

GLOBAL DIGITAL LIBRARY DEVELOPMENT

in the New Millennium

Fertile Ground for
Distributed Cross-Disciplinary Collaboration

Edited by
Ching-chih Chen



TSINGHUA UNIVERSITY PRESS
<http://www.tup.tsinghua.edu.cn>

G250

C518

GLOBAL DIGITAL LIBRARY DEVELOPMENT IN THE NEW MILLENNIUM

***Fertile Ground for Distributed
Cross-Disciplinary Collaboration***

Edited by
Ching-chih Chen

Tsinghua University Press

(京)新登字 158 号

Global Digital Library Development in the New Millennium——Fertile Ground
for Distributed Cross-Disciplinary Collaboration

©2001 by Tsinghua University Press

All rights reserve. No part of this publication
may be reproduced without prior written
permission of the publisher and the editor

Tsinghua University Press

Beijing, 100084, China

[Http://www.tup.tsinghua.edu.cn](http://www.tup.tsinghua.edu.cn)

Fax: 010-62771137

出版者：清华大学出版社（北京清华大学学研大厦，邮编 100084）

印刷者：北京人民文学印刷厂

发行者：新华书店总店北京发行所

开 本：787×1092 1/16 印张：39.5 彩插：8

版 次：2001 年 5 月第 1 版 2001 年 5 月第 1 次印刷

书 号：ISBN 7-302-04473-2/TP·2638

印 数：001~800

定 价：198.00 元 (US\$ 59.95)

PREFACE

In recent years, we have experienced how new information technology has transformed the way we live, learn, work, do research, play, and communicate with each other across countries' borders; and will be much more so in the future. The fast and dynamic advances in computing and communications technologies have created a new infrastructure for education, research, commerce, and social interaction.¹ As a result libraries and information centers can make available to their users a previously unknown level of service, featuring powerful global information access, retrieval and delivery capabilities.

In addition to the fast developments in telecommunication networks, electronic publishing, interactive multimedia technologies, integrated information systems, and digital libraries, the explosive and widespread use of the Internet and World Wide Web have enabled librarians and information specialists to offer unprecedented information services. In the areas of digital libraries, there are incredible opportunities for global information access and sharing with distributed information systems. On the other hands, cyberspace is becoming a dominant medium through which people can create and share information and ideas without the formal information delivery structure. This has posed great challenges to libraries and information centers.

The coming of the Next Generation Internet in this new millennium will offer hundreds and thousands times more capabilities for faster and better digital communications. Thus, we will see more and bigger distributed digital libraries or information systems with multimedia information developed everywhere in the world. Yet, in order to create these widely distributed systems effectively, we shall need to know how to create them in such a way that they can operate in multiple languages, formats, media, and social and organizational contexts so that they can be easily connected, interoperable, and retrievable. In addition, we shall also need to know how to protect intellectual properties in this fluid and uncontrollable cyber marketplace. These high-order research problems can only be tackled by international experts from many related interdisciplinary fields – computer science, multimedia, communications, networking, library and information science, linguistics, law, social sciences, and the like. This is what this invaluable book hopes to address and include. This resourceful reference book includes contributions from a long list of international authoritative experts representing most of the interdisciplinary fields just mentioned. This collection of the timeliest papers are only possible because of the *NIT 2001: The 12th International Conference on New Information Technology* held in Beijing, May 29-31, 2001, because it contains the fully edited versions of those papers presented at that conference.

Since 1987, I have organized successfully eleven *International Conferences on New Information Technology (NIT)* in Bangkok, Singapore, Guadalajara (Mexico), Budapest,

¹ U.S. President's Information Technology Advisory Committee. *Information Technology Research: Investing in Our Future*. Washington, DC: National Coordination Office for Computing, Information, and Communications. February 24, 1999. Executive Summary.

