

# Report on ICDL 2004 (International Conference on Digital Libraries)

[Organized by TERI (The Energy and Resources Institute), New Delhi in partnership with Department of Culture, Ministry of Tourism and Culture, Government of India]

**New Delhi, 24-27 February 2004**

[www.teriin.org/events/icdl](http://www.teriin.org/events/icdl)

## Introduction

TERI in partnership with Department of Culture, Ministry of Tourism and Culture, Government of India organized the ICDL 2004 (International Conference on Digital Libraries) during 24-27 February 2004 in New Delhi. Ministry of Communications and Information Technology, Government of India, Commonwealth Educational Media Center for Asia (CEMCA) and Centre for Development of Advanced Computing (CDAC) were also associated with the conference. The Conference was inaugurated by His Excellency Dr A P J Abdul Kalam, Hon'ble President of India. Shri Jagmohan, Hon'ble Minister of Tourism and Culture, who is also the patron of the Conference, was also present and spoke at the inaugural function. Mr Dhanendra Kumar, Secretary, Ministry of Tourism and Culture, Government of India, welcomed the audience while Dr R K Pachauri, Director-General, TERI gave an introduction to ICDL 2004. Prof. N Balakrishnan, Professor and Chair, Division of Information Sciences, Indian Institute of Science, Bangalore delivered the Keynote address. Prof. Hsinchun Chen, Professor, Department of MIS, University of Arizona, USA gave the special address. Mr Vinod Bhargava, Associate Director, Information Technology and Services Division, TERI, proposed the vote of thanks.



Inauguration of ICDL 2004. From left: Prof. Hsinchun Chen, Professor, Department of MIS, University of Arizona, USA; Dr R K Pachauri, Director-General, TERI; Prof. N Balakrishnan, Professor and Chair, Division of Information Sciences, Indian Institute of Science, Bangalore; **Dr A P J Abdul Kalam, Hon'ble President of India**; Mr Jagmohan, Hon'ble Minister of Tourism and Culture, Government of India; and Mr Vinod Bhargava, Additional Director, Division of Information Technology and Services, TERI.

With more than 740 registered participants it was clear that the Conference had a great impact in the digital library circles the world over. The registered participants also

included about 100 from 33 countries in Asian, African, European, and American continents and Australia. The 55 invited speakers (belonging to 16 countries) consisted of eminent speakers and information scientists in the digital library field. It was also remarked that ICDL 2004 was the biggest ever digital library conference organized so far in the world. Due to constraints of space, registration had to be frozen a fortnight before the Conference.

## **Background**

Digital libraries and knowledge management are innovations, the implementation of which is nascent in the developing countries. However, they hold the promise of becoming key technologies for knowledge creation and management in the future. It is under these circumstances that TERI proposed the organization of an international conference that would deal with all subjects related to digitization and knowledge management.

In February 2003 the Government of India launched the National Mission for Manuscripts, aimed at preserving the valuable manuscripts in India. The mission is designed to undertake elaborate surveys, run awareness programmes, and impart training in conservation and preservation techniques. Since the objectives of the proposed conference and the Mission were on the same wavelength, we approached the Department of Culture, Ministry of Tourism and Culture, Government of India, with the proposal for organizing the ICDL 2004. And we are grateful that they wholeheartedly supported the idea.

## **Objectives**

The objectives for organizing this conference were to (a) bridge the knowledge gaps between developing and developed countries; (b) initiate capacity building activities in digital libraries; (c) evolve a road map for the digitization of archives, manuscripts and libraries; (d) provide a forum for facilitating interaction amongst participants; (e) formulate recommendations on digitization technologies, Acts, and policies.

## **Themes**

Since this was going to be one of the important conferences in this part of the world, we were very careful in selecting various themes. The themes most suited to fulfil the conference objectives were selected. Expected future developments in both developed and developing countries in this field were also kept in mind while doing this. Based on the various themes to be covered in the Conference, papers were invited from different countries so that those who are already working in this area or those who would be working in the near future could brainstorm various issues at length during the Conference. The following topics and speakers were selected for the tutorials: 'Finding information through the Internet and the WWW' (by Paul Nieuwenhuysen, Vrije

Universiteit, Brussels, Belgium); 'Thesauri and Ontologies' (by Nicola Guarino, National Research Council, Padova, Italy); 'Institutional repositories' (by T B Rajashekar, Indian Institute of Science, Bangalore, India; 'Building digital library using open source software (GSDL)' (by Ian H Witten, University of Waikato, New Zealand); 'Knowledge management systems: development and applications' (by Hsinchun Chen, The University of Arizona, USA); 'Automated text classification' (by Fabrizio Sebastiani, Consiglio Nazionale delle Ricerche, Italy); 'Dublin Core' (by Stuart Weibel, OCLC Office of Research, USA); and 'Web technologies and persistent information and knowledge repositories (XML, RDF & Co)' (by Erich J Neuhold, GMD-IPSI Darmstadt, Germany).

The themes selected for the conference included 'Digital library concepts'; 'Digital divide'; 'Country initiatives'; 'Digital library policies'; 'Digital library services'; 'Digital library communities and Digital library consortium'; 'Digital library security'; 'Copyright issues and digital rights management'; 'Archiving and sharing cultural heritage and history'; 'Manuscript: archiving and digitisation'; 'E-publishing and multimedia'; 'Metadata issues'; 'Multimedia and multilingual'; 'Content management and Knowledge Management'; 'E-learning'; 'Semantics, Thesauri and Ontologies'; 'System scalability and interoperability'; 'Technology: planning, development, and Management'; 'Digital library architecture, network and technology'; 'Storage and retrieval systems'; 'Human-machine interface and User interaction'; and 'Futuristics'.

### **Invited speakers**

The first and foremost task before us after deciding to organize the conference was to get well-known and knowledgeable scientists in the field of digital library all over the world. Participation would largely depend on who the speakers are. We received very useful advice from the Conference Advisory Committee in this regard. We also sought and received useful assistance from reputed scientists in the field of digital library in this regard. We further collected data of such speakers from the Internet, proceedings of similar conferences and from published papers. Invitations were sent to them along with a copy of the background paper.

### **Registration**

The delegates of the conference belonged to three major categories, those who attended only the tutorials (these were very few), those who attended only the conference sessions, and those who attended both the tutorials and conference sessions. Ample facility for registration was made for each category of delegates. The registration opened a day before the actual conference day, to avoid last minute rush and to enable them to register at their leisure. Registration was also open on the mornings of all the conference days.

## **Contributed Papers**

In response to our invitation and call for papers, 235 extended abstracts were received from India and abroad. These abstracts were sent to members of a review committee, which consisted of experts in the field from different parts of the world. The abstracts were sent for review with the name and affiliation of the author removed, to facilitate an impartial assessment. Taking into consideration the remarks of the review committee, 80 papers were accepted for presentation in the conference. All of them were requested to send the full paper within the specified time for inclusion in the pre-conference proceedings. Apart from these, 75 entries were selected for poster presentation. The posters presented were judged by an international panel and six posters were selected as the best of the lot. The authors of these six posters were given awards.

