# Challenges for Developing a World Digital Library and Gateway: The Case of Global Memory Net*

*Ching-chih Chen*
Professor
Graduate School of Library and Information Science
Simmons College, Boston, MA 02115
E-mail:  chen@simmons.edu

*Abstract:* With the exciting convergence of content, technology, and global collaboration in this digital era, there are unprecedented potentials as well as challenges for developing digital libraries of all kinds.  In the case of Global Memory Net (www.memorynet.org), its newly developed Interactive Multimedia Content Retrieval System (tentatively called i-M-C-S) has managed to incorporate many innovative concepts and functions to meet the challenges of a world digital library and gateway.  This paper will articulate these challenges, including those selected ones as listed: (1) Using the Web as a platform to integrate seamlessly all types of multimedia resources. (2) Digital images can not only be retrieved in traditional ways by metadata fields, but also by cutting-edge content-based retrieval methods (CBIR). (3) Once the desired image(s) is(are) located, associated sound, digital video, textual annotation, as well as expanded bibliographical and Web resources can be obtained by the simple clicks of the mouse. (4) Available resources can be retrieved by searching a single collection, or by cross-collection (or multi-collection) search. (5) Multilingual presentation as well as retrieval should be possible. (6) Geographical retrieval should be provided. (7) While universal access is the goal, the users should be possible to contribute their resources as well.

## INTRODUCTION

Since 1993, I have advocated the concept of a world digital library [Chen, 1993].  But, in those early days, the concept was started mostly from the "technological" point of view, contemplating the possibilities of linking the world digital collections together with the availability of high-speed network like the Internet as we commonly knew that a couple of years later.  In the early beginning, there was little idea regarding how to deal with the digital collections.  As time goes on, such a concept has moved rather rapidly to include many other components.  Content has been considered to be just as important, if not more, as technology. In other words, with all the rich information resources in the world, if they are not digitally available, they will not be able to be accessed, retrieved, shared and enriched on the Web. Research activities related to content have proliferated in the areas of digital collection creation and development, organization and standards, interoperability and scalability, and many others.  One important fact in relation to content is also becoming clear – no institution can provide everything, thus global collaboration in digital collection development becomes essential. In addition, it is natural to ask what are these digital resources are developed for.

---

The component of users and uses has surfaced as another important part of the digital library research.
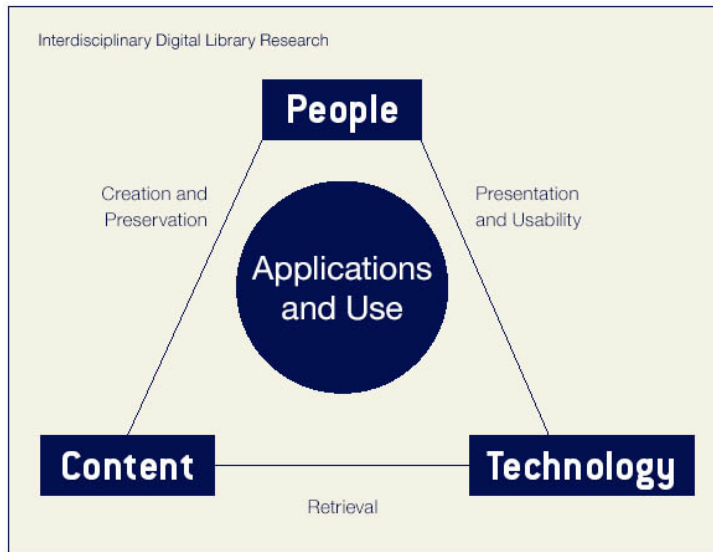


Figure 1. Conceptual Model of an Interdisciplinary Digital Library.

The above macro introductory description, though over-simplified, is sufficient to provide a sketchy conceptual model, which coincides well with an interdisciplinary digital library research model, collectively developed by the members of the *DELOS-NSF Working Group in Digital Imagery for Significant Cultural and Historical Materials* [Chen, 2005]. This model was presented at LIDA 2004 at the keynote speech [Chen, 2004] – Figure 1. This illustration shows clearly the triangular relationships among *people*, *content*, and *technology*. In other words, an interdisciplinary digital library needs to develop technologies to enhance the way all kind of users can use, share and contribute to the world digital content.