Hong Kong, Puerto Rico, Alexandria (Virginia, USA), Riga (Latvia), Pretoria (South Africa), Hanoi (Vietnam), and Taipei. This series of conferences has helped to force global consensus on key problems and issues facing library and information professionals during this exciting but confusing time. As we have entered the new Millennium, the latest *NIT 2001* in Beijing with its exciting technical environments at Tsinghua University is to do the same. Yet, in addition to the *NIT*'s original goals and objectives, *NIT 2001* is held on the impressive campus of the world-renowned Tsinghua University as one of the major celebration events for its 90th Anniversary. In line with the pioneering spirit of Tsinghua, it is most fitting to choose such a groundbreaking theme for many international professionals from different fields interested in global digital libraries to contemplate the great challenges facing them in this new Millennium. It is also most fitting to pay tribute to this great university in this special way because this gathering, with its groundbreaking events, will mark the beginning of some exciting global digital library development in China and elsewhere. It will, for sure, contribute to encourage productive interchange of data, information, and knowledge of distributed digital systems around the world, and to prevent the development of fragmented digital systems with unproductive duplication of efforts. Similarly, in making these edited papers available in printed form, I hope that this book will be as an inspiring volume to those readers who are interested in collaborative global digital library development, and willing to accept the challenges lying ahead of them.

While I have advocated the "global digital library" concept for over a decade, we have only begun to enter the threshold of this development. We realize that with world's vast amount of information resources, only a very minute percentage of them are digitally available for global sharing. It is impossible for any country, even the rich ones, to consider the development of a Universal Library alone. We must learn to collaborate with each other -- to pull our limited financial resources and technical expertise together, to benefit from others' experience and knowledge, and to build bigger and richer contents that are supplemented by each other's collections. It is my hope that this book can inspire many to keep this type of global networks growing and to work collaboratively toward the development of a truly functional Global Digital Library!

As we are at the dawn of a new millennium, while excited by the fascinating future the information technology will bring to us, many of us are very much perplexed by the ever changing world brought upon by the dynamic technological changes, and wondering what the direction the libraries and information center should be. I hope that the thoughts and ideas expressed in this rich volume will confirm to many of us that we are not alone. Together we will find a workable collaborative path. Only this, we will have the potential to fulfill the vision for universally accessible collections of human knowledge as articulated by the PITAC's Panel on Digital Libraries, "*All citizens anywhere anytime can use any Internet-connected digital device to search all of human knowledge.*"²

We know that we have a long way to go, but we must begin to work toward that.

Ching-chih Chen

Newton, Massachusetts

May 2001

² U.S. President's Information Technology Advisory Committee. *Panel on Digital Libraries. Report to the President.* February 2001. (see Appendix 1 of this book).

ACKNOWLEDGMENT

During my professional career, particularly in the last two decades, I have been very fortunate to have opportunities to create my own powerful but informal network consisting of hundreds of professionals in over 30 countries. This informal network is also growing as time goes on both in size and scope, as well as the subject areas covered. The way we communicate and do our business among members of this informal network have certainly changed and transformed substantially by the ever powerful Internet and World Wide Web. We have truly benefited from the exciting time we are in with all the technological advances!

My earlier interest in applying technologies to "information" related areas has offered me endless opportunities to work with people of all walks of lives in the society, as well as accomplished professionals in almost every subject field and capacity. These societal crosswalks and promise for interdisciplinary collaboration are certainly reflected well in this book with contributors coming from many different countries and subject fields. They work in different capacities and at different levels -- from top leaders of countries directing and implementing their countries' network development and/or content resources to the top administrators of several world-class universities, from first-rate computer science, communications, and library and information science researchers and educators to the managers of national distance education consortium, from information service providers servicing all types of libraries to the customers of these profit-making organizations, from large information company chiefs to the consumers of these information products, etc. This book is a melting pot of ideas and human knowledge, research results, and expirations on the global digital library development. Thus, clearly it would not be possible without the papers from these authors, who deserve my most sincere gratitude for their thoughtful and timely contribution.

This book is based heavily on the papers and posters presented at the *NIT '2001* at Tsinghua University, Beijing. But, it is much more than a *Proceeding* of a conference. It is not a publication with a format template for the authors to follow, and simply a collection of what will be presented. As an editor, in order to ensure the quality of the papers, numerous communications with authors took place prior to and during the production process. In order to enhance the readability and value of some papers particularly those from non-English speaking countries, I have taken the liberty to edit all papers, and edit some heavily or even rewrite several of them. Thus, if I have mistakenly interpreted some of the authors' intents, the faults are mine, and my advance apology for the possible mistakes.