## **Pre-conference Proceedings**

It was decided that the pre-conference proceedings would be published with all the invited papers, contributed papers and poster presentations. We were glad that most of the papers could be included in the proceedings. Working backwards, 31 December 2003 was fixed as the deadline for receiving the papers. The date was further extended later as special cases in view of specific requests received from some speakers. A few papers, even then, could not be included since these were not received in time. The proceedings consisted of three volumes, one for tutorials and the other two for invited and contributed papers and poster presentations. A total of 115 papers and 73 abstracts were included in the proceedings. The proceedings were distributed to the delegates at the time of registration. (A few copies of the proceedings are available for sale. Please contact [outreach@teri.res.in](mailto:outreach@teri.res.in) or [icdl2004@teri.res.in](mailto:icdl2004@teri.res.in) for details.)

## **Programme**

The programme was divided into tutorials and conference sessions. The first day of the conference (day 1) was devoted for tutorials and inauguration. Conference sessions were conducted during days 2, 3, and 4. As per the practice followed by most such international conferences, the programme was divided into three parallel tracks A, B, and C. This was also necessitated due to the large number of papers received. Track A dealt with less technical topics whereas Tracks B and C dealt with technical and highly technical subjects, respectively. Furthermore, each track had papers moving up from simple ones to complex ones to enable participants to catch-up with the higher-end papers. The conference sessions were divided into three tracks, A, B, and C. With more than 130 presentations including invited papers scheduled to be made in four days, time was a great constraint. Once the final draft programme was ready, it was sent to members of the review committee and international programme committee and a few others to get their opinion. Their views were taken into consideration while finalizing the programme. Invited speakers could be given only 15 minutes and contributed

speakers, 10 minutes. It was also made sure that in each session, enough time was given for questions and answers.

We realized that the time given to invited speakers was not enough. Fifteen minutes were certainly not ideal for some of them who could speak for a much longer period. But the tremendous response received both from the invited speakers as well as in terms of papers received and accepted for presentation, we could not do better. The realization came much later that four days were not enough for a conference when one receives such a tremendous response!

The conference also witnessed product presentations by various globally known software companies and database vendors and 24 exhibitors.

The delegates were also treated to two typical Indian classical dance programmes during the conference.

**Conference bulletins** for all the three conference days were prepared and uploaded on to the conference web site. Bulletins of the first two days were distributed to the participants during the conference. These were prepared based on the rapporteurs' report, and give a gist of the discussions held in each track. This will, therefore, help delegates to be informed of discussions in those tracks, which they could not attend due to the sessions running parallel.

TERI's digitization initiative as well as knowledge management activities in TERI were presented in the conference.

**Inaugural address by His Excellency Dr A P J Abdul Kalam, President of India** (Full text of the President's speech is available at <http://presidentofindia.nic.in/scripts/sllatest1.jsp?id=282>)



In his inaugural address, the President requested the audience to evolve methodologies to share our Digital Contents without any barriers. This would truly speak of the technology as an integrator of people. He further urged them to pledge that we would share information of value freely to those citizens in the less privileged nations across the world.

Some excerpts from the Hon'ble President's speech:

"In the 21st century, a new society is emerging where knowledge is the primary production resource instead of capital and labour. Efficient utilisation of existing knowledge can create comprehensive wealth for the nation in the form of better health, education, infrastructure etc. for improving the quality of life. Ability to create and maintain the knowledge infrastructure, develop knowledge workers and enhance their productivity through creation, growth and use of new knowledge will be the key factors in deciding the prosperity of this Knowledge Society.

Digital library is an important component for capturing the explicit knowledge. This has to be supplemented with the implicit knowledge to the digital library system, which will eventually get transformed into a knowledge management system.

The future Digital Libraries would have speech interface. Though this technology world over is progressing, still many more things have to be done. I would urge the learned audience to work in close collaboration to ensure introduction of Natural Interfaces to the Digital Libraries world over. It is important that we take on this mission of integrating all forms of knowledge and culture into our Digital Library.

The digitization initiative has picked up momentum in the country; this is the time we have to make consistent national policies and procedures, which will lead to effective management and control of the data leading to enhancement of national knowledge base like what we have done in IT. Policy makers should take into account of the standardization requirements, inter-operability, copyright issues, classification of documents and selection and use of number of library information systems available with various organizations in the country in different standards. I would recommend deliberation of this issue in the conference and constitution of a multi-disciplinary task force for working out the draft policy document for implementation. The policy should take into account the dynamics of the web services technology and keep provision of on-line improvement and variation in future digital library systems. The task force should have access to all available resources so that they can ensure prevention of duplication of efforts.

We have to create a Knowledge Management Grid with the Central Digital Library Data Center equipped with the comprehensive Virtual Digital Library and Knowledge Management System into which all the participating organizations are connected with broadband along with Internet connectivity.

For enabling knowledge connectivity in our rural areas, we need to have a comprehensive plan for developing new infrastructure for extending the digital library services in regional languages. These include – development of OCR (Optical Character Recognition) Software in all the Indian languages, language independent operating

system, database servers, search engines, web servers and messaging servers. This will enable the digital library initiative to percolate to the rural masses in the form of e-governance, tele-education and tele-medicine. This has to be done on a mission mode with the active participation of Government, educational institutions, R&D organizations along with private sector enterprises.

It will be useful to digitize and store textbooks in the library, which can be accessed by the students whenever required. This additional facility will enable easy availability of books to the students for instant reference and study through Internet and intranet access. We need to integrate the tele-education system with the Digital library, so that the students can read and refer to the books suggested by the teacher from a distant location through on-line e-learning services.

There are large numbers of files and records occupying valuable space in our State and Central Government offices. It is essential to segregate the important files which have to be preserved for a long duration. These segregated files have to be digitized and stored in the digital library.

Land records have to be digitized and verified with satellite imagery and stored, which needs to be linked with e-governance applications for issue, transfer, conversions and additions and deletions. It should be linked with the revenue collection, estate management, and municipal records. Digitization and information flow has to go parallel to get the real advantage of e-governance workflow in the land records and its management domain.

Digital data of a voter should be made available on-line over the Internet for all the States and Union Territories as a digital book for reference. Each State should have a dedicated site for presentation of the data accessible to all voters as well as election officers.

In this millennium when the rate of flow of new books/journals has increased substantially, there is a need to have a re-look at the lock-in period of copyright documents. This International Conference could initiate action for reducing the copyright duration substantially.

We should see that all the schools, colleges, and universities digitize their libraries in their own native languages and connect to the outside world within the next 4 years. We have to ensure availability of fiber optic cables, satellite communication and wireless infrastructure especially in remote areas. It is also essential to realize high bandwidth technology like Multiple 10 Giga Bits connectivity across the country."

The President also wished that the proceedings of this International Conference on Digital Libraries 2004, reinforce our efforts in transforming nations into knowledge societies leading to their faster development and growth.

### **Keynote addresses**

The gist of all three Keynote addresses on the three conference days are given below.