## CHALLENGES FOR DEVELOPING A WORLD DIGITAL LIBRARY

Conceptual model as described above is essential for developing a digital library because it provides a map to guide us in our planning and developing processes. Yet, when actually creating a digital library application, one will face with many difficult tasks due to the need to get down to the granularity of creating and developing the building blocks. Like completing a puzzle, all small and large pieces have to fit well together to create the complete whole.

In the case of Global Memory Net, since it is intended to be a world digital image library and gateway, we have conceptually set various requirements as our challenges for developing such a world digital library. The following are some of the selected ones:

- Instead of using web for publishing, we need to use web as a platform to enable more user participation,
- We need to provide seamlessly integrated multimedia information services in order to enrich user experience,
- Our system needs to provide innovative information services, and it is not to offer packaged software for building databases,
- User's need and behavior in information seeking should not be pre-determined, thus the search and retrieval capabilities have to be able to accommodate all kinds of needs,
- User needs to control the use of data, thus he/she can choose to search one single collection or multi-collections,

- What we offer should be a knowledge base and not database(s), therefore once simple information is found, user can and should be enriched with more related multimedia information, and mixable or re-mixable resources if desired,
- We need to provide freely accessible and searchable web resources, and not just links with them,
- To enhance user experience and ability to consume the obtained information, multilingual capabilities are essential for both presentation as well as retrieval,
- User should be able to lead from one tiny useful information to play big and link that to the world collections as well as world bibliographic and web resources,
- User should be able to address the content in any granularity,
- Geographical access to content should be provided,
- Intellectual property of the content provider should be protected,
- User should be able to use the discovered and retrieved data to develop his/her own project(s),
- User should be able to actively contribute to the library,
- Etc.

## THE CASE OF GLOBAL MEMORY NET

At LIDA 2004 [Chen, 2004], Global Memory Net (GMNet) was presented with a macro conceptual model with a small number of novice functionalities. Since then, fast system and content development has taken place in the short two-year period. Particularly, in the Fall of 2005, with the National Science Foundation supported community building effort, 7 visiting researchers from China, Croatia, and Vietnam joined the GMNet research team full-time for 3 to 5 months each. The project really took off. Boris Badurina of the University J. J. Strossmayer, Osijek will present his experience in developing the Museum of Slavonia Collection in GMNet, as well as his continuing technology development effort for the project [Badurina and Chen, 2006]. Other visiting researchers have also contributed to the content development from China and Vietnam, as well as technology development [Zhang and Chen, 2005].

From the technology side, we have developed our own in-house Linus/MySQL/PHP-based *i*nteractive **M**ultimedia **C**ontent retrieval **S**ystem (*i*-M-C-S) with much added functionalities from Version 1.0 to the current Version 4.5 since 2004. The later versions have been developed based on the basic system structure of Version 2.0 with functions as shown in Figure 2 [Zhang and Chen, 2005]. The only exception of our in-house development relates to the content-based image retrieval (CBIR) capabilities, for which we collaborate with James Z. Wang of Penn State University by using his SIMPLIcity, developed when he was at Stanford University.

