

## Digital Library Initiatives at Higher Education and Research Institutions in India

N. Varatharajan

IGM Library

University of Hyderabad

Hyderabad -500 046, India

M. Chandrashekara

Dept. of Library Science

University of Mysore

Mysore -74, India

### Introduction

Most higher education and research institutions in India are funded by the central and state governments. Those institutions have made a significant contribution to the transmission of knowledge and to research in all fields and disciplines. Universities and research institutes have played a leading role in transforming the country into a modern industrialized and technologically-advanced state. The green revolution and tremendous progress in dairy development have made India a major food-producing country. Its development of space technology, the production and launching of indigenous satellites, and the development of peaceful nuclear energy have brought it into the forefront of technologically advanced nations to which a large number of developing countries look for training and guidance. Indian universities and institutes of higher learning support the needs and aspirations of Indian students and scholars. The libraries of those institutions also play a vital role in acquiring and disseminating information for academic and research activities. Digital libraries are a way of making educational and research data and information available to faculty, researchers, students, and others at the institutions and worldwide.

### Digital Libraries

Traditional methods of collecting, storing, processing, and accessing information have undergone a massive transformation due to the growth of virtual libraries, digital libraries, online databases, and library and information networks. Digital technology, Internet connectivity, and physical content can now be dovetailed, resulting in a digital library. Digital libraries and the digitisation of print materials can preserve resources in art and culture, education, science and technology, literature and humanities, media and entertainment, and cultural heritage and history. In India, a substantial number of libraries and information centres have initiated digital library projects including databases and e-journals, or by digitizing their own archivally-valuable collections. Hundreds of thousands of ancient books and manuscripts, scores of them still preserved in palm leaves, urgently need digitization to preserve the cultural heritage of India .

### Digital Library Projects

## **Digital Library of India**

The Indian Institute of Science (IISc), Carnegie Mellon University (CMU), the International Institute of Information Technology, Hyderabad (IIITH), and many other academic, religious, and government organizations in India, a total of more than twenty "Content Creation Centres," have become partners in the Digital Library of India (DLI). The DLI seeks to preserve Indian heritage that is contained in books, manuscripts, art, and music. Each centre brings its own unique collection. This digital library is also a test-bed for Indian language research. The DLI is a leader in worldwide efforts to make knowledge free. A pilot project to scan some 10,000 books was initiated at CMU and then followed up at IISc, IIIT-H, and other organizations. All the processes involved have been perfected. The vision is to preserve all the knowledge of the human race in digital form and make that content searchable, independent of language and location, and to ensure that the cultural heritage of countries like India is not lost during the transition from paper to bits and bytes, as they were lost during a former transition of cultural content from palm leaves to paper.

So far, more than 289,000 books have been scanned, of which nearly 170,000 are in Indian languages. More than 84,000 books (25 million pages) are available on the DLI web site at the [Indian Institute of Science](#), and more than 149,000 books (43 million pages) are available on the DLI web site at the [International Institute of Information Technology](#). The link to other partner sites is also provided through a [commonly accessible website](#).

Funding for the DLI comes from multiple sources. The Office of the Principal Scientific Advisor to the Government of India is funding the project at the Indian Institute of Science. The Ministry of Communication and Information Technology (MCIT) is funding the project at various DLI partner centres. The National Science Foundation (USA) is providing funding for scanners and software research and development through Carnegie Mellon University. The First Citizen of India, His Excellency Dr APJ Abdul Kalam, President, who himself is one of the contributors to this vision, has personally taken a keen interest in making the Rashtrapathi Bhavan one of the major centres of the DLI.

## **National Mission for Manuscript Digitization**

India has one of the oldest and largest collections of manuscripts in the world. These are in different languages and scripts, and written on materials such as birch bark, palm leaf, cloth, paper, etc. They are in libraries, museums, monasteries, and in the collections of individuals. A significant portion is not archivally preserved. Experts estimate that most palm leaf manuscripts will perish due to wear and tear over the next 50 to 100 years. The National Mission for Manuscripts has taken a significant step. The Department of Culture, and Ministry of Tourism and Culture, Government of India, launched the National Mission for Manuscripts in February 2003. The objectives of the mission are to facilitate conservation and preservation of manuscripts through training, awareness, and financial support; to document, catalogue, and promote access to Indian manuscripts and to encourage scholarship and research. In addition, the Central Secretariat Library (CSL), in the Department of Culture, has undertaken the massive task of digitizing government document resources. Expected benefits are the creation of a National Directory of Custodial Institutions and Individuals and Subject Directories; a National Manuscript Library to provide central access; raising awareness of the rich intellectual heritage of India; providing policy inputs to conserve, preserve, digitized, improve access, and save manuscripts for posterity; creating interest among scholars and institutions to for training in traditional Indian languages and subjects; and improving accessibility to all the stakeholders.