#### **Keynote Address 1: 25 February 2004**

#### **Mr Kiran Karnik, President, NASSCOM**



Digital library in a broader context is nothing but a database. The objective of the digitization process should be 'empowerment of people'. The Indian IT (information technology) industry has done remarkably well globally in recent years. But, its orientation being towards export, it has given the industry a negative image within the country.

There are technologies available, but use of the appropriate ones, not necessarily high-end ones, is the key. While conceptualizing any digitization project, it is important to have some measurement criteria for outputs as well as outcomes. Outputs should be in terms of number of documents digitized; outcomes should be in terms of ease of accessibility, the spread or the reach, and the ability to access information and knowledge. Both qualitative and quantitative goals should be incorporated into the project.

Another important dimension of the digitization process is the coordination between information technology and library science professionals. These professionals will have to work as a team to achieve the desired objectives.

A few successful examples from India were cited. In Madhya Pradesh, in an attempt to empower the rural populace, several kiosks were set up to provide information. These kiosks are totally entrepreneur-driven and a small fee is charged for their use. In Karnataka, about 20 million land records were computerized, bringing in more transparency and authenticity of ownership. In Andhra Pradesh, through the e-*Seva* project in Hyderabad, electricity bills, telephone bills, property taxes, etc., can now be paid through the Internet. In Rajasthan, many government policy documents and others have been digitized and they are accessible to the public.

### **Keynote Address 2: 26 February 2004**

**Prof. Ching-chih Chen, Simmons College, USA**



Whereas digitization and digital libraries have closer a connotation to storage and retrieval of text and images, digitization has immense potential to facilitate access to and study of cultural artefacts, be they stored in museums or in actual use in communities spread the world over. The Project Emperor 1 in China was the first major project of this kind and captured 108 000 pictures on a videodisc. Though this was offline, with the spread of the Internet, such contents can be made available to any corner of the world. Universal access implicitly assumes a largely unidirectional flow of information from the developed world to the developing world but the latter is often richer in traditional wisdom, crafts, and cultural artefacts—digital technologies have the potential to preserve that heritage and make it accessible. The concept was neatly illustrated through a triangle with people, content, and technologies as its three vortices. There is a wealth of material (content) and technologies to capture, store, and distribute it; what is needed is the willingness to share. The talk elicited queries from members of the audience about image-based retrieval versus metadata-based retrieval, about appropriate changes in the present curriculum of library and information science education, and about a national policy on sharing content.

### **Keynote 3: 27 February 2004**

**Prof. N Balakrishnan, Indian Institute of Science, Bangalore**



If technology push, crashing costs, and global connectivity are the signposts for the coming decade, digital libraries are the next step in the convergence revolution. The overall goal is to have all public knowledge online, available for free to all, everywhere. At present, there are only 50 000 000 books in the world and a few billion dollars could bring these online. However, there are several obstacles to reaching the goal of universal access. Use of copyrighted material, privacy, reliability of information, and, most important, the cost-effectiveness of the entire operation are still issues to be discussed and resolved.

In the 20th century, governments created BBC (British Broadcasting Corporation), PBS (Public Broadcasting Service), AIR (All India Radio), and also the public library system in the interest of public good. These provided compensation for artists and writers while providing free access to public. The total global expenditure in public broadcasting and public libraries exceeds a hundred billion dollars. Our ancient kings supported all poets and scholars so that people at large can benefit. It is time to find the 21st century equivalent of BBC, AIR, and PBS.

Perhaps the solution lies in the four Cs (Consortium for Compensation of Creative Contents), which suggests setting aside 25% of the current national expenditure on public broadcasting and public libraries. Authors are encouraged to put the work on the Web after a few years of commercial exploitation and in return may be exempted from tax.

Knowledge multiplies whenever bits are circulated on the web. The Internet and the four Cs are necessary to push each other towards the exponential increase in connectivity and knowledge. Technology follows the law of accelerating returns, not the law of diminishing returns. The 'expanding universe theory' means that the time span of a technology will be shortened.

## **Technical Sessions**

The gist of discussions during the technical sessions are given below:

### **25 February 2004**

#### **Digital Library Concept**

Development in information and communication technology has helped not only in managing classical libraries better, but also in creating, distributing, and accessing information resources in digital formats through computer networks. Libraries follow different standards and guidelines for digitization. Developing a set of adaptable yet comprehensive standards to meet the needs of all libraries is not a viable proposition. However, greater efforts are needed to ensure more standardization rather than more varieties of standards.

In the US (United States), one-third of academic libraries and one-fourth of public libraries have been digitized. Although preservation of historical resources is a primary goal in a number of digitization projects in the US, the common thread shared by many digital projects is to provide greater access than ever before to these collections.

For digitization to succeed in the developing countries, the socio-economic conditions have to be studied while choosing the correct approach or a road map to a digital library. Other challenges include technological obsolescence, media fragility, hardware and software compatibility, copyright, and problems related to formats and security.

## **Digital Divide**

Digital divide refers to the gap between individuals, households, businesses, and geographic areas at different socio-economic levels with regard to the opportunities available to them to access information and communication technology. Five types of digital divides have been identified: (1) between rich and poor nations, (2) between haves and have-nots within a country, (3) between those who can speak English and those who cannot, (4) between commercial and individual rights, and (5) between digital elite and the urban or rural poor in terms of digital development itself.

The factors that produce this division are (1) technology, (2) economics, (3) information, and (4) tools. Metasearching and link resolution are two new tools that make it easier for end-users to search for information and remove much of the drudgery in searching for information and using it. Also, we need technologies to break language barriers. Participation of non-governmental organizations and public-private-academic interaction are also important in the digitization process. Specialized training programmes to upgrade skills of the professionals are essential.

The case study of Gulbarga in Karnataka highlighted the nature and extent of digital divide in an urban setting in India. For bridging the digital divide, the developing nations could use both online and off-line digital libraries. Offline digital libraries are less expensive compared to online libraries.

## **Case Studies - I: Special Libraries**

Metadata are the key component of digital libraries. To build digital library repositories, caution and cooperation are needed. Risks can be minimized by using open-source software like Fedora, which is available for free on the Internet. While digitizing is easy, scalability and quality control can become a problem if these issues are not addressed earlier. Test cases on digital library developed by C-DOT include digitization of newspaper clippings, sharing technical knowledge, and technical review and updating. Open-source software can be modified to suit organizational needs, and reusing designs is cost-effective. The Tata Memorial Hospital's case study showed that doctors relied on the library for immediate and critical information. Doctors had a tendency to search less and browse more. Journals were accessed either in portable document format or as HTML, so that figures could easily be adapted. The Tamil Nadu case study highlighted

the importance of digitizing ancient manuscripts and showed how two initiatives, namely Project Madurai and TADILNET, have been successful. It was observed that there is a need to develop an OCR (optical character recognition) package for Indian languages; there is also a need for standardization.