### Features of Global Memory Net

In the last years, we have addressed the challenges outlined above, and have developed our system, i-M-C-S, with many features and functionalities in line with our initial requirements. Although these would be best to be experienced with the real-time presentation and discussion, the following summary with some of the selected feature is provided for the benefit of the readers:
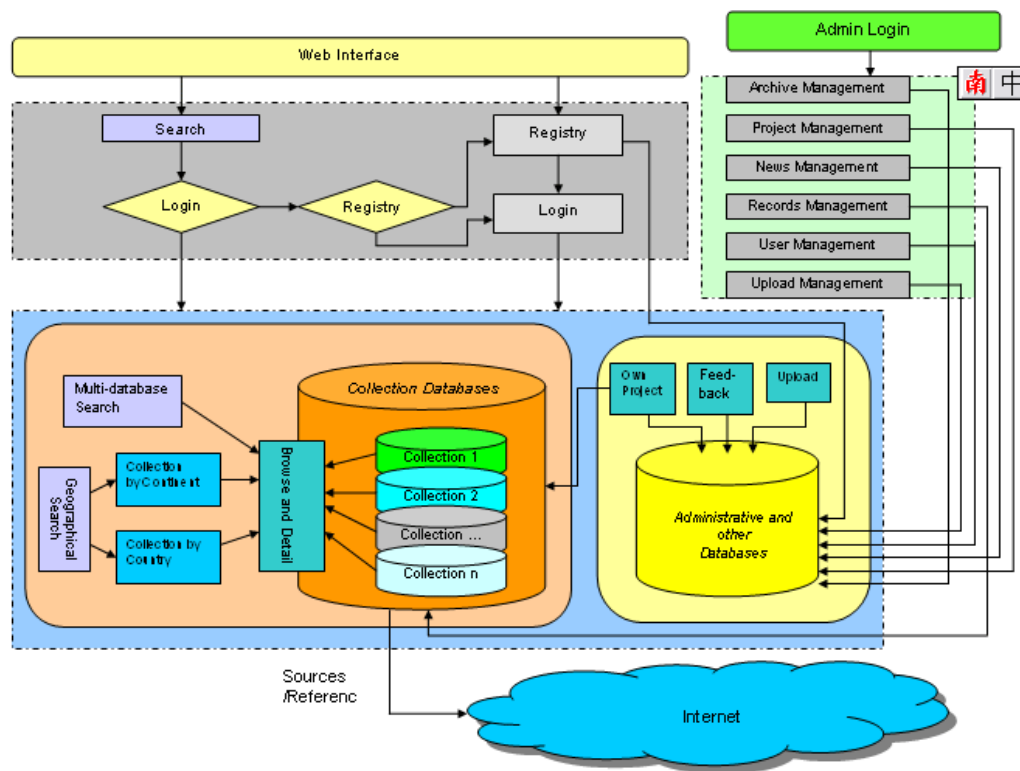
Figure 2.  Functionality Scheme of Global Memory Net [Zhang and Chen, 2005].

- *Web is used as a platform*, and not as a publishing medium.

- *Instant access to rich image collections* – One can access to over 30 image collections in GMNet with over 20,000 images as well as over 2400 digital collections from over 80 countries.  Although GMNet's own current collections have focused in culture, history, and heritage, the World Digital Collection in GMNet has included over 2400 digital collections cover all subject areas.  This includes over 100 collections from UNESCO's Memory of the World, over 290 collections from the US Library of Congress, as well as those from other major national libraries, archives, museums, academic institutions, etc. In other words, the world's rich resources are instantly accessible at a simple click of the mouse.  Figure 3 shows how these collections can be accessed by selecting the collection from the left blue panel of "Collections".

- *Easy and flexible traditional search* – Select a collection of interest, and use traditional search by metadata fields when one knows the precise information to search in this collection.

