer Library Center), a worldwide library cooperative, provides access to its 1 billion records in the full World Cat, reflecting the holdings of some 9 000 libraries in the world

image collections from academic educational and research partners in the world, it is important that it can link to the world digital collections for instant access by the users. Considering the enormous cost involved by an institution in developing its digital collection(s), it becomes clear to us that it is not only impossible but also conceptually unsound to consider any centralized digital collections. For this reason, we have judiciously identified over 2300 digital collections around the world, and all of these are instantly searchable and linked from GMNet. Online users are provided with similar search capabilities to enable them to retrieve these world digital resources—images, sound, and videos—with complete backup of textual descriptive information and referral to the world rich bibliographic resources, such as the OCLC's World Cat. In GMNet, users can also access all collections including the World Digital Collections by geographical locations. Fig.8 is the screen showing that over 230 countries and geographical areas are listed for instant retrieval of the available digital resources on and from those areas.

It is impossible to cover all the features of GMNet, but two additional ones are worthy of brief



Fig.7 Seamless integration with OCLC's World Cat, Universal Library, Google, etc.



Fig.8 User can search collections by countries and geographical areas

mention:

(1) Multi-collection, multilingual and multimedia search capabilities. Currently users can search all the collections at once or going to each individually. When there are languages other than English present in any of the collections listed, language search selection will be shown. Currently, Chinese, Vietnamese, Thai, Croatia and Italian are present. For example, we can retrieve the ancient Waka poem from the Tsurumi Collection in both English and Japanese. The PDF icon can provide the contemporary Japanese translation because the ancient Japanese are generally not readable by non-specialists, and the sound icon will play the current reading in Japanese as shown in Fig.9.

(2) User's own project creation. When a user found images of his/her choice and would like to keep them for whatever purposes—report writing, lecture presentation, showing to friends, etc., they can create up to three own projects on any subjects for a selected time period, such as 3 months. Upload functions are

also possible for them to be a contributor to the system pending on the resolution of security issues.

CONCLUSION

The above is a quick birds' eye view of Global Memory Net (GMNet) and its latest development. Our next steps will consider the development of simpler system version to enable the involvement of more organizations in the world with lesser technological capabilities, the expand of subject areas beyond the culture, history and heritage, etc.

When we view our world from a global dimension, it is really much smaller than in the past. To be able to fully leverage the world's rich information resources for the benefit of all citizens in the world, we have shown GMNet's way to seamlessly integrate all types of resources together, and make the use easy and transparent to users. To accomplish this, we need

The screenshot shows a Netscape browser window displaying a search result from the 'Multi-Collection Searches' website. The search criteria are set to 'All' and 'Japanese'. The search results show 257 images matching the selection. The main content area displays a 'Detail information' window for the document 'Ihon Heike monogatari Dankan'. The document is attributed to Sesonji Gyohshun (?-1407) and is dated to the 14th century. The location is listed as Asia, Japan. The description provides a detailed overview of the work, including its historical context and the battle between the Taira and Minamoto families. The document is available as a PDF file and an audio file. The reference section lists the source as 'Ihon Heike monogatari Dankan' by Sesonji Gyohshun, published by Tsurumi University Library in 2004.

Fig.9 Multilingual and multimedia retrieval—Japanese Waka poem in English and Japanese with PDF and sound files for contemporary Japanese translation and reading

to seek international partnerships and collaboration. This is what the ICUDL and the Universal Library (Million Book Project) are all about, and it is also what Global Memory Net is all about.

As I have traveled from country to country with special “global” lens focused on the potential sharing of cultural and heritage resources in the cyberspace, I was struck with the great potential for meaningful collaborative activities. Global Memory Net has a space holder for everyone. Instead of building localized digital libraries, let us work on building up large-scale digital contents in multimedia format together beyond national borders for sharing with the world, and let us also explore together the fertile future for distributed cross-disciplinary collaboration (Chen, 2001).

ACKNOWLEDGMENT

Global Memory Net is supported by the US National Science Foundation/International Digital Library Program with the Grant Nos. NSF/CISE/IIS-9905833 and NSF/CISE/IIS-0333036.

References

- Chen, C.C.(Ed.), 2001. Global Digital Library Development in the New Millennium: Fertile Ground for Distributed Cross-Disciplinary Collaboration. Tsinghua University Press, Beijing, China.
- Chen, C.C., 2004. The Promise of International Digital Library Collaboration for Innovative Use of Invaluable Resources (Keynote Speech). Proceedings of LIDA 2004 (Library in the Digital Age): Human Information Behaviour & Competences for Digital Libraries, Dubrovnik and Mljet

- Island, Croatia. University J. J. Strossmayer in Osijek, Osijek, Croatia, p.7-15.
- Chen, C.C., 2005. 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., Kiernan, K.(Eds.), 2002. Report of the DELOS-NSF Working Group on digital Imagery for Significant Cultural and Historical Materials. http://dli2.nsf.gov/internationalprojects/working_group_reports/.
- 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.
- Chen, C.C., Wactlar, H.D., Wang, J.Z., Kiernan, K., 2005. Digital imagery for significant cultural and historical materials: An emerging research field bridging people, culture, and technologies. *International Journal on Digital Libraries—Special Issue on EU-NSF Working Group Reports*.
- Larsen, R., Wactlar, H.(Eds.), 2004. Knowledge Lost in Information: Report of the NSF Workshop on Research Directions for Digital Libraries. Chatham, MA. School of Information Science, University of Pittsburgh, Pittsburgh, PA. <http://www.sis.pitt.edu/%7Edlwkshop/report.pdf>.
- PITAC (U.S. President's Information Technology Advisory Committee), 2001. Digital Library Panel. Digital Libraries: Universal Access to Human Knowledge. <http://www.nitrd.gov/pubs/pitac/pitac-dl-9feb01.pdf>.
- Zhang, S.Q., Chen, C.C., 2005. Global Memory Net and the development of digital image information management system: Experience and practice. *Journal of Zhejiang University SCIENCE*, **6A**(11):1216-1220.

Welcome visiting our journal website: <http://www.zju.edu.cn/jzus>
Welcome contributions & subscription from all over the world
The editor would welcome your view or comments on any item in the journal, or related matters
Please write to: Helen Zhang, Managing Editor of JZUS
E-mail: jzus@zju.edu.cn Tel/Fax: 86-571-87952276