

# Development of Digital Libraries in India: a Survey of Digital Collection of National Digital Library of India

**Mohd Iqbal Bhat**  
Assistant Librarian  
J&K Cultural Academy  
Jammu and Kashmir, India  
[Iqbalnabi23@yahoo.com](mailto:Iqbalnabi23@yahoo.com)

## Abstract

*Digital preservation is ensuring the long term maintenance of a byte stream sufficient to reproduce the document and provide continued accessibility of the contents over time and through evolving computer technology. Digital Library of India (DLI) is a digital collection of freely accessible rare books and manuscripts collected from various libraries in India of different states. DLI presently is providing access to 550583 digital rare books to users throughout the World through web. But thousands of rare books and manuscripts are still there in Indian libraries without digitization. DLI needs to reach every nook and corner of India to achieve its goal of maximum digitization of Indian heritage and historical record.*

**Keywords:** Digital library, Rare books, Manuscripts, Digitization, National library

## Introduction

In India, poor access to international digital resources and the low visibility of papers are major problems facing Indian researchers and general users. Open access to digital resources is viewed as a solution to this problem. As a result, there has been a concerted effort on the part of local open access advocates and others to actively promote open access technology to maximum extent. ICT infrastructure is usually considered with respect to its main areas such as computer hardware and software application, telecommunication and networking technologies. Software installation is mandatory to create bibliography, content pages, abstracts and full-text electronic databases and their networking through Internet. It is also needed to create various union catalogues and to develop efficient software's to handle online user queries. In modern Globe, a library which wants to be part of the resource sharing network should have computer, printers, modems, scanners, database creation and its networking, software of particular standard

agencies, telephone connectivity, network access, photocopying machines and fax machines. The Open source digital library software derives its strength from several enabling technology and metadata based interoperability protocols, which have become available recently in markets of India. Digital Library of India (DLI) is a digital collection of freely accessible rare books and manuscripts collected from various libraries in India of different states.

### **What is digital Preservation**

Digitization is the presentation of an object, image, sound, document or signal (usually an analog signal) by generating a series of numbers that describe a discrete set of its points or documents. The result is called digital presentation or specifically, a digital image, for the object, and digital form, for the signal. In modern practices, the digitized data is in the form of binary numbers, which facilitate computer processing and other operations, but strictly defining, digitizing simply means the conversion of analog source material into a numerical format. The decimal or any other number system can be used instead.

Digital preservation has the two dimensions i.e. preservation of materials that are born in digital format and another is digitization for preservation of the printed documents and manuscripts, which are touched by the hand of deterioration. Digital preservation facilitates the WAN accessibility of any form of documents. For the digital preservation, two processes are followed generally ; one is creating image file of the documents and second is scanning the documents and then using the OCR process for making the text file free of error. These files are then stored in different types of secondary storage devices.

### **Literature review**

A number of studies regarding the digitization of manuscripts have been conducted in different settings, different times and for different manuscript libraries. For this paper some significant studies in the field that focus on different aspects of manuscript digitization in the Indian context have been thoroughly reviewed. Kumar and Shah (2004) have discussed in detail the Scindia Oriental Research Institute (SORI) a pioneer manuscript library of India. Some 4190 manuscripts of importance have been microfilmed by IGNCA at SORI and it has been recognized as one of

the MRCs for accessioning, cataloguing and launching of an awareness programme in Madhya Pradesh. Kumar and Shah (2004) also discussed UNESCO's digitization project 'The Memory of the World' initiated in 1993 and the manuscript digitization pilot project 'Down Memory Lane' at the National Library of India. Majumdar (2005) has described the history of artistic heritage, history of literary heritage and recorded knowledge of India and viewed that past literary heritage in the form of manuscripts available on palm leaves, cotton, silk, wood, bamboo and copper plates and has also discussed the initiatives taken by the Indian Government in introducing the NMM towards preserving and digitizing these culturally significant works. Ramana (2005) has given a brief overview of India's largest and ancient manuscript collections, the forms and places of availability of these manuscripts. He also described some indigenous methods of preserving palm leaf manuscripts and has highlighted the important benefits of digital preservation in dissemination of information, the manuscript collections of the NLI and the process of digitization of manuscripts at NLI. Nair (2006) has depicted the valuable recorded knowledge housed in different museums, archives, art galleries and manuscript libraries that are affiliated to Kerala University and has pointed out that development of a campuswide information system and opting for digitization of the valuable content would help their wider accessibility. Maltesh et al. (2007) have discussed digitization of cultural heritage, particularly manuscripts of India and other parts of the world including the UNESCO project 'Memory of the World', Czech National Library, National Library of Australia, etc. This paper also highlights the organizational role of metadata for information retrieval and access as regards manuscripts. Kumar and Sharma (2007) pointed out that digitization of manuscripts in the Indian set up is a bigger challenge than it appears. However, in the area of manuscripts, the Department of Culture, GOI made an ambitious plan in 2003 by constituting the National Mission for Manuscripts to Preserve, conserve and digitize manuscripts for posterity and described how Punjab University, Chandigarh is utilizing NMM guidelines to digitize its multilingual holdings. Mazumdar (2009) has described the manuscript collection in Assam as well as initiatives for digital preservation in Assam with reference to the Krishna Kanta Handique Central Library of Gauhati University which has about 4500 valuable manuscripts written on sanchipat, tulapat and paper. Gaur and Chakraborty (2009) have asserted that the glorious past of Indian culture lies in the ancient manuscripts which represent the basic historical evidence with great research value. It is

estimated that India possesses more than five million manuscripts, making her the largest repository of manuscript wealth in the world. In order to preserve this knowledge resource and to make these accessible to scholars, IGNCA initiated the most important manuscript microfilming programme in 1989. Gaur and Chakraborty (2009) also discussed topics like the tradition of preservation and access in India, institutional efforts in the fields of preservation and access, initiatives taken by IGNCA and NMM and challenges of manuscript preservation in the 21st century. Saikia and Kalita (2011) have highlighted the digitization process of manuscript collections in the Krishna Kanta Handiqui Library, Guahati, Assam which has 4500 copies of manuscripts on important branches of knowledge written in Assamese, Sanskrit, Bengali, Nepali and Tibetan scripts. The study also describes digitizing tools like scanners, digital cameras, image-processing software, file compression and OCR software along with