- *User can explore the unknown collection and learn its coverage* – For an unknown collection of curiosity to the user, one can browse the collection, but more significantly, the "random" feature permits one to explore and learn the coverage of this collection.  In seconds, one can learn how to retrieve effectively by the variety of images displayed as well as the words showing up for the titles as suggestive keywords.  Figure 4 shows the
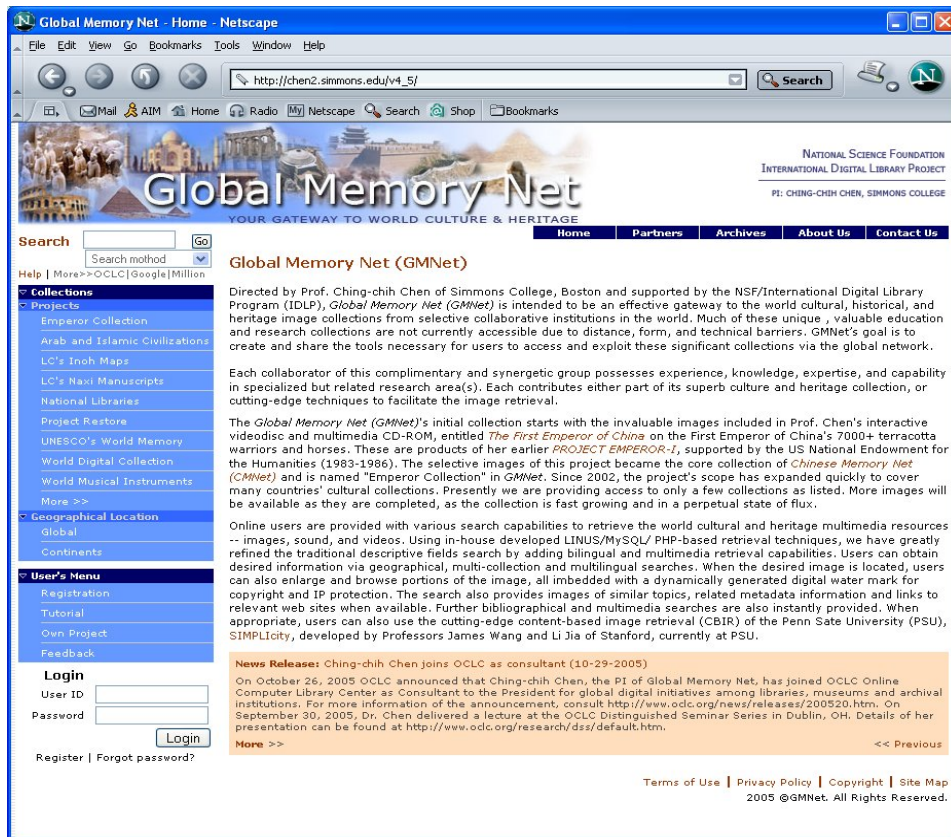
Figure 3.   Home Page of Global Memory Net with Collections Panel on the Left.
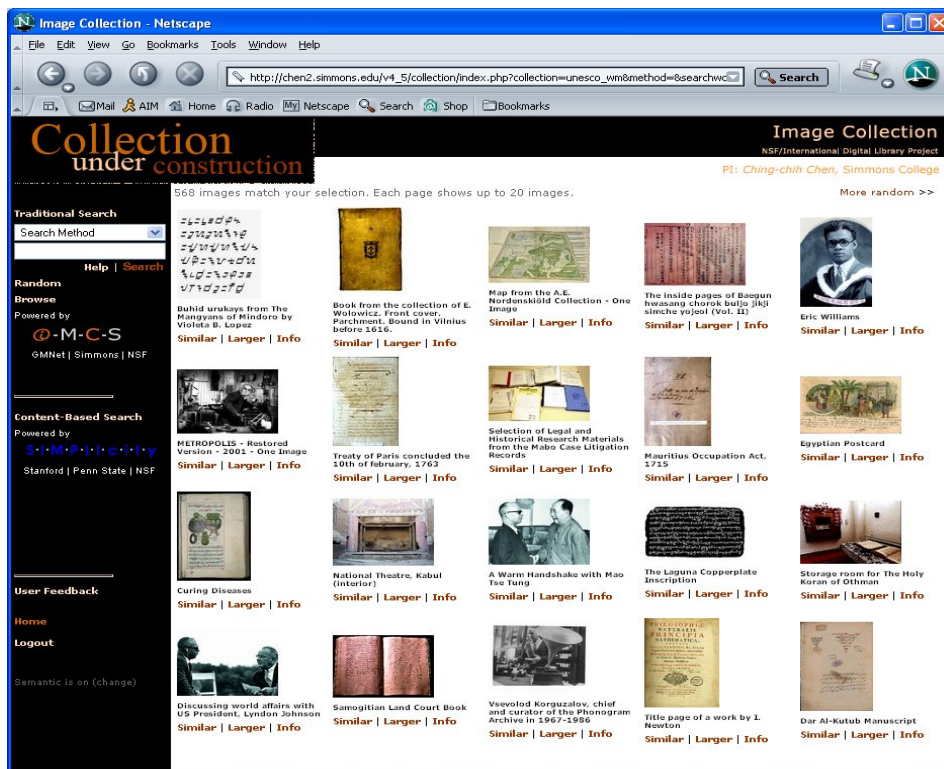
Figure 4. For UNESCO's World Memory, "Random" feature provides an instant display of
a great diversity of topics covered

great diversity of images covered in the UNESCO's Memory of the World collections. A quick glance of this screen, one can have a good idea on the subjects involved.

