World Library and Information Congress:
71th IFLA General Conference and Council

"Libraries - A voyage of discovery"

August 14th - 18th 2005, Oslo, Norway

Conference Programme:
http://www.ifla.org/IV/ifla71/Programme.htm

June 15, 2005

Code Number: **132-E**
Meeting: **143 Asia and Oceania**

Digitization, sustainability and access in the Indian context

Kalpana Dasgupta
Member, RSCAO, IFLA

ABSTRACT

Digitization and digital libraries are the present buzz words in the LIS profession in India. It is the 'in' thing of today just as thirty years ago all librarians only spoke of automating the library activities. This paper will describe some major digital library initiatives in India in different LIS sectors. The Digital divide within the LIS scenario due to diverse situations will be highlighted in the paper. Most of the present initiatives are in the area of Science and Technology, in the Government sector, and in Institutions of advanced research. Most of these efforts are project based which make them very vulnerable to changes in government policies, technological changes, economic support and socio-educational changes. The need for trained personnel to undertake the new responsibilities is also a major issue. Since there is no holistic policy regarding digital content generation, management of digital libraries, standards and services, most of these efforts are isolated ventures generally restricted only to the needs of a special group of users. As most of them are started by getting a one time grant without a follow up plan, sustainability is always in doubt. Also preservation of digital data is often not focused with the seriousness that it deserves. Digitization of India's vast and rich collections for access and preservation will need very special attention from many quarters. Sustaining the present efforts and continuing these activities in all the LIS sectors will need special efforts by the decision makers, librarians and the Information Industry. Sustainability of digital libraries in India will need continuous technical and economic support. The influence of social and educational sectors and awareness at the grass root levels will be significant in sustaining digitization initiatives and digital libraries. Manpower development is a necessary factor in continuing service-oriented access through digital libraries. However, the most important factor remains the political will of the government and the sustained well informed interest of the decision makers and the library professionals.

Introduction

India has been the cradle of knowledge for thousands of years. Today there is a felt need to preserve and propagate indigenous knowledge and culture alongside the acceptance of universal knowledge through globalization in this age of technological developments. The IT oriented environment has brought in opportunities of a revolutionary nature in archiving and accessing knowledge in the digitized form which were known to exist in conventional libraries mainly in the print form. India has undertaken many initiatives to digitize its documented knowledge base and set up digital libraries for better access to its diverse clientele. However, these initiatives are seen to be sporadic and projects with one time funding. India needs well planned and policy based digitization efforts to be effective in the present situation of acute digital divide between the urban educated techno savvy minority and the vast majority who are not so fortunate. Only when the fruits of digitization of available information/documents reach the people through digital libraries, information centers and similar agencies can we consider ourselves true beneficiaries of technological advancements.

India's Digitization Initiatives

Digitization has become the buzz word in every field of information generation, processing, preservation and access. There are many prestigious nationally and internationally funded activities which are being undertaken simultaneously by many institutions and agencies. Some of these are mentioned as follows:

- The National Science Foundation of the USA has initiated the "million book project" at Carnegie Mellon University in USA, with India as partner. The Indian Institute of Science, Bangalore is the focal point of this activity in India. The project is an international cooperative attempt to digitize a million seminal books in fields that happen to be in the public domain or are copyrighted but out of print, and make them accessible through the Web globally. This will surely help in bridging the barriers in terms of time, geographical location and most importantly the economic strata of the users. The project aims at digitizing one million books in 2005 in India alone.
- The Department of Culture, Government of India has launched the 'National Mission for Manuscripts' in 2003 with the main objectives to facilitate conservation and preservation of Manuscripts for posterity as well as give central access to MSS to sensitize people about the rich cultural heritage of India. India being the largest repository of MSS in the world has innumerable ancient MSS as well as rare books, classics etc., which urgently need digitization to preserve as well as give access to. The National Informatics Centre, Government of India has prepared detailed guidelines for digitization of MSS. These guidelines cover all aspects of the digitization process starting with the selection process and includes all technical issues, quality control, collection management, organization of images, output specifications etc.
- The Central Secretariat Library under the Department of Culture has undertaken work on digitizing the government resources such as the Annual Reports, Commission and Committee Reports of the Central Government of India and Gazettes of India.