## **Case Studies II - Special Libraries**

The process of creating digital libraries is similar to that of the print media. Interoperability is the main challenge in the development of digital libraries. Different concepts of digital libraries exist and a clear-cut definition needs to be evolved. There is a need to develop well-developed national policies along with policies on preservation. Copyright is another issue that needs to be addressed especially from the users perspective. However, copyright of academic work should be addressed differently as against entertainment work. The challenge of digital libraries is to arrive at a consensus as regards technical and workflow formats. Vidyaniidhi was one such project developed in Kannada, which overcame the problem of fonts wherein the user did not have to worry about downloading fonts.

The digital library maintained at IIT, Kharagpur, provides access to online journals, e-books, digital documents, portals, web sites, and so on. The Dublin Core has been strongly influenced by librarians and is referred as the library standard. The library has initiated to preserve existing digital resource for long-term use by applying modern preservation techniques.

The TERI library initiative is an automated, networked library with bar coding done in house. The library also maintains various specialized online networks such as ENVIS, Mycorrhiza Network, Regulateri, and many others.

## **Digital Library Services**

As the increasing number of DLs (digital libraries) expands the overall size of the Web, with many of them not included in search engines, it becomes difficult to connect potential audience with available resources. Understanding linking and meta environments, and keeping up to date with search engines will make the available resources visible to users. This is necessary as the search engine industry continues to change its patterns of indexing, ranking, and providing access to information.

The digitized/computerized information services provided by the Indian Institute of Technology, Roorkee, India, was a case in point in the discussion, which highlighted the benefits in shifting to a digital environment for information dissemination.

Panellists discussed the 'cache architecture' for LFDL (lightweight federated digital library) to address networking with other libraries and improving performance and usability. Caching allows processing the metadata, improves search services, and enables the exposure of metadata via 'open archive initiative protocol for metadata harvesting'. This way, it is feasible to build a federated search service that is compatible with non-cooperating DLs. It was stressed that there is need to investigate various approaches to caching and also the management of cache including replacement policies.

### **E-Publishing and Multimedia**

Context-driven personalization of DL reduces the gap between the content provider and the user. The session highlighted various personalization methods that enable context-driven library access and illustrated three advanced approaches: (1) personal reference libraries, (2) collaborative content annotation, and (3) modelling and exploitation of personal web context.

Rather than seeing a 'document' in the conventional sense of being a paper-based entity, it was urged that we consider a document as 'an entity consisting of any media type appropriate to store or exchange information in a given context'.

In the context of the changing process of scholarly publishing, it was observed that there were tremendous expectations from electronic publishing. Achieving higher quality and quantity of retrieval can meet these expectations. This could be realized by creating a 'new culture': new rules, standards, and technology. The discussion also covered the development of DLs by utilizing freeware like Apache as web server, PERL packages and modules, and the Greenstone Digital Library software.

In the developing countries, where general literacy level is low, electronic literacy is a dream. But considerable effort is being made to connect to most regions of the world. Application of information technology to the collection, selection, distribution, and preservation of archival contributions offers the possibility of fostering international cooperation in this regard.

### **Metadata Issues**

As organizations go digital, thanks to the all-pervasive technology, the challenges to provide meaningful information to users in the shortest possible time become immense. That is why metadata, which is defined simply as 'data about data', are becoming important in the context of the World Wide Web.

Metadata are well-organized, descriptive information about web resources, like catalogue cards in the case of printed documents. Constantly changing content and information resources on the Internet and the lack of a single format make cataloguing, search, and retrieval difficult. A universally accepted standard format for identification and organization of data elements of Internet resources is, therefore, crucial.

XML, a mark-up language for structured information, is the basis for many new and emerging technologies. Whether it is for bibliographic databases – the print equivalents of indexes and abstracts – or RDBMA (relational database management) for storing and managing different forms of digital content, XML becomes the preferred syntax for library standards. It is a very efficient format for publishing and exchanging data and a powerful tool to access the contents of any type of relational database.

### **Digital Library Communities**

The advent of IT has revolutionized the way a library functions, with digital libraries emerging as a faster way to search, read, and disseminate information. A step beyond are digital library communities, which bridge the gaps in global information access and dissemination. While the developed countries have taken great strides in developing and strengthening information access to communities, the developing countries need to undertake more initiatives to reap the benefits.

To give a fillip to the development of a digital library community at the national level, there is an urgent need for countries like India to put in place a national apex body to coordinate and support strategic, technical, and infrastructural plans.

The key to success would lie in the selection of contents, collaboration with other partners in a sustained and organized way, putting in place proper infrastructure, and a composite approach to the digitization process.

Apart from academicians, designers too are using digitization to bring on one platform text, audio, and visual resources for multiple reference, quality presentation, and archival storage.

### **Technology: Planning, Development, and Management - I**

In the current digital era, when most of the information is easily accessible despite barriers of space and time, many libraries are also going digital. Monash University in Australia, for example, has taken a lead and digitized all its collection and is now able to provide e-books, e-journals, etc. to its students, faculty, and other professionals. According to them, such features as search and better definition of metadata have

helped their users. GSDL (Green Stone Digital Library) software is an open-source software. With features like data management and publishing for the Web and CDs, it has become one of the most popular and most used softwares around the world. Integration of CDS/ISIS database with Green Stone data has become easy with different kinds of approaches different people have developed. Library professionals are now required to widen the range of their skills if they are to effectively operate the software available for digital libraries. Information technology and library professionals have to go hand in hand to make the digital libraries a reality and a success.

## **Technology: Planning, Development, and Management - II**

Digitization of the library has become a buzzword. Several softwares are available off the shelf that can help libraries in managing and publishing data. DSpace is an open-source institutional repository system developed by Hewlett-Packard Labs and Massachusetts Institute of Technology libraries in response to the need to cope with increasing amounts of born-digital information artefacts in institutional settings. Since its release, thousands of organizations have downloaded the DSpace source code, and many DSpace sites have sprung up around the globe. The extensive functionality of DSpace, which includes data model, metadata, user management, and authorization, in addition to a robust architecture, has helped many organizations. The 'COLLATE Project', on the other hand, is set out to design and implement a Web-based collaboratory facility for archivists, researchers, and end-users working with digitized historical or cultural material. The COLLATE system offers a comfortable working environment and user interfaces for domain experts and end-users for indexing, annotating, and searching multi-format, multimedia archive material. On top of the distributed digital repository, it functions as a 'collaboratory in use', supporting distributed user groups through such dedicated knowledge management facilities as content-based access, comparison, and in-depth indexing and annotation of digitized sources.

## **Digital Library Architecture, Network and Technology - I**

Networked digital libraries must have a vision, namely 'access any document from anywhere without language and format barriers at any point of time at a low cost'. It was categorically stated that the lack of high bandwidth connectivity and proper training within the country would hamper the spread of digital libraries to remote and rural parts of the country. Creation and exploration of 'musical information spaces' for retrieval of music and musical information and searching of content effectively and efficiently are new dimensions in digital libraries. High-speed networks are essential for migration to high-tech libraries. However, for effective digitization of libraries, a 'traditional-level' (10 mbps) network has to be set up to begin with, which can later be converted into a 'middle-level' (100 mbps) network before moving into a 'high-level'

(1000 mbps) network. The session, however, concluded by accepting the need for high-end networking devices and powerful IT infrastructure for developing intangible (digital) libraries from the tangible (printed books) ones.