- *Instant retrieval of similar images of interest by using cutting-edge retrieval techniques* – Once a picture of interest is spotted, one is given three choices - to find pictures of the same color or shape using the cutting-edge content-based image retrieval (CBIR), or to zoom the image for larger sizes and more details, or to have some descriptive information, as shown in Figure 5 on the right. Our CBIR uses SIMPLIcity of Prof. James Z. Wang of Penn State University.



Fig. 5. Chosen image

  - "Info" – Want to know more about this image? "Info" will yield instant descriptive information about it, as shown in Figure 6. This information can be in multilingual forms. In addition, a simple click on the URL of the source, the user will go imme-diately to the Web page so he/she can browse and search information there.



Figure 6. Descriptive information (metadata) of the image are available in both English and French of this Islamic image

  - "Similar" – Curious about what other similar images of same color and shape? Figure 7 shows that a simple click on "Similar", all similar images of the same colar and shape, or the same search term will show up on the screen for immediately for the user to browse further.

    Undoubtedly this opens up a new horizon for learning.



Figure 7. Instant display of all similar images of the same color and shape
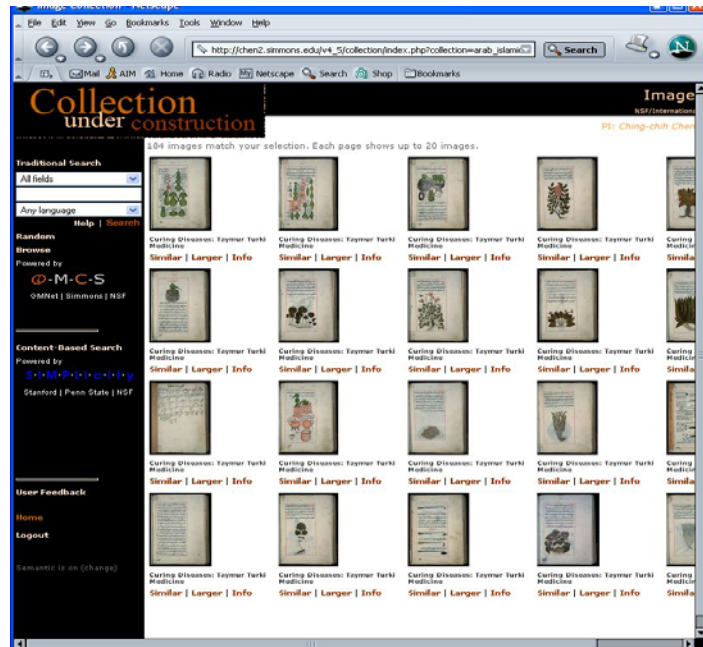
o "Larger" – By clicking this, instant magnification of a chosen image in defined areas is shown. Depending on the resolution of the image, some images can be zoomed over a dozen of time. Note the small icon of the rare Japanese Inoh Map iof the Library of Congress' Asian Division (Figure 8). It is so small that nothing can be seen. Yet, this tiny icon-image can be magnified to 12 times with a great deal of significant information to the scholars.
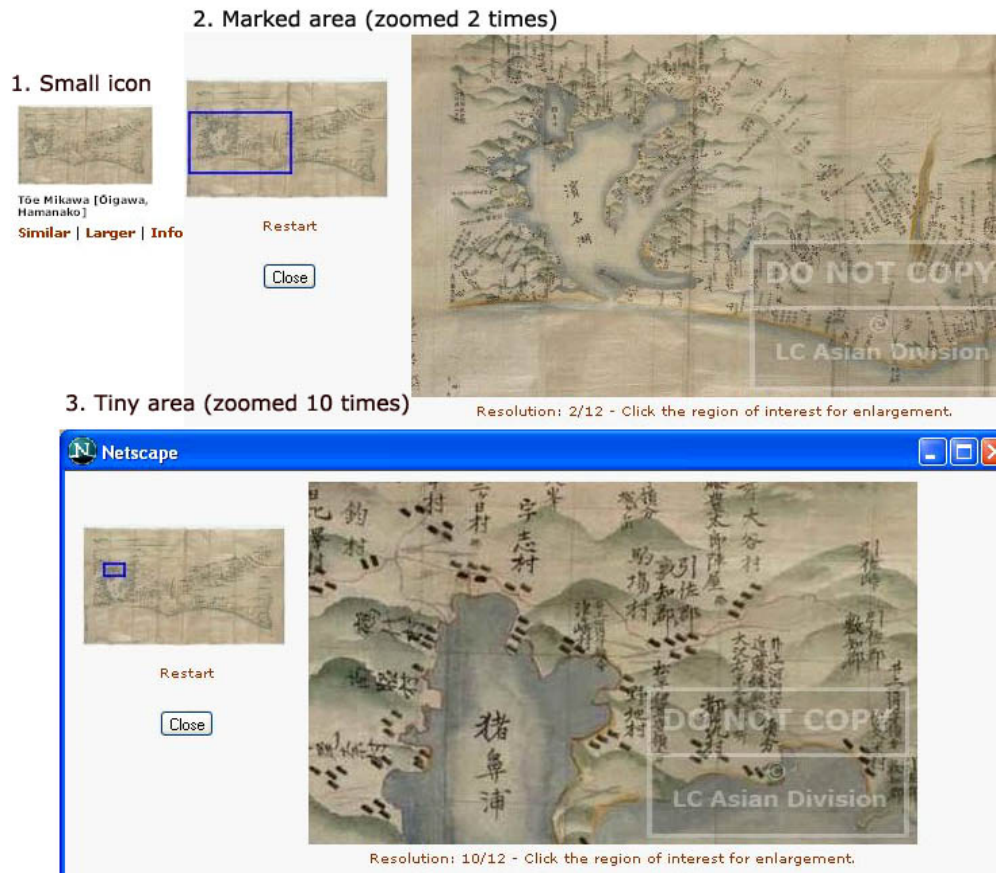