- The National Library of India has undertaken digitization of often used rare books and documents such as East India Co. records, diaries, etc. as well as MSS on paper and palm leaves.
- The Parliament Library has digitized all debates, questions, Committee Reports, bio data of present and past Members of Parliament including photographs and addresses etc.
- The Khuda Baksh Oriental Public Library has initiated digitization of Arabic and Persian MSS of medieval India.
- The Indira Gandhi National Centre for the Arts is pioneering digital treasure of Arts, digital images etc. The IGNCA will also act as the anchor for the National Mission for Manuscripts of the Government of India.
- In the field of Science and Technology efforts are many and are likely to be more consistent.
- Scholarly science journals of the Indian Academy of Sciences have been digitized from the first volume.
- The Indian National Science Academy has digitized many of its journals with support from the Department of Scientific and Industrial Research.
- MedIND Biomedical Journal of India i.e. full text of important medical journal have been digitized with support from the Indian Medlars Centre and the National Informatics Centre.
- Theses and dissertations are being digitized under different projects undertaken by the Indian Institute of Technology, Delhi and Mumbai; Physical Research Laboratory, Ahmedabad, and a special project named 'Vidyanidhi' of the Department of Library and Information Science, University of Mysore.
- There is also the e-print archives of research papers, book chapters, technical reports, unpublished findings, conference papers, journal articles etc., of the Indian Institute of Science, Bangalore.
- The National Chemical Laboratory has the digitized National collection of Industrial Micro organisms; culture collection of about 4000 strains of Algae, bacteria, fungi, yeast, etc.
- National Health Information Collaboration (NHIC) is a WHO/ICMR project for health data and Information and URDIP (CSIR Unit for Research and Development for Information Products) gives digitized open access to Indian patents and medicinal plants.
- Traditional knowledge Digital Library (TKDL) is a joint venture of the National Institute of Science Communication and Information Resources (NISCAIR) and Department of Indian Systems of Medicine and Homeopathy (ISM&H) which gives focus to preservation of community knowledge resources. It also is supposed to give legitimacy to the existing traditional knowledge and enable protection of such information.

There are interesting digital library initiatives in India:

- Honey bee: The Honey bee experiment of the National Innovation Foundation and the Indian Institute of Management, Ahmedabad focuses on innovations made by villagers to solve their problems. These are captured on video and stored as off-line digital library to disseminate knowledge to villagers in multimedia form.

- BAIF Development Research Foundation, Pune in collaboration with IDRC, Canada has set up a GIS based information base of traditional knowledge of ground water availability in districts of Maharashtra in India. This knowledge base is derived from the traditional knowledge of tribal folks of the region regarding the existence of particular tree which implies the probability of locating ground water in the area. This digital library contains mapping of ground water locations which is of great use not only to the tribal folks but also to the governmental agencies for taking effective measures for water management.
- The Mukhtabodha Project is an attempt to build a digital library in Indian languages specially the ancient texts on palm leaves, birch tree barks etc.
- Digital libraries are being set up by important institution as part of their developmental process such as Indian Institute of Technologies at Delhi, Kharagpur etc. The Energy Research Institute, New Delhi, Tata Memorial Hospital, Mumbai, etc.
- The Indian National Digital Library in Engineering Sciences and Technology (INDEST) under the Ministry of Human Resources Development is a consortium of 38 centrally funded Government institutions in the field of Science and Technology. This is a very ambitious initiative, which will subscribe to electronic resources to achieve optimal use of these resources through the participating libraries. The consortium will have cooperative subscription, resource sharing/ILL, Training and technology support as well as other shared activities for better user services and help in sustenance and continuity of the digital libraries.
- The National Resource Centre for Women set up by the Government of India is conceptualized as a Virtual Resource Centre on Women's issues. The portal would function as a decentralized, participatory and partnership oriented entity aiming at giving access to digital catalogues of different libraries dealing with women's issues, reports of diverse nature, statistics, events, legislations pertaining to women etc using different media to reach a very broad based clientele at different levels.