The time allowed to complete this incredible task is extremely short. In fact, some of the revised full papers did not arrive via the network until 10 days to the printing deadline, less than one month from the conference, and one final abstract arrived on the date of mailing. While this attests to the power of technologies, one can well imagine the tremendous stress placed on me in producing this book during every step of the process -- final reformatting, editing, rewriting some, producing graphics and hardcopies, indexing, and designing the book jacket, etc. I wish therefore to thank those who have helped me

during the process. Dr. Diane Tebbetts, Associate Director of University of New Hampshire Libraries, sacrificed several weekends to come down from New Hampshire to help me. Xiaohui Zheng of Tsinghua University Libraries and her assistant reformatted the initial pre-edited versions of the papers from China. My assistants, Ting Ying and Alissa Miller followed my instructions and markings in preparing the early drafted indexes. Their helps made my task less painful. I wish to thank Mr. Wang Chuanchen of the Tsinghua University Press for his communications on numerous versions of the book jacket design as suggested and revised by me. These have resulted in a much more attractive cover prepared by the cover designer. With the large number of two-way text and graphic file transmissions between Beijing and Boston, as expected, system incompatibilities crashed my computer system several times always during the most crucial moments. The help from Paul Aloisio, the Manager of the Technology Lab, Graduate School of Library and Information Science of Simmons College, and his staff, as well as the staff from the College's Computer Center has been much appreciated. The Tsinghua University Press's contribution to the printing of this book is gratefully acknowledged.

As mentioned earlier, this book is closely related to the *NIT 2001*. For this reason, it is impossible not to thank those who have made the conference possible. I am grateful to Steve Griffin, the Program Director of US National Science Foundation's (NSF) Digital Library Initiative of the IIS Division, and William Chang, Senior Program Manager of the NSF's Division of International Programs for their effort in facilitating the much-needed support from NSF to enable several leading digital library researchers in the US to contribute their experience and knowledge at *NIT 2001*. These speakers' willingness to prepare their papers in an impossible timeline is very much appreciated. The small carry over fund from the World Bank has made some participation from developing countries possible.

I am also thankful to the Ministry of Education, Ministry of Culture, Ministry of Science and Technology, and Natural Science Foundation of China for their generous financial support to the *NIT 2001*'s Local Organizer, Tsinghua University, for running this important conference. Their enthusiasm, endorsement, and support together with the effort from each of the Local Steering Committee members have made the conference a great success. The generous contributions from numerous publishers and vendors, particularly those from Elsevier Science, Engineering Information, Inc. and ISI-Thomson Scientific, are very much appreciated.

From conference organization point of view, I cannot find a better Local Committee to work with. Vice President Hu Dong-cheng of Tsinghua University has ably chaired this Committee and offered strong support in every step of the planning and implementation process. I wish to acknowledge the needy support and help offered by Prof. Li Yan-da, Dean of the School of Information Science and Technology of Tsinghua University, and Prof. Gao Wen, Deputy President and Professor of Computer Science, University of Science and Technology, Academy of Sciences of China, with whom I shall have the pleasure to work with on the China-US Million Book Digital Library Project.

While it is impossible to thank individually each of those Chinese leaders in the fields of computer science, and library and information and science for his/her contribution to *NIT 2001*, I wish to thank Prof. Zhou Li-zhu, Chair of the Department of Computer Science and Technology of Tsinghua University for his role as the Co-Chair of the Chinese Paper Review Subcommittee together with Prof. Gao Wen. His contribution

to the final program planning has also been helpful. Dr. Xing Chunxiao, who has worked diligently throughout the conference planning process, has ably supported Prof. Zhou's effort particularly for those work related to the Chinese participation of the conference.

I wish to credit the smooth running of the conference to my friend, Prof. Liu Gui-lin, Director of the Tsinghua University Libraries, who, in his role as the Vice Chair of the Local Organizing Committee, together with his Vice Director, Ms. Yang Yi, have shouldered all conference logistical responsibilities with care and devotion. In this regard, I must thank the invisible Internet for making our thousands of emails bouncing back and forth between Boston and Beijing possible. I can't help wonder how NIT was ever run without the Internet during those early days of 1980s and 1990s? One other aspect of the conference logistics is the arrangements for the participants. For this, I must acknowledge the assistance from Mr. Yu Hesheng, Deputy Director of Tsinghua's Center for Overseas Academic & Cultural Exchanges, for his hard work in making both the participants' stay in Beijing and after Beijing more smoothly and enjoyable.