## **Digital Library Architecture, Network and Technology - II**

Digital environment encourages users to discover, support, create, and interact with resources from any place at any time. The basic requirements for creating and operating a digital library are collection, infrastructure, deployment, evaluation, and monitoring. Development of interactive learning objects was projected as the unique feature of the model described in the given session. In BITS (Birla Institute of Technology and Science), Pilani, metadata have to be very urgently digitized to move towards a digital library. Two types of trends, namely techno-driven (driven mainly by computer scientists) and user-driven (driven mainly by library professionals), were elaborated. However, the techno-driven changes had an edge over user-driven changes because of greater power and efficiency. The discussion covered the growth and evolution of medical digital libraries in our country. It is necessary to integrate the existing libraries with digital libraries. The session concluded by agreement on the positive aspects of digital libraries and their networking.

## **26 February 2004**

### **Country Initiatives**

The integrated digital science library of the Chinese Academy of Sciences aims to make available at every researcher's desktop exhaustive databases pointing to information held in all major STM (science, technology, medicine), including those that cover published literature on a common platform. The digital library has also established a reference service to answer users' queries online. The integrated library currently combines the resources of over 70 libraries. Besides being a repository of knowledge, the library adds to it by commissioning its own research on such topics as intellectual property and digital library standards. A cautionary note was introduced in importing STM databases: 'modify your library's policy of collection development and know how to evaluate them.' Equally salutary were the lessons learnt in another digitization project: (1) outsourcing can prevent the loss of productivity due to staff turnover, (2) it is better to scan all images to a uniformly high resolution, and (3) power spikes can be a major setback. The digital library movement may originate with the objective of serving users in one country – be it a small country like Nepal or a large one like Iran – but universal access will be the eventual goal of the movement.

## **Digital Library Policies**

As the session devoted to policy-related matters, the emphasis was naturally on broader issues and not the operational details of digitizing. In focusing all available resources on meeting technological challenges, the significance of cultural and organizational issues is sometimes lost sight of. Just as technology has brought people and resources closer, so it can separate them because of incompatible technical standards. Planning digital libraries at the national level requires a holistic and long-term perspective, participatory approach, and commitment; without that, technology can waste precious resources. 'Assessing and preserving our past for the future depends on the decisions taken today.' At the local level, the head of a library needs to appreciate that the library's mission and goals should be in harmony with those of the organization of which the library is a part. This understanding acknowledges that planning a library is rooted in the policies of the academe. As the Internet and digital documents become a part of everyday life, users increasingly tend to believe that any workstation is as good as any library: it is up to librarians to prove them wrong by making users aware of the added value that a library offers through careful selection and organization of knowledge.

## **Archiving and Sharing Cultural Heritage and History Globally**

Preservation is about linking the past with the future. Technology is fast changing. And digitization as a preservation option is not yet tested. Hence, it is safe to have multiple preservation options of data. A hybrid model of microfilms and digitization with fungibility option is ideal in the given context: microfilms could be converted to a digitized version and vice versa. Another option is to maintain the microfilmed and digitized versions together. While digitization has tremendous advantages, it is a huge task that involves managerial and labour resources and technological involvement of a large number of partners. The Russian example juxtaposed to the Albanian one shows how complex the former one turned out to be, spreading over a decade. Maintaining a standard protocol such as International Standard for Archival Digitization (General) was also difficult as the manuscripts varied in structure and character. The developing nations, when they attempt digitization, must have adequate financial resources lined up. Preservation of images like the Beijing Alleys would be ideal if they can be used for multiple applications such as sociological research, community study, and tourism industry.

## **Futuristics**

Digitization will be a leveller in terms of democratization of information. However, it could also offer unequal access to information. This is because of the varying penetration levels of digitization among societies. One of the studies found that Internet

penetration in the USA in 2001 was as high as 50%. Comparable numbers for India and countries in sub-Saharan Africa were 0.7% and 0.3%, respectively. The future of the books will be in the order books to bits to e-books. The computer science department at the University of Watako, New Zealand, has already begun working on some futuristic projects. These include realistic documents that enable users to browse and flip through pages of books by clicking the mouse, a book without words (for first-aid purposes), and hierarchical phrase-browsing. As digitization becomes the norm, machine translation will play a crucial role in digitizing large volume of data lying in all parts of the world. One of the challenges then would be to locate the needed information. Also it will be a challenge to access such information in one's own language. For these purposes, a general-purpose search engine or machine translation systems that are domain-driven need to be developed. Countries like India need a multi-language open-source software and should also work according to the principle 'rise, raise, and race', realizing the short time left to bridge the digital divide.

### **Content Management and Knowledge Management - I**

The discussion over the definition of 'IM (information management)' versus 'KM (knowledge management)' is never-ending and overlapping, but its success (whether of IM or KM) will largely depend on how efficiently we can organize, and provide access to, all the different varieties of information that are required for the day-to-day business operations as well as for making strategic business decisions. And, there are various approaches to organizing information. E-journals present new integration challenges for libraries and have certainly been a value-added service. E-journals support paper-based and electronic methods of knowledge dissemination, provide remote access, permit simultaneous use by more than one user, provide instant access, offer search capabilities, and provide such feature as links. Knowledge management is not a new concept to India; the '*gurukul*' system, the oral tradition of transmitting texts, and such treasures available today in the form of manuscripts as the *Vedas* and the *Puranas* are evidence enough of that. Digital technology is a powerful tool to disseminate, present, and possibly preserve this heritage. Economies the world over are moving from being industrial economies to information economies and finally to knowledge Economies. Transforming information and knowledge into intelligence is the fundamental basis of competition to achieve excellence and world-class performance.

### **Content Management and Knowledge Management - II**

The Internet is changing the way we live and do businesses. It gives opportunities for libraries, governments, and businesses to deliver information and services more efficiently and to interact with stakeholders: citizens, patrons, other businesses, and other government agencies. Digital governance makes the government a major user of

information technologies, a collector and maintainer of very large data sets, and a provider of critical and often unique information services to individuals, states, businesses, and other customers. The need for language-centric research in India is ever increasing because of the needs of a multilingual society. Information overload, greater demand for quality information, and movement towards knowledge-based wealth creation are some of the reasons that give rise to the need for a KM-based digital library. The ACADO digital library software developed by IIITM-K is a popular KM-based digital library software. Local commercial banks in Malaysia, by adopting KM, have increased knowledge sharing horizontally (through departments, and functional and business units) and vertically (through tiers of organizational hierarchy). The integration of KM with other existing enterprise systems in an organization, data integration, and other technological complexities make the implementation of KM in an organization a challenging task.

### **E-Learning - I**

The session started with a presentation on the Tamil Virtual University, which has emerged as an effective online platform for meaningful education. The online university offers various levels of courses for children and is a meticulously-organized repository of knowledge, with audio-visual tutorial tools and a user-friendly search facility.