The initiatives undertaken by different organizations in India are mainly academic and research oriented. However, useful information for the people such as land records, cartographic materials, legal judgements, policy documents, parliamentary discussions, official forms for applicants etc., are also being digitized. Corporate houses too are digitizing their in-house working documents for creating institutional archives. However, most of these digitization initiatives are done in isolation without any coordination with other similar institutions.

While reviewing these initiatives it is seen that most of these efforts are funded either by the different government departments or important and solvent institutions. Many of them are also undertaken as one time projects with a grant, often without any proper planning for future continuity. Therefore, if one analyses these activities many unaddressed issues may be raised such as:

- the objectives of the projects;
- the rationale for content selection;
- preservation methods;
- access mechanism;

- coordination with other similar organizational initiatives;
- cost factor and sustainability etc.

Just as library automation was done in a very fragmented manner in India thirty years ago, the digitization activities are equally fragmented and diversified. This is in sharp contrast to some countries which have undertaken digitization and setting up of digital libraries in a planned manner with policies, strategies and structures well thought out, keeping in mind the holistic view of national need.

Problems and barriers

While undertaking digitization activities the Information and technical professionals have to face multiple problems and barriers in the Indian context:

1. Lack of policy framework at the national level.
2. Technological problem of obsolescence in terms of software and hardware and difficulty in upgrading the same as a recurring need.
3. Non-availability of cost beneficial new technological advancement.
4. Lack of multiple Indian language OCR facilities.
5. Non-standard technical activities, data description and transmission characteristics.
6. Non-availability of well-trained personnel with necessary skills to fully participate in the new environment.
7. Lack of proper preservation policy to sustain digitization efforts and digital libraries.
8. No IPR policy for content development of digital information for research and decision making purposes.
9. No well thought out views on the various aspects of sustainability and long-term availability of digitized material.

All these factors affect the success or failure of digitization initiatives and the creation of digital libraries. These problems have to be realistically tackled to ensure planned digital resource development programmes and digital libraries to contribute to the common goal of India's holistic socio-economic development. Therefore, the following three major issues have to be addressed in this context:

I. National level policy formulation for Digitization and creation of Digital libraries.
II. Address the sustainability issues for long term access and preservation of digital resources.
III. Ensure access to Indian information at all levels to eliminate the Digital Divide in the country.

I. Digitization Policy

The Digitization Policy for India cannot be formulated in isolation but should be within the broad spectrum of Information Policy in the new environment. Therefore, the Digitization policy also has to consider the priorities of information flow and services in the broader context of development. The main factors for development relating to the following have to be taken into consideration while preparing a policy framework:

a) Education and life long learning for all citizens

b) Information to enhance participation in the socio-economic arena in a democratic set up.
c) Business and economy, training and employment.
d) Traditional knowledge and cultural heritage preservation
e) Historical evidence and community history.

The policy document has to include the following:

1. Selection criteria – Intellectual significance of the content in terms of importance, uniqueness, timeliness and user demand. In India the form and physical material of the content is also significant considering the fact that India's heritage is available in different formats and media.
2. Technical concerns of standard infrastructure, standards, benchmarks and guidelines to be used for converting into digital documents as well as for long-term preservation.
3. Quality issues for Digitization, access and preservation.
4. Access provisions suitable for Indian material. This will include standards for metadata creation keeping in mind the diversity of Indian documents, different languages and scripts as well as the information needs of different types of clientele. Developing standard OCR facilities for Indian languages and other infrastructure for giving access to special material.
5. Legal aspects have to be included in the policy document keeping in mind the vastness of information sources; legal problems of IPR, misuse of technology for piracy, plagiarism, copyright issues of digital documents etc.
6. Preservation of the original documents such as MSS, rare books etc. after digitization.
7. Financial matters and costs of digitization, infrastructure development, sustenance of digital libraries and virtual centers, resource mobilization and recurring expenditure, pricing of services etc.
8. Development of skilled manpower to cope with the new environment and contribute in all the activities of the digitization process and preservation/sustenance of digital documents and libraries.
9. The administrative details of governance in the context of Digitization programmes and digital libraries.
10. Setting up a National Repository of Indian digital material and its organizational structure.