Finally but not the least, no activities of mine can be accomplished without acknowledging the moral and assured support from all my family members. All of them understood the importance of international work to me and have full-heartedly supported me all these years without little complains. In particular, my husband, Sow-Hsin, has always been there for me and we have mutually supported each other's academic and professional activities throughout the years with much pride and enthusiasm. His love and understanding have made it possible for me to carry out this kind of strenuous and stressful activity, like all those involved in NITs, without additional worry and justification. I have taken granted much for so many things that are luxuries to many.

Ching-chih Chen



Figure 5. The Group Window showing the selected images of The First Emperor of China Exhibition Collection (Chen, Ching-chih, p. 21)



Figure 6. The Image Workspace showing how the selected image of *The First Emperor of China Exhibition Collection* can be zoomed and studies (Chen, Ching-chih, p. 21)

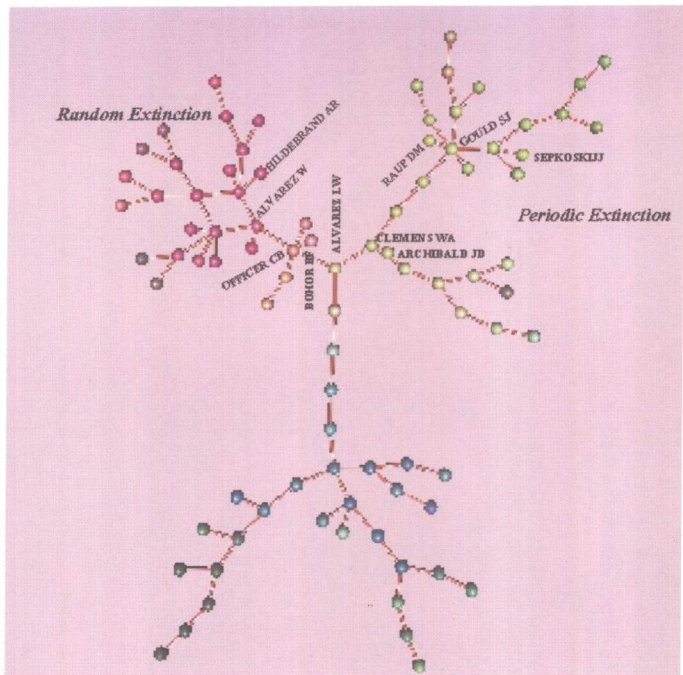


Figure 3. An overview map of author co-citation network based on the dinosaurs extinction(Chennawasin and Chen, p. 55)

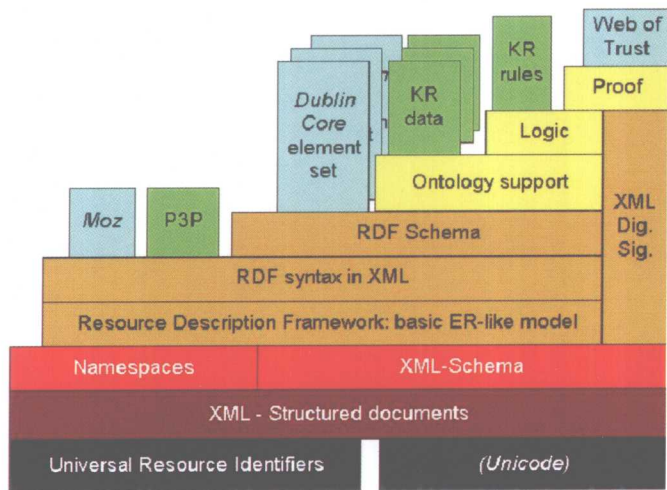


Figure 1. XML-based digital resource description framework [Lee, 2000]
(Gao, Huang, Duan, and Wang, p. 71)