Technology dictates the pace at which digital libraries are developed but the manner in which technology is used depends on information professionals. A web-enhanced delivery model for information dissemination was discussed, and the development of a scalable, highly interactive model for need-based training in future was stressed upon. The existing network of distance-education institutions – such as the Indira Gandhi National Open University – with their available infrastructure could be used to realize the targets.

The success of the AVU (African Virtual University) was cited to highlight the benefits in integrating ICTs (information and communication technologies) in information services. The AVU's electronic library support service has shown the way for other institutions to enhance the quality and usefulness of their services. Modes of cost-benefit analysis and evaluation of e-learning projects were discussed. Feedback from users was considered to be a reliable index to measure the success of such efforts.

### **E-Learning - II**

Technology, by itself, is morally neutral and does not transmit values: it is a vehicle for whatever contents we choose. ICTs offer a great opportunity to remove barriers in education and to facilitate learning. However, there still persists a wide gap between the

potential of ICTs and the effectiveness of their use. There is a need to integrate the use of ICTs with formal education systems, since the notion of formal education also works as a barrier to acceptance of e-learning by a learner.

SCRAN (the Scottish Cultural Resources Access Network) – a nationwide consortium of museums, libraries, and archives – was cited as an example of a massive project that sustained itself even two years after the grant ran out. Flexibility in adapting to the technical requirements of its target users, combined with aggressive marketing, has resulted in SCRAN being licensed to every publicly-funded school in Scotland, several in England, and over 90 educational institutions worldwide.

One significant manifestation of electronic publishing is the arrival of e-books. However, there is an air of uncertainty on part of publishers as regards the nature of the products and the market in the coming days. Citing the evolution of the e-Book Working Group – a UK national body assigned to formulate a national strategy to address the issues – the session provided insights on devising frameworks, which various countries could consider for adopting.

### **Semantics, Thesauri and Ontologies**

The session focused on the role of thesauri and ontologies for searching and accessing desired information. It is now believed that only word-based approach cannot yield precise and most-desired information. Lexical knowledge; synonyms, contextual mapping; and multi-linguality aspects together play a role in ontology-driven information access and knowledge retrieval.

Being a heterogeneous collection of data in multiple forms, a digital library should be annotated. Various graphs, acyclic ordered annotation, and key graphs may be deployed. Ontology maps an object to concepts, which is represented via symbolic language to achieve language-independent retrieval. Various technologies of ontology representation were suggested in the session and XML standard with variations specific to digital libraries were presented.

### **System Scalability and Interoperability**

In the session, while discussing system scalability and interoperability, statistical methods to achieve machine-aided translation were explained with the help of sample results. Speakers talked about metadata schemes and interoperability versus localization. A possible model to enhance interoperability and localization of metadata and metadata schemes was deliberated upon. Basic concepts of the Dublin Core metadata, Warwick framework, layered model for metadata, and registry were

explained. It was stressed that no single metadata set is sufficient; a layered model with a combination of various approaches is needed. Besides, mediation, cumulative information model mapping method, OAI-PMH, and digital library network were covered in the discussion.

### **Digital Library Security**

Reliable 'access management' was projected as a key to develop security in digital libraries. From the security point of view, the attribute-based approach is better than identity-based approach. New technologies do facilitate new data management techniques and accessibilities methods but pose enormous threat to the digital libraries in many ways. There are many weaknesses with regard to these technologies; TCP/IP, proper configuration, and network policy often result in security problems. Implementation of firewalls and additional securities were put forward as a solution to these threats. It was also suggested to put in place a disaster-recovery plan for all digital libraries to meet out emergencies such as data loss due to unforeseen reasons. When we further move up, the network traffic would also move up, which could again cause security flaw. To overcome this, a mesh topology – even for nodes – was suggested. Implementation of security policies for every organization, and following and modifying that at regular frequencies was strongly recommended during the session.

### **Human Machine Interface**

Panellists discussed hierarchies generated by humans to humans and machines to machines. The automated text categorization is essential for organizing digital libraries. Programmes capable of assigning one or more categories from a predefined set based on content of the document would lead to construction of automatic text classifiers, and as a result, formation of text categorizations would be achieved. A new authoring mechanism called 'virtual authoring' for multimedia documents was discussed. This can be used to create documents by integrating and customizing segments of existing documents. Virtual authoring provides the framework for the new information layer on top of the existing multimedia documents. Usage of existing documents in developing new documents without replication data in fair use applications (without violating copyrights) was suggested as a chain for information reuse. Virtual authoring could act as an alternative – as a new mechanism – to link multimedia information sources.

**27 February 2004**

## **Manuscript: Archiving and Digitisation**

Medieval manuscripts hold a high potential to be preserved in a digital format for knowledge organization and accessibility. However, it needs planning and execution at the government level, so that it is done in a systematic way. At present, Indian manuscripts have a global audience and their invaluable nature calls for preservation, conservation, and distribution. There are two ways of doing this: through microfilming and through digitization.

Digitization is cost-effective; it also allows bibliographic control, unhindered access to contents, and effective retrieval. Many initiatives have been taken to digitize manuscripts in the past. Oxford University, UNESCO Memory of World Heritage, and DELNET are a few players who have gone ahead and digitized their manuscripts collection.

Digitization of manuscripts also calls for image enhancement. Pre-processing includes enhancing and refining the raw image, which involves identifying and extracting blurred, stained, faded, bled-through, or transferred characters. The new PRISM method specifically identifies and analyses linear structures (line strokes). This technique works in both 3D (CT, MRI) and 2D (images) domains.

But at the ground level, especially in India, all these applications are possible only if there is a strong restructuring of all universities and public libraries. So far, a librarian has been a guardian of books, but now her role may change. Going by the speed of digitization, she may become a broker between information and the user.

Library collections of Indian universities are in transition; a new era is beginning. The Information and Library Network – INFLIBNET – has so far helped 142 university libraries that are on the way to computerization. ERNET India and UGC (University Grants Commission) are setting up the UGC Infonet. Under this programme, it is proposed to use ICTs (information and communication technologies) and the Internet to transform the learning environment from a mono-dimensional to a multi-dimensional one.

University libraries in India have come a long way. Their collection has grown in quantity as well as quality with the changing times, which will also include digital documents in the near future.

As scholarly communication changes and as users' needs change, librarians and libraries must also change to meet those needs. They will have to be flexible enough to continue changing, adapting to change as they have done over the past decades, but more rapidly and more creatively.

## Copyright Issues and Digital Rights Management

Copyright is a legal device that provides the creator of a work of art or literature, or a work that conveys information or ideas, the right to control how the work is used. Intellectual property is a capital asset albeit an intangible one but it is protected by contracts, patents, and, sometimes, by encryption and watermarking (fingerprinting) technologies. However, certain uses of a work protected by copyright do not require permission of the copyright owner when the purpose is essentially non-commercial (a teacher taking photocopies for distribution in a class, for instance).