II. Sustainability issues

Indian digitization programmes are in their initial stages and it is now considered the most favoured activity in the field of information generation, processing, dissemination and preservation. In this situation of hype it is hardly understood that applying digital technologies will be a complex process of experimentation with gains and losses, triumphs and failures. It has become fashionable to undertake digitization projects to preserve our cultural heritage. But little is being done to prepare long-term strategy to sustain these efforts and preserve the digital documents for posterity and make them available for access in the future.

Issues concerning Sustainability

Sustainability is a broad term which refers to many factors starting from technological issues of preserving digital data to social and economic questions for long term accessibility of information to the people. The focus will have to be on the broad spectrum comprising

i. Policy and Strategies
i. Technical issues
ii. Economic aspects
iii. Social and educational issues
iv. Political and administrative will
v. Capacity building
vi. Cooperative ventures
vii. Changing role of librarians and information professionals

i. Policies and strategies

The policies and strategies for a digitization project and digital library should be such as to not meet only the proximate goals of content development, metadata creation, technical guidelines, and access to a limited clientele. The strategies should enable the professionals to create an enterprise worth sustaining. The enterprise should be so user focused that it will have value over time. The policies have also to lay down criteria for preservation because not all on-line material needs to be preserved on a long-term basis.

ii. Technical issues

Digital preservation is still a new field which is poorly understood and poorly funded. The awareness of digital preservation issues is very low among decision makers and stakeholders. Even the digital publishing arena is unclear about the responsibility of archiving. There has to be proactive actions for the preservation of digital content and media. Technically, preservation of digital information has three major aspects: medium preservation; technology preservation and intellectual preservation. The preservation of the medium on which the information is stored such as tapes, disks, optical disks, CD ROMs DVDs etc. will need the technique of 'Refreshing' which is copying to other devices of similar kind. Taking backups on a regular basis is also appropriate and required.

To overcome technological obsolescence, information will have to be migrated to the successive generations of technology. It may also require using software to emulate the behaviour of older machines; preserving original hardware and software to run obsolete programmes and creating hard copies (paper or microform) of digital objects. Each of these strategies may comply some but not all preservation goals. Moreover, both strategies of migration and emulation will need ongoing commitment and substantial resources. Just as in the analog realm, a combination of approaches will be required to ensure that digital information survives. To facilitate effective preservation a three point approach is required:

a. Use current state-of-art standards to create digital objects
b. Monitor standards as they change
c. Migrate to new standards as they are established.

Preservation of intellectual content of the digital library/centre addresses the authenticity and integrity of the information as originally recorded. Just as technology is a boon in many ways it can be very problematic when it comes to intellectual property preservation. A flawless copy with undetectable changes is always a possibility in the present situation of technological advancement.

iii. Economic aspects

Presently, the Indian Digitization scenario is very favourable. The Government of India and other important institutions and major libraries are taking significant initiatives to start digitization programmes, set up digital libraries, and host websites and portals on different areas of research. Many of these initiatives have been undertaken with one time institutional, national or international grants, with very little planning for recurring expenditure for sustenance.

The cost of digitization will have to be worked out on the basis of:

a. Nature of the original document.
b. Format, size and characteristics of the digital file.
c. Post digitization processing and preservation of the original copy.
d. Changing cost of technology.
e. Preservation and maintenance costs.
f. Prices for users' services which is dependant on the size and type of user groups.

Cost assessment will have to be done keeping in mind the functionality, demand and expense. Resource mobilization will be a major issue and different possibilities will have to be considered to deal with the on-going expenditure for sustaining the programme for the future. Economic sustainability may prove to be the toughest part of sustainability which will need serious budgeting and pricing policy also. The pricing policy well be largely dependant on the ability of users to pay for the services. There should be clear guidelines as to what to charge for as well as the type of users who will be charged.Keeping in mind the diversity of users in India, this aspect will be very significant.