## **Traditional Knowledge Digital Library (TKDL)**

TKDL is a collaborative project of the National Institute of Science Communication and Information Resources (NISCAIR), the Council of Scientific and Industrial Research, the Ministry of Science & Technology and Department of AYUSH, and the Ministry of Health and Family Welfare, which this is being implemented at NISCAIR. An inter-disciplinary team of Traditional Medicine (Ayurveda,

Unani, Siddha, Yoga) experts, patent examiners, IT experts, scientists, and technical officers are involved in the creation of TKDL for Indian Systems of Medicine. The project documents the public domain traditional knowledge related to Ayurveda, Unani, and Siddha, in five international languages: English, German, French, Japanese, and Spanish. Traditional Knowledge Resource Classification (TKRC), an innovative structured classification system for systematic arrangement, dissemination, and retrieval has been developed for about 10,500 subgroups of a single International Patent Classification (IPC), i.e. AK61K35/78 for medicinal plants. TKDL is India's effort to protect its traditional medicine from foreign pharmaceutical companies who might try to copyright such medicine. TDKL will serve not merely as a source of protection for intellectual property but also as a means by which its researchers can further study and document the scientific underpinnings of the medicines and remedies in the collection.

### **Digital Library Initiative at National Library of India**

The National Library of India is a permanent repository of all material produced in India and written by Indians, and also about India written by foreign authors, wherever published in any language. The library has a large collection of publications in English and other European languages, as well as Chinese, Japanese, Arabic, and Persian. There is also a rich collection of Sanskrit, Persian, Arabic, and Tamil manuscripts and rare books. One function of the National Library is to conserve the printed heritage for future generations. The library has separate divisions for physical, chemical, reprographic, and digital conservation. Rare and brittle books and other documents are being scanned and stored on compact disc. English books and documents published before 1900 and Indian publications before 1920 are considered for digitisation. So far, 6,600 books in Indian and English languages have been scanned, with a total of over 25,000,000 pages.

### **Centre for Development of Advanced Computing (C-DAC) Digital Library of Art Masterpieces**

This is the first initiative of its kind in Asia and it will digitize 200 rare paintings of Rabindranath Tagore and Amrita Shergill from the National Gallery of Modern Arts (NGMA). The infrastructure to host this digital library will be located at the C-DAC Bangalore. C-DAC and Hewlett Packard launched the joint initiative "When Art Meets Technology" for digital preservation, restoration, and dissemination of art from the NGMA at Bangalore on February 4, 2003.

### **Indira Gandhi National Centre for the Arts (IGNCA)—Kalasampada**

IGNCA has taken up the Kalasampada Digital Library - Resource for Indian Cultural Heritage (DL-RICH) project, which is sponsored by the Ministry of Communication and Information Technology (MCIT). This project aims to develop software that will allow users to interact and explore images, audio, text, graphics, animation, and video in an integrated approach to the study of Indian art and culture. Kalasampada will facilitate access for students, scholars, artists, and the research and scientific community. The materials include several hundred thousand manuscripts, more than one hundred thousand slides, thousands of rare books, photographs, audio, and video.

### **V. V. Giri National Labour Institute**

The Archives of Indian Labour were created by the V.V. Giri National Labour Institute and the Association of Indian Labour Historians (AILH). The archive preserves documents, builds collections, and initiates research in labour history. The collections include documents from different organisations. Documents from labour movements are included, as well as personal accounts and memories of labour leaders and workers. The archive uses Greenstone, an open-source digital library system, to integrate text, audio, and video.

### **Indian Parliament Library**

This library serves members of Parliament and officers and staff of Lok Sabha Secretariat. Large databases were initially developed by the computer centre. The data are stored and available now in PARLIS (Parliament Library Information System).