In the pre-digital-copying age, photocopying was private, labour-intensive, and unremunerative until high-quality dry processes and copy shops came into being. By then, consumers had acquired explicit legal rights over intellectual property owned by others. Today, many rights holders view the basic technology of the Internet as their enemy. Analogue sources provide no unambiguous clues about possible infringement. As long as any digital work can be printed or played in analogue form, it can be copied. The line between copying for private use and copying for worldwide distribution is dangerously blurred.



The panel discussion in progress. Seen on the dais are (from the left): Dr Om Vikas, Ministry of Communications and Information Technology, Government of India; Dr Leena Srivastava, Executive Director, TERI; Mr K K Jaswal, Secretary, Ministry of Communications and Information Technology, Government of India (Chairperson of the Panel); Prof. N Balakrishnan, Indian Institute of Science, Bangalore, India; Prof. Norman Wiseman, University of Nottingham, UK; Mr B P Sanjay, Indian Institute of Mass Communication, New Delhi; and Prof. Tengu Mohd. T Sembok, Universiti Kebangsaan, Malaysia.

Copyright enforcement is a purely private matter within civil law. Some rights holders' organizations have taken on an active vigilante role. Universities are working to increase copyright awareness among students and faculty.

Widespread access to digital reformatting tools and the Internet has changed the balance between the publisher and the consumer. Whatever legislation is passed, and whatever technological protection is used, the management of intellectual property cannot eliminate reasonable exercise of consumer rights without risking widespread popular resistance and potential economic harm.

Computer specialists and library science specialists must be represented on all authorities, councils, committees and similar bodies concerned with formulating the Public Library Act. The Press and Registration Act, the Delivery of Books Act, and the Copyright Act must be suitably amended. Depositing a digital version of every document should be made compulsory. In the beginning, all digital versions can be offered at a fee to cover costs.

### **Storage and Retrieval Systems - I**

The quantity of information available is increasing rapidly every year. Hence, managing and archiving huge data and answering user queries are the main problems one comes across while extracting information across multiple documents. The main aim of information retrieval technologies is to provide a higher level of information to the user. IEMD (Information Extractor from Multiple Documents) is a step in this direction. IEMD synthesizes information common to a set of related documents, either qualitatively or quantitatively, and helps users find relevant information without having to go through the entire document. It is an improvement over the conventional information extraction methods. The potential applications of IEMD are identifying research trends, expanding cited references, etc. However, issues that still need to be addressed are information extraction, document standardization, and handling more abstract-level queries.

IR or information retrieval is concerned with automatic storage and retrieval of documents. IR systems and the Library of Congress Classification are currently being used. There are different mining techniques available (for example, sequence, web, or text mining). Clustering is one of the techniques used in data mining. Clustering groups similar types of data to facilitate easy access. Hence, based on the composition of keywords, simple and investigative methods to analyse and compare documents are possible.

## **Storage and Retrieval Systems - II**

Improving access to information, so that it addresses the needs of people, is a major issue in the storage and retrieval systems in digital libraries. Removing unnecessary data while keeping only relevant information, especially when there is a heavy information load (with the use of the World Wide Web and digital libraries), is also vital. In order to address these needs, a 'query preview system' and a 'collaborative filtering' method have been useful. Also, collaboration and communication are vital to successful information retrieval. Query previews give an overview of the data distribution in a data collection, preventing users from applying fruitless queries. Collaborative filtering keeps a tab on how users formulate queries. These methods help improve information access. But, in order to overcome the flaws of both the systems, a new CQP or collaborative query preview system is useful.

'Federated search' is also a path to better digitization. In this process, a user can search a library without logging on to a portal.

The use of multilingual search engines is also vital, especially with regard to countries like India where the rural masses are deprived of the knowledge available through the Web because most of it is in English. An integrated search system will also help make search engines more user-friendly.

### **User Interaction**

One of the principles declared at the WSIS (World Summit on Information Society; Geneva, 10–12 December 2003) is related to access to information. With this background, the session began by raising such questions as: Do users have access to information? If so, is information easily accessible or intractable?

When a survey (55 postgraduate students were asked to carry out a literature survey on hybrid libraries) was conducted to understand the human information-seeking behaviour, it was found that users need some basic skills in ICTs, different interfaces for e-books and for e-journals, good connectivity, and the ability to formulate a proper query.

It was suggested that the user should be trained in using digital libraries; there should be a one-stop window for searching various collections, a single gateway to search the Internet, intranets, digital libraries, and KM (knowledge management) repositories. There should also be a provision for a user-driven information access system and also a task-based information access system.

A UCD (user-centred design) should involve capturing user requirements through user research. It would help in developing user profiles. Before releasing the end-product, there should be a prototype, which would be thoroughly tested by the users. In other words, users should be involved from the beginning until the end and at all stages of designing a digital library system.

### **Digital Library Consortium**

A consortium of digital libraries will help librarians, users, vendors, and publishers tremendously, saving not only money but also effort and time. The INDEST (Indian National Digital Library in Engineering, Science, and Technology) consortium, which has 125 members, was set up in 2003 by the Ministry of Human Resource Development on the recommendation of an expert group appointed by the ministry under the chairmanship of Prof. N Balakrishnan, Indian Institute of Science, Bangalore.

Apart from regular services such as offering access to subscribed e-resources and providing technical help and in-house training for optimal use, INDEST offers specific services such as shared technology (federated searches, an application service provider, a union catalogue), joint storage (journal archives and back files), shared core collections, and much more.

The consortium also regularly monitors usage in several ways including analysis of statistics and asking users to fill questionnaires.

There is no universally acceptable pricing model for e-journals; subscriptions are available in different forms, such as electronic, print and electronic, pay-per-view, membership-based access (to current and archived files), and extra charges for value-added services.

### **Panel discussion**

The panel discussion on the last day dealt with 'digital library policy and strategic planning'. Chaired by Mr K K Jaswal, Secretary, Ministry of Communications and Information Technology, Government of India, members of the panel included reputed scientists and bureaucrats in this field, such as Prof. N Balakrishnan, Indian Institute of Science, Bangalore, India; Dr Om Vikas, Ministry of Communications and Information Technology, Government of India, Prof. Tengku Mohd. T Sembok, Universiti Kebangsaan, Malaysia, Prof. Norman Wiseman, University of Nottingham, UK; Mr B P Sanjay, Indian Institute of Mass Communication, New Delhi; and Dr Leena Srivastava, Executive Director, TERI.

The panel discussion touched upon the current status of digitization in India and went on to ascertain why and how India should build up its digital libraries. The panelists highlighted approaches and concerns that policy formulation should take into consideration.

It was felt that digital libraries should not be just about the abundance of knowledge but should also consider whose knowledge, what kind of knowledge, and at what cost it is embedded in digital formats. As these libraries are meant for different groups of people, their usefulness to society should be ascertained and some mechanisms put in place so as to plough back the benefits to communities that are the real holders of knowledge.