iv. Social and Educational issues

Technology advances will affect the life of all the people of the country in some way or the other. The society will be torn apart and the digital divide will restrain all developments if the educational issues are not addressed on time. Necessary knowledge and skills will be a precondition for full and effective participation in the new environment by all strata of the society. India is already handicapped by the existing hierarchical caste ridden social divide which needs special attention in this regard. This is the main issue which will test the political and administrative will of the country. The governments at the Centre and at the State levels must gear up the administrative machinery to give all support not only to start projects with great fanfare but also making the planning machinery aware of the responsibilities to prepare proper strategies to make these initiatives successful and sustainable. Only such massive countrywide educational and related initiatives can ultimately help in improving the quality of participation by the

people to utilize the newly found technology driven information for the betterment of the society as a whole.

v. Capacity Building

There is an urgent need to develop skilled manpower to handle the newly created resources. Effective human resources development depends on empowerment and knowledgability. A better strategy for sustained long term human development is to create the capability of creating relevant information collections, understand the technical aspects of information generation, processing, dissemination and preservation. A heterogeneous group of ICT specialists, subject experts, library professionals and decision makers will be needed to make digitization programmes successful and sustainable. This will allow developing countries like India to actually actively participate in the Information society. In the Indian context training will have to be given in different aspects and at different levels. Important national agencies may have to collaborate with educational institutions to impart such diversified training. Training programmes may be designed for:

a) Training of trainers
b) Training of content developers and content organizers
c) Training of grass root level workers for dissemination of information relevant to the people and tailored for their use.

vi. Cooperative Ventures

It has been mentioned earlier that many efforts towards digitization of specialized materials are being done in isolation and digital libraries are being set up keeping in mind mainly the institutional needs. But the vastness of India and the diversity of its information needs often render such initiatives incomplete.Many of these are often abandoned due to paucity of funds and necessary expertise as well as negative attitude of the decision makers. Many of these ventures also cannot justify their existence and sustenance. Therefore cooperation amongst institutions of similar nature, related subjects and common goals can be brought together for setting up consortia to undertake the following:

1. Cooperative subscription or content development.
2. Resource sharing in Readers' services including Inter library loan at the virtual level.
3. Technology support and shared technology system.
4. Joint archives and storage facilities.
5. Shared core collections.
6. Funding for sustenance through different agencies involved with the consortium.
7. Need based management strategy and shared vision.
8. Cooperative training facilities for HRD.

Also, cross sectional approaches for partnerships may be productive and have access to different funding sources. Therefore, links between different types of educational institutions and public libraries or business sector enterprises and local organizations may be developed. In the Indian context such alliances may be able to address the social-economic and educational needs at different levels.

vii. The Charismatic Librarian

Librarians cannot remain the unassuming, quiet and withdrawn technical personnel if digitization programmes and digital libraries have to be planned and launched successfully and sustained as useful projects. Librarians are traditionally not drawn to public relations, networking and marketing. But in the present situation when the new environment expects strong, charismatic, confident project leadership, these techniques have to be acquired like the other professional techniques. However, the new librarians should be able to set up a product which will prove to be valuable and user friendly. The project has to be demand driven so that people will see their interests in the survival of that initiative over the years. The campaign for sustenance should be for a product which extends such services to its users which no other initiative has done so far.

III. Access to bridge the Digital divide

Indian society essentially believes in hierarchy and there are more than one divide within the Indian scenario. The age-old socio-religious caste system, economic class system, literates and illiterates, gender divide etc. are all fissures which still prevail within the Indian socio-cultural and politico-economic systems. Digital divide is a new phenomenon which will have long term effect as India gradually enters the Information society in a big way.

Culturally India has had oral tradition for transfer of knowledge in the early periods of civilization. However, with the advent of the written word different material such as palm leaves, birch tree bark and other natural products were used for writing which gradually evolved into writing on paper after paper was invented and then printing came into existence. Just as printing technology assured mass production of the written document and knowledge transfer to a much larger group of people, the present IT development enables a much wider reach specially through the internet. But the benefits of these facilities are restricted to the urban elite which is a miniscule minority in India.

Therefore, the digitization programmes and the digital libraries have to be planned to bridge this knowledge/information divide in the most effective manner. The major issues to be addressed in India before actually creating digital libraries with long ranging access facilities are:

a. Access to Internet and no access – Availability of telecommunication and other basic infrastructural facilities.
b. Internet user and non-user – Availability of technology and economically feasible.
c. Those who know the English language and those who do not – Educational and language barriers.