Figure 8. This figure shows the multiple levels of magnification.
Note also the dynamically generated digital watermark.

- *Dynamically generated digital watermark for copyright protection of the content provider* – See Figure 8 above.

- *Seamless integration of multimedia information* – If relevant resource information on a chosen image is available in formats other than textual annotation, the user can then retrieve the relevant audio, video, etc. again by a single click of the mouse, as shown in Figure 9.



Figure 9. The "Info" screen shows the availability of 3 digital videos and 2 PDF document files for instant retrieval. Some also include sound and other files

- *User can choose to conduct either a single collection or multi-collection search* – For example, if one wants to find information on Naxi, one of the 56 minoriy ethnic groups in China, one can choose to search Library of Congress' Naxi Manuscript Collection, or one can do multi-collection search. The later will search all the 30 some collections in GMNet, and the search results will instantly be shown (see Figure 10). Thus, in addition to 1027 images found in LC'Naxi Manuscripts, one can also see additional resources and images in other collections.

- *User can have much more in-depth learning of the subject by using expanded bibliographical and web resources* – For example, one can instantly find relevant books from the half-billion bibliographical records of OCLC/World Cat, or web resources like Google, Wikipedia, Internet Archive, Million Books etc…
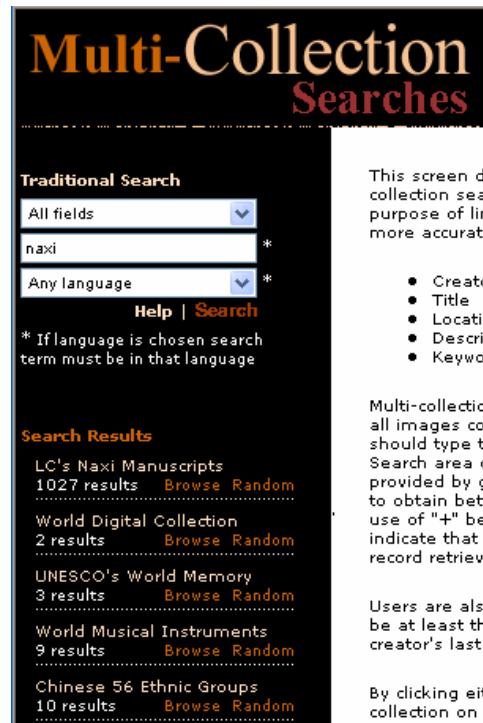