### **Khuda Baksh Oriental Public Library**

This library has a rich collection of manuscripts in Persian, Arabic, Urdu and other languages. The descriptive catalogue is available in a 30-volume set which appeared in 1923 and was reprinted in the 1970s. Print catalogues were converted to machine-readable form by NICNET, which has undertaken the digitisation project. Some 400 thousand pages are now available. Documents are accessible as JPEG files. There is no retrieval mechanism for the text, since documents are treated as collection of image files. The catalogue is not searchable using a metadata scheme or any descriptors.

### **Indira Gandhi Memorial Library, University of Hyderabad**

TAs well as being the first fully automated library in India, it was the first to begin a digital library program. Since 2002 the library has digitized around 250,000 pages, primarily theses and dissertations, as well as 300 books in English and Indian languages. The library has access to about 170,000 electronic journals. The library preserves discs that accompany printed books and journals by uploading them to the CD server, which is linked to the digital library system. The library scans printed journals from Indian publishers and maintains them in the digital library as well. The library uses the open source software Dspace for its institutional repository.

Maintenance of a digital library includes content creation, designing and updating webpages, metadata creation, uploading and linking digital content. The digital library uses MARC-21 for metadata and PDF for digital content.

### **ETD and Institutional Repository**

Theses and dissertations are the bedrock of graduate education. Thesis and dissertation research is guided by experts in the field and frequently funded by highly competitive scholarships and grants. Theses and dissertations are useful sources of secondary information, particularly in the humanities, where texts are important and ideas stay current longer. Most of these works languish in college and university libraries and archives. Electronically publishing of theses and dissertations brings this valuable material more prominence. An Institutional Repository (IR) is a digital archive of the intellectual output of a university. Theses and dissertations are one basic category of material for an IR.

### **Vidyanidhi Projects**

Vidyanidhi (which means "treasure of knowledge" in Sanskrit) is a digital archive of dissertations, as well as a set of resources for doctoral research in India. Vidyanidhi is being developed as a national repository and a consortium for electronic dissertations, through participation and partnership with universities, academic institutions, and other stakeholders. Vidyanidhi began as a pilot project in 2000 with governmental support, well as support from the Ford Foundation and Microsoft India. The Ford Foundation support is for focusing on Social and Human Sciences. The Microsoft support is for the implementation of Unicode for Indian Languages. Vidyanidhi is a member of the [Networked Digital Library of Theses and Dissertations](#) (NDLTD), and [UNESCO](#) and [other](#) efforts in this direction. UNESCO supports ETD initiatives worldwide.

### **Electronic Theses and Dissertation Project of INFLIBNET Centre**

INLIBNET hosts a bibliographic database 200,000 dissertations from about two hundred Indian universities going back to 1905. The Repository uses DSpace, which complies with the Open Archives Initiative (OAI) framework allowing publications to be easily indexed and searched by web search engines and other indexing services.

### **Indian Institute of Astrophysics**

The Indian Institute of Astrophysics has its origins in the Madras Observatory, which was created in the late 18th century. Today the Institute is a national research centre for physics and astronomy. Its repository includes dissertations from researchers associated with the Institute, as well as papers from the Bulletin of the Astronomical Society India beginning with volume 1 (1973), journal articles, and conference papers. Archival materials from the 18th, 19th, and 20th centuries have recently been added. These materials are manuscripts, photographs, annual reports, instruments and their descriptions. The repository uses DSpace.

### **Raman Research Institute**

The Raman Research Institute Digital Repository allows the Institute community to deposit pre-prints, post-prints, and other publications and organizes these publications for retrieval. It also contains the annual reports of Institute and newspaper clippings from its archives. The Repository uses DSpace.

### **National Chemical Laboratory**

National Chemical Laboratory is an interdisciplinary research centre focusing on polymer science, organic chemistry, catalysis, materials chemistry, chemical engineering, biochemical sciences, and process development. It partners with industry, and some 400 graduate students are pursuing doctoral degrees. About 50 Ph.D. degrees are awarded each year. The institute has the second largest number of papers in chemical sciences (ca. 430), files the largest number of patents, both in India (60) and abroad (60) and produces the largest number of PhDs in chemical sciences in India. The repository uses DSpace. There are currently 500 theses, project reports, and journal articles available.