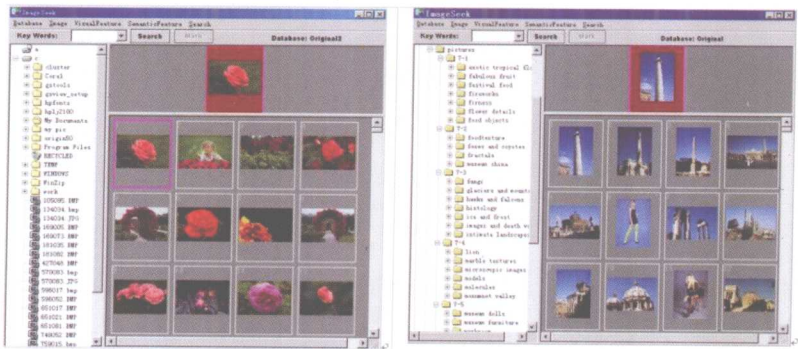


Figure 4. Search by example image (Gao, Huang, Duan, and Wang, p. 71)

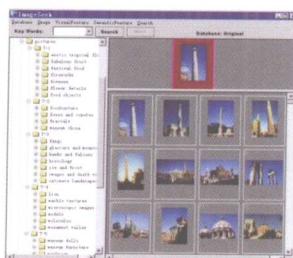
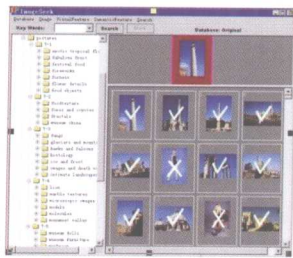


Figure 5. User interaction and feedback result

Figure 6. Search by keywords

(Gao, Huang, Duan, and Wang, p. 71)



Figure 8. NewsParser's human-machine interface (Gao, Huang, Duan, and Wang, p. 71)



Figure 9. An example of the pictorial catalog
(Gao, Huang, Duan, and Wang, p. 71)



Figure 11. Some images labeled as outdoor images
(Gao, Huang, Duan, and Wang, p. 71)



Figure 12. Some images labeled as indoor images
(Gao, Huang, Duan, and Wang, p. 71)

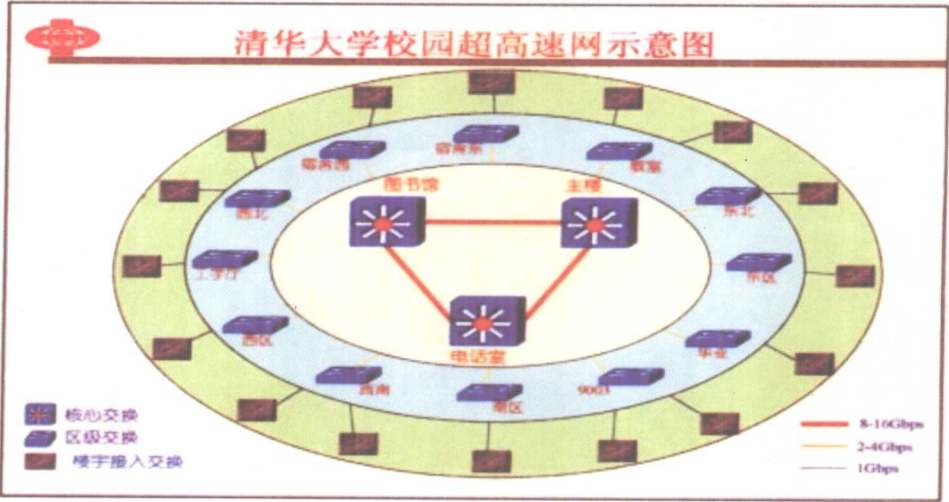


Figure 1. THU's network sketch map (Jiang and Zhang, p. 107)