Figure 10. Multi-collection retrieval results

- *Multilingual display of descriptive information as well as multilingual retrieval* – Currently GMNet already has contents in English, Chinese, Croatian, French, Italian, Japanese, Spanish, Thai, Vietnamese, etc.

- *User can conduct geographical searches by continent and/or country* – Information on over 230 countries can be searched.

In addition to the features mentioned above, it is worthy to stress that GMNet positions its user to:

- Participate actively during his/her own information seeking process,

- Decide how he/she would like to use the discovered or retrieved information,

- Develop his/her own project(s) by choosing, mixing and remixing the retrieved images, and

- Be an active content contributor to GMNet.


**CONCLUSION**

Chinese Memory Net, started in 2000, was supported by the National Science Foundation's International Digital Library Program. Since 2002, Chinese Memory Net has been expanded to be Global Memory Net. The conceptual model of this world digital image library was

conceived as early as 1993, and further refined in the late 1990s. The development of GMNet began at a time when there was no mention of Web 2.0.

The concept of "Web 2.0" began with a conference brainstorming session between O'Reilly and MediaLive International in 2004, and "in the year and a half since, the term "Web 2.0" has clearly taken hold, with more than 9.5 million citations in Google." [O'Reilly, 2005].
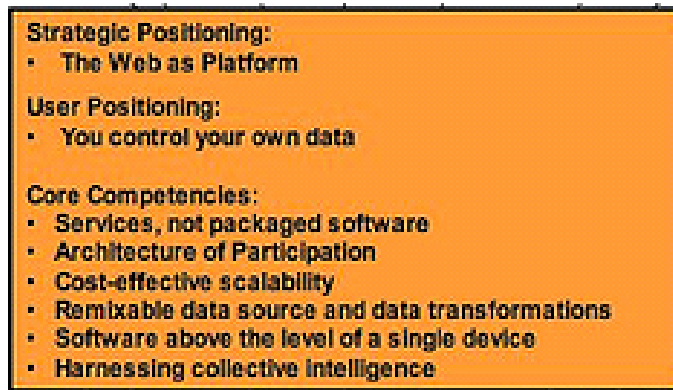


Figure 11. Part of the O'Reilly "meme" map showing the Web 2.0 Core

Although there are still a lot of disagreement on what Web 2.0 is all about and for, for GMNet, we have actually found considerable agreement between the core concepts of GMNet and the Web 2.0 core as shown in Figure 11, as well as some of the ideas that radiate out from this core.

It is clear to us that the keywords for next-generation web-based application, including digital Libraries, should be "user", "services", "participation," "knowledge base," etc., and NOT "publishing", "packaged software", and "database(s)".

## ACKNOWLEDGMENT

## REFERENCES

[Badurina and Chen, 2006]. Badurina, Boris and Ching-chih Chen, "Osijek Treasures Are One Click Away at the Global Memory Net," Paper presented at *LIDA (Library in the Digital Age) 2006*, Dubrovnik, Croatia, May 30-June 2, 2006.

[Chen, 1993]. Chen, Ching-chih, "Technological Potentials for the Global Digital Library," In: *Proceedings of the International Conference on National Libraries - Towards the 21st Century, Taipei, Taiwan, April 20-24, 1993*. Taipei, Taiwan: National Central Library. pp. 859-868.

[Chen, 2004]. Chen, Ching-chih, "The Promise of International Digital Library Collaboration for Innovative Use of Invaluable Resources," In: *Human Information Behaviour & Competences for Digital Libraries*. [*Proceedings of the Libraries in the Digital Age, Dubrovnik and Mljet Island, Croatia, May 25-29, 2004.*] Osijek, Croatia: University J. J. Strossmayer in Osijek. pp. 7-15.