Experts, while discussing the issues of technology development and compatibility, did acknowledge that the issues of scaling, speech input and output, resource allocation, and of scaling intellectual property access to all kind of users still need to be resolved and so do the research issues of OCR (optical character recognition) for Indian languages, machine translation, speech recognition, automatic summarization, search engines, and handling large distributed databases, to name a few.

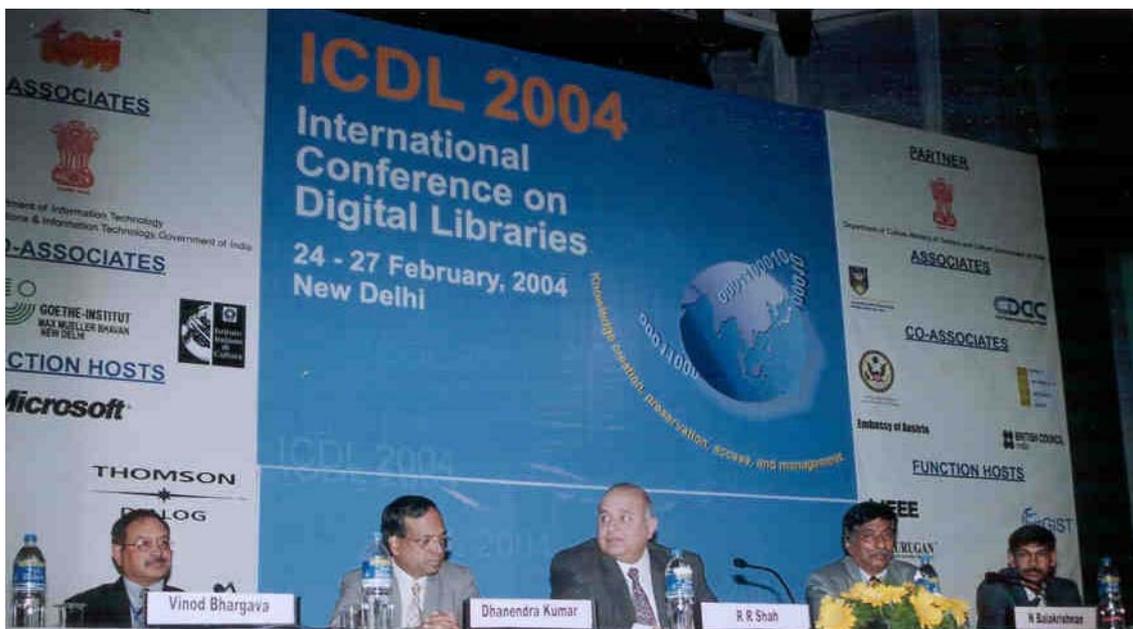
Issues identified from the perspective of the user were: identifying what user wants, structure of digital libraries, ownership and accountability of data available in these libraries, and a system in place to digitize indigenous knowledge without jeopardizing the owners' intellectual property rights so that digital libraries become a vehicle of the transformation towards a knowledge society.

### **Valedictory session**

The valedictory session was chaired by Mr R R Shah, Secretary, Planning Commission, Government of India. Prof. N Balakrishnan presented the gist of discussions of the task force constituted to make India digital. The details of this presentation are given below. Mr Dhanendra Kumar, Secretary, Department of Culture, Ministry of Tourism and Culture also spoke on the occasion. While Mr Vinod Bhargava, Additional Director, Division of Information Technology and Services, TERI welcomed the audience, Mr Debal C Kar, Conference Coordinator, ICDL 2004 proposed vote of thanks.

### **Outcome**

In his inaugural speech, the President expressed the opinion that '... we have to make consistent national policies and procedures, which will lead to effective management and control of the data leading to enhancement of national knowledge base ... Policy makers should take into account of the standardization requirements, inter-operability, copyright issues, classification of documents and selection and use of a number of library information systems available with various organizations in the country in



Valedictory Session of ICDL 2004. (From left) Mr Vinod Bhargava, Additional Director, Division of Information Technology and Services, TERI; Mr Dhanendra Kumar, Secretary, Department of Culture, Ministry of Tourism and Culture, Government of India; Mr R R R Shah, Secretary, Planning Commission, Government of India; Prof. N Balakrishnan, Professor and Chair, Division of Information Sciences, Indian Institute of Science, Bangalore; and Mr Debal C Kar, Fellow and Area Convener, Library and Information Centre, TERI and Conference Coordinator, ICDL 2004.

different standards.' He further recommended '... deliberation of this issue in the conference and constitution of **a multi-disciplinary task force** for working out the draft policy document for implementation.'

Based on these recommendations, a task force, under the Chairmanship of Prof. N Balakrishnan, Director, SERC, Indian Institute of Science, Bangalore, India, was formed and met during the conference. Several meetings of the task force were held during the conference, but separated from it, to consider the action plan needed to be adopted after the conference. These meetings were attended by a host of well-known figures in the field of digitization, officials from the Department of Culture, Ministry of Tourism and Culture, Government of India, and colleagues from TERI. Four sub-groups were constituted during these deliberations to deal with (a) resources; (b) technology; (c) users; and (d) policy and management. Another sub-group on economics and sustainability will be formed later. The gist of the recommendations made during these meetings were presented by Prof. N Balakrishnan at the valedictory session of the Conference. The salient features of the presentation are described below.

**Sub-group on Resources** [*Members - Dr S Majumdar (Central Secretariat Library), Mr H K Kaul (India International Centre), Dr S K Dhawan (National Physical Laboratory), Mr S Venkadesan (Indian Institute of Science), Dr T B Rajashekar (Indian Institute of Science), and Dr Sudha Gopalakrishnan (Indira Gandhi National Centre for Arts).*] Identification of resources need to be undertaken under cultural and literary heritage of India; Indian history and science; environment;

tourism, folklore and music; adult literacy and government publications; demography; learning; and medicine.

**Sub-group on Technology:** *[Members of this group include Mr V Ponraj (Rashtrapati Bhawan), Dr Jean-Marc Comment (Archives fédérales suisses), Dr. Usha Mujoo Munshi (Indian National Science Academy) and Dr. Gobinda G Chowdhury (University of Strathclyde).]* The main task of this sub-group will be to develop suitable technology for adoption for the task at hand. A technology should be developed for each resource.

**Sub-group on Users:** *[Members of this group include Mr Debal C Kar (TERI), Prof. Ching-chih Chen (Simmons College), and Mr Subrata Deb (TERI).]* This group will identify various users and stake holders and their needs. The users include activists from all walks of life. The group identified about 20 types of users and their requirements.

**Sub-group on Policy and Management:** *[Members of this group include Dr Michael Seadle (Michigan State University), Dr Om Vikas (CDAC), and Dr Harsha Parekh (SNDT Women's University).]* This group will deal with policy related matters, including intellectual property rights, etc.

It was intimated that an amount of 2.50 billion rupees has been set aside for the above activities for the next three years. Work on the aspects mentioned above should start immediately. The project intends to take at least 1.5 million documents of approximately 200 pages each on to the web during the period. Adequate storage facility should be provided with moderate connectivity between resource centres and users.