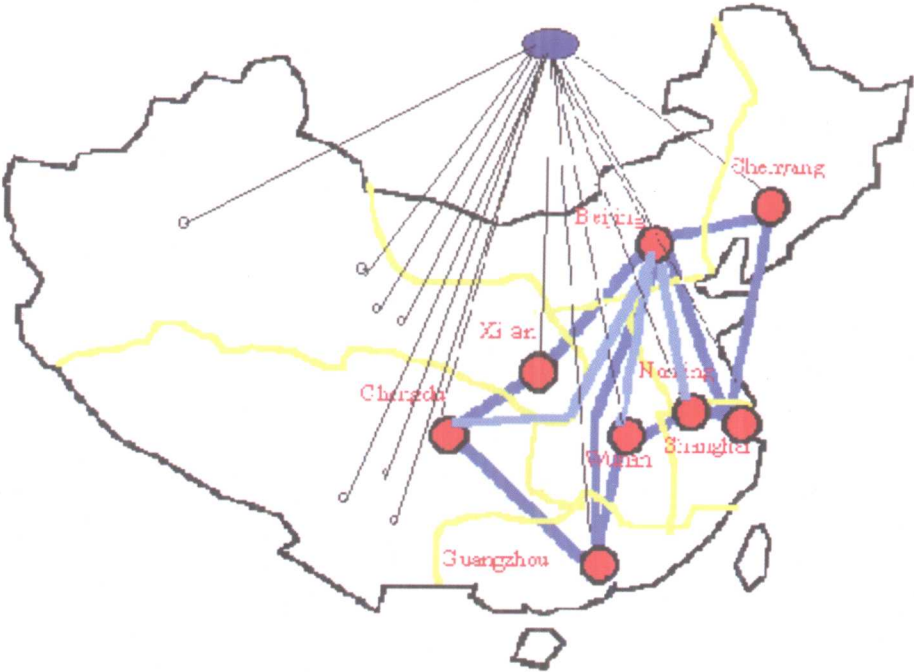


Figure 2. The CERNET Topology in 1998 (Li and An, p. 153)

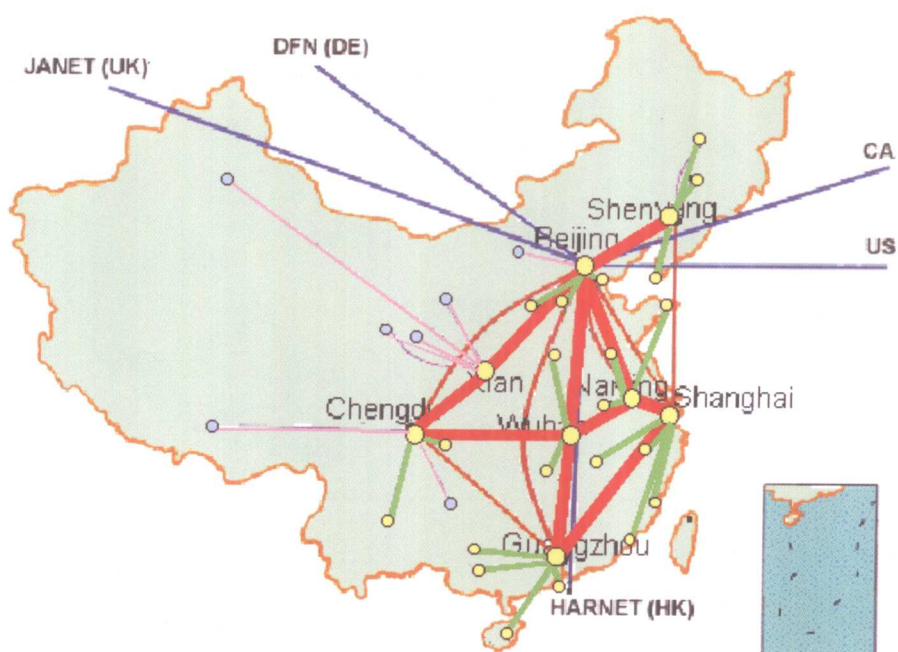


Figure 3. The current CERNET Topology (Li and An, p. 153)