[Chen et al, 2005]. Chen, Ching-chih, 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*, 5: 275-285 (2005).

[O'Reilly, 2005]. O'Reilly, Tom, "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, " September 30, 2005. http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

[Zhang and Chen, 2004]. Zhang, Shengqiang and Ching-chih Chen, "Global Memory Net and development of digital image information management system: Experience and practice," *Journal of Zhejiang University SCIENCE*, 6A (11): 1216-1220. November 2005.

**Ching-Chih Chen** is Professor of the Graduate School of Library and Information Science, Simmons College, Boston, and is a consultant and speaker to over 40 countries. She is the author/editor of more than 35 books and over 200 journal articles in areas of new information technologies, such as global digital libraries, multimedia technology, digital imaging, interactive videodisc technology, global information infrastructure, information management, and information resources, etc. She produced the award winning interactive videodisc and multimedia CD entitled *The First Emperor of China*, supported by the US National Endowment for Humanities (NEH). Currently she is leading two major NSF/International Digital Library Projects (IDLP): (1) Global Memory Net, a gateway to the world cultural, historical, and heritage multimedia resources, with collaborators from different part of the world, and (2) International Collaboration to Advance User-oriented Technologies for Managing and Distributing Images in Digital Libraries. She is also co-PI, with Prof. Raj Reddy of Carnegie Mellon University, of the China-US Million Book Digital Library Project.

Photo credit: Justin Knight

A Fellow of the American Association for the Advancement of Science, she was appointed by President Clinton in February 1997 to serve as a member of the U.S. President's Information Technology Advisory Committee (PITAC). PITAC was established by a new Presidential Executive Order. Under both Presidents Clinton and Bush during 1997 to December 2002, she co-chaired the PITAC Subcommittee on International Issues, and was a member of the PITAC Subcommittees on Next Generation Internet (NGI) and IT*2 Initiative Review; and Panels on Digital Divide, Digital Library, Learning of the Future, and Individual Security. She also chaired the PITAC's activity on Digital Divide for Smaller Institutions. During 1987 to 2001, Dr. Chen was Chief Organizer of a series of 12 International Conferences on New Information Technology (NIT) in many continents of the world. The outcome of NIT '99 (Taipei) and NIT'2001 (Beijing) are the two-volume books related to the development of Global Digital Libraries – *IT and Global Digital Library Development* (1999) and *Global Digital Library Development in the New Millennium: Fertile Ground for Distributed Cross-Disciplinary Collaboration* (2001). She is a recipient of many major awards, and was also elected in 1985 as Fellow of the American Association for the Advancement of Science. She served as an Honorary Professor of Tsinghua University in Beijing from August 1999 to 2002 and University of Hainan, China since 2004. Active in the digital library area she was the co-Chair of the 4th ACM/IEEE Joint Conference on Digital Libraries (JCDL) of 2004 held in Tucson, Arizona in June 2004. She was on the Advisory Board of DELOS (the European Digital Network for Excellence), serving as the US Co-Chair of the NSF/DELOS Working Group in Digital Imagery for Significant Cultural, Historical and Heritage Materials, and served as the co-editor for the *Journal of Digital Library*'s Special Issue on Multimedia Contents in Digital Libraries (February 2006). A sought-after international speaker, in the last two years alone, she delivered keynote speeches and made presentations at many international conferences including those in Delhi and Bangalore, India; Dubrovnik, Croatia; Buenos Aires, Argentina; Beijing, Shanghai and Hangzhou, China; Kawasaki, Japan; Tainan, Taiwan, etc. She is on the advisory board of the major China Digital Library Project of the National Library of China; in October 2005, she was appointed as a consultant to OCLC for its Global Digital Initiative (http://www.oclc.org/news/releases/200520.htm); and receivedthe coveted LITA/OCLC Kilgour Award from the Library Information Technology Association in June 2006 (http://www.ala.org/ala/pressreleases2006/april2006/2006KilgourAward.htm).