### **National Institute of Oceanography**

The digital repository of the National Institute of Oceanography collects and preserves institutional publications (journal articles, conference proceedings, technical reports, theses, dissertations, etc). Some of the completed and ongoing projects are:

- [Marine boundary layer characteristics during a cyclonic storm over the Bay of Bengal](#)
- [Variation of wave directional spread parameters along the Indian coast](#)
- [Study of Goa and its environment from space: A report on coastal sand dune ecosystems of Goa: Significance, uses and anthropogenic impacts](#)
- [The coastal regulation zone of Goa: Oceanographic, environmental and societal perspectives](#)
- [Marine pollution detection through biomarkers in marine bivalves.](#)

The repository uses DSpace.

### **Indian Institute of Technology, New Delhi**

Digital library initiatives began in 1998 with an upgrade to a faster Internet connection. The high-speed Internet connection led to a number of digitized collections. IITs receive grants from government bodies such as AICTE (All India Council of Technical Education) and the Ministry of Human Resources

Development and Management (MHRD) to develop digital libraries. Online courseware has been developed and older volumes of journals have been digitized, among other projects. More than 500 dissertations are available in the repository. The campus has facilities for submitting material to the repository. More than 25,000 pages of journals were scanned and are available on the Institute intranet.

### **Indian Institute of Technology, Kharagpur**

The Central Library, IIT Kharagpur, created an electronic library in 1994, which is now called a digital library. Older documents have been digitized, and it has large number of electronic resources such as *EiTech* index, *Compendex*, IEEE / IEE journals in full text, *INSPEC*, *Current Contents*, *Chemical Abstracts*, *Biotechnology Abstracts*, *Agricultural Abstracts*, *Library and Information Science Abstracts*, ASTM standards and ABI. The institutional repository collects, preserves, and disseminates research output. At present, access is restricted to the IIT Kharagpur campus LAN only and submission of documents to this repository is also limited to the IIT Kharagpur research community. The repository uses DSpace.

### **Indian Institute of Science, Bangalore**

The Institute uses [e-Prints](#), an institutional repository of research output. The archive is maintained by the National Center for Science Information (NCSI) and it supports self-archiving in various file formats (pdf, Word, html, etc.) Around 5,000 articles are available.

### **Indian Institute of Technology, Bombay**

The repository has bibliographic information and abstract for dissertations beginning in 1965. The masters thesis database has bibliographic information and abstract from 1999 on. More than 3,000 full text theses and Dissertations are available in the [ETD database](#). The repository uses Greenstone, open source software, which complies with the Open Archives Initiative (OAI) protocol.

### **Indian Institute of Management, Kozikode**

The IIM-K institutional repository uses GNU E-Prints software, which was developed at the University of Southampton . The community can archive preprints, postprints, and other scholarly publications. Anyone can access the archive, but submission of documents is limited to the IIMK research community. At present around 200 full-text documents are available in the repository.

### **National Institute of Technology, Calicut**

"Nalanda" was initiated in 1999 and is one of the largest digital libraries in the country. It serves the campus with research and other academic information in science, engineering, and technology. The software used was developed by the institute itself. Nalanda is accessible from anywhere on campus. The repository contains theses and dissertations, course materials, articles, and annual reports.

### **National Institute of Technology, Rourkela**

Formerly known as Regional Engineering College (REC), this is one of the premier institutions for technical education in the country. NIT is a joint undertaking of Government of India and Government of Orissa. This Institutional Repository uses DSpace. At present around 343 documents are available in the repository.

## Librarian's Digital Library

This repository is at the Documentation Research Training Centre, Indian Statistical Institute, Bangalore. It is aimed at librarians world-wide, and uses DSpace. It contains articles, theses and dissertations, presentations, multi-lingual documents, photographs, etc.

### **Need for Collaboration and Sustainability**

Among the initiatives described above, only a few government institutes have shown their interest. Out of three hundred universities in India, only two, the University of Hyderabad and the University of Mysore, have taken up digitization initiatives. The remaining institutions need funds, manpower, and guidelines from the UGC and their state governments. The vision should be an Indian information infrastructure linking education, research, government, and business. Questions of funding and governance, as well as technical issues, require the participation of state and national governments.

### **Conclusion**

Digital libraries and digitization are crucial for disseminating and preserving knowledge. Digital library activities are gathering momentum in developing countries, especially India. Since most higher education and research institutions in India are funded and controlled by the central and state governments, clear-cut national plans and policies are needed for infrastructure, standards, metadata, interoperability, multi-lingual databases, training, co-ordination, copyright, and archiving and preservation methods, so that our heritage of knowledge and culture can the ravages of time and present and future generations can benefit and be guided by them.