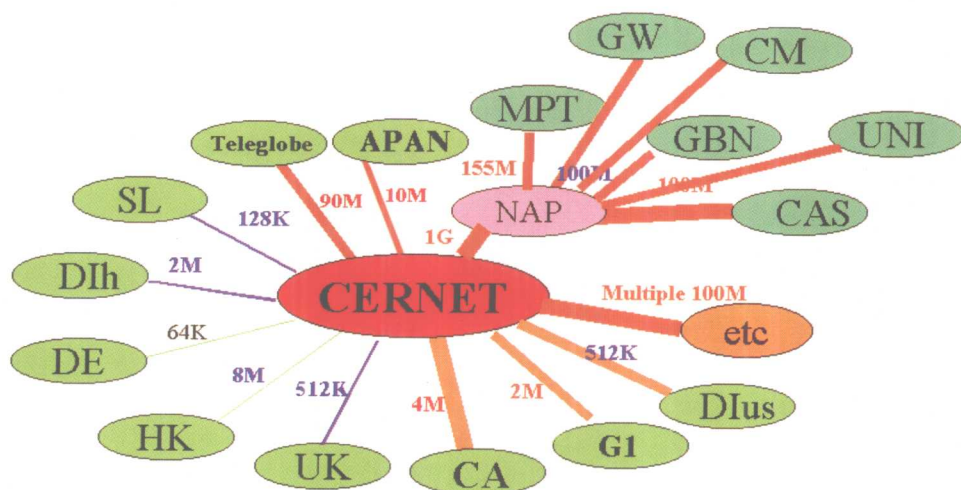


Figure 4. The CERNET External Links (Li and An, p. 153)

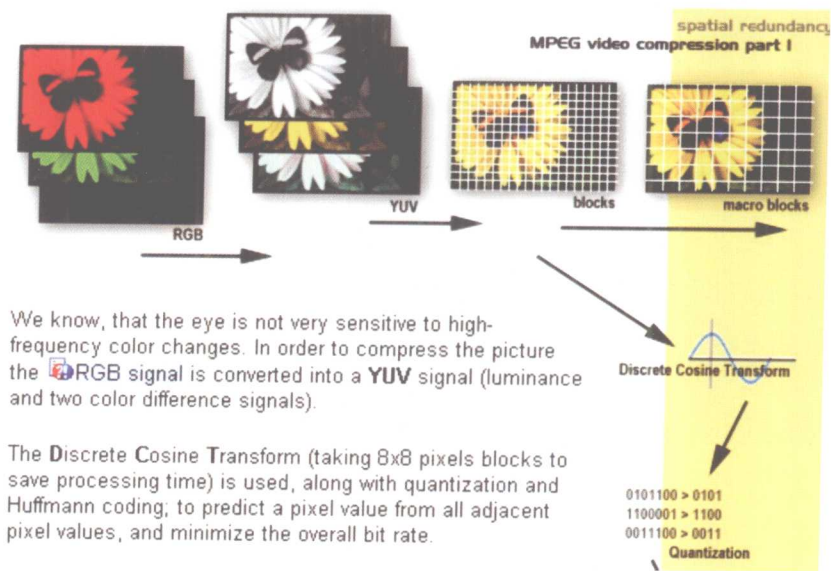


Figure 2. The "RGB Signal" icons which can produce the Q&A screen of Figure 3.
(Maurer, p. 165)

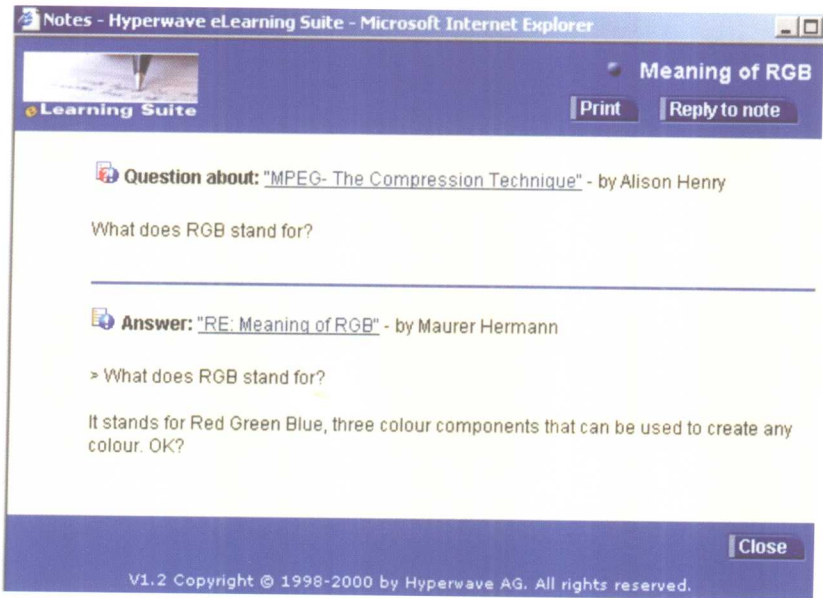


Figure 3. The question/answer dialogue attached to the document in Figure 2.
(Maurer, p. 165)