Digital libraries can over come many problems faced by traditional libraries by giving access to the digital surrogate while preserving the original; separating the information from the physical medium is a need based repackaged version; and get liberated from the difficulty of preserving all reading materials from the poor storage facilities and environmental harshness of the tropical weather. Through well planned access facilities it can also reach to those rural parts of India where traditional libraries have not reached. The digitization projects which have been undertaken to preserve our cultural heritage

and fascinating historical evidence, can provide access to such information which will bring India's diverse heritage to all citizens and thereby promote unity in diversity which is the essence of Indian civilization.

To implement all these ideas the digitization programmes and the digital libraries must have some important features which will enable users of different types and in different locales to access relevant information in a comprehensible manner.

a. Digital cataloguing have to be done in the virtual environment by adopting proper standards to ensure compatibility for retrieval and dissemination.
b. The virtual content should be available at different levels – metadata or bibliographic details of documents, abstracts, full text documents, multimedia information etc. Access criteria for digital documents would have to be worked out as per the importance of the information and user needs.
c. All available media, technology and languages should be used to reach out to the rural population. Agencies and organizations working at the grass root levels should be exposed to these facilities to empower the people with information pertaining to their livelihood.
d. Distance learning demands for online delivery of documents have to be met to make digital libraries a driving force behind the changes occurring in higher education.
e. The digital library should be so developed that it will not only be relevant in contents but its functionality and quality of interaction should make it valuable to the user on a long term basis.
f. Just as most libraries are becoming hybrid in their approach to acquisition and services, digital libraries can also become both on-line and off line according to the available infrastructure and user needs at the semi-urban and rural levels.
g. The digitization programmes and newly developed portals/websites on socio-economic issues should have technical facilities for uploading data from different levels of information generating agencies so that there is a two way approach to information access.

Conclusion

Digital technology has raised the hopes and expectations of people to face the challenges of not only bridging the gap between the information rich and the information poor in the country, but also uplifting the level of development in all its different facets. Major responsibility now rests on the decision makers, technological experts,librarians, educationists, social workers, legal experts, publishing industry as well as the local institutions to play their respective roles in bringing digital information in need based comprehensible form and language to the diverse clientele of the country. No agency can really work in isolation to reach the expected goal in the right manner. Therefore coordinating agencies may have to be established on a distributed regional basis to understand local requirements and thereby assist policy planners in preparing proper guidelines for useful and sustainable digitization programmes. The available technical

infrastructure and the networks in existence may now be utilized while initiatives for more sophisticated technology becomes successful in creating proper infrastructure to deal with the multi-lingual and multi-sectoral information required for the vast majority of Indians. Just as the audiovisual media such as TV and radio have reached every corner of India, digital technology will one day become a household facility in distant parts of the country.

Since Indian decision makers have now understood that Information is power and information based decision making has become the order of the day, the Government of India and other agencies are taking necessary steps to improve the telecommunication and other technical facilities to make IT based Information access a reality in the true sense so that there can be substantial improvement in the quality of life of every Indian.

References

1. Hamilton, Val. Sustainability for digital libraries. pp 1-5 http://eprints.relis.org/archive/00001930/01/sustainlrfinal.doc

2. International Conference on Digital Libraries 2004: Knowledge creation, preservation, access and management. 24-27 February 2004, New Delhi. Vols. 1 & 2…. New Delhi, TERI, 2004. xxiv + 1098p.

3. Limb, Peter. Digital dilemmas and solutions…. Oxford, Chandos publishing, 2004. xvii + 193p.

4. Peters, Dale and Pickover, Michele. DISA: Insights of an African model for Digital Library Development. D-Lib Magazine, Vol. 7, no.11, November 2001. pp 6-11.

5. Vyas, S.D. and Singh, D.K. Digital Libraries: Problems, Issues and challenges. *in* Proceedings of the National Conference on Information Management in e-libraries 26-27 February 2002 edited by S. Parthan and V.K.J. Jeevan. New Delhi, Allied publishers Ltd. pp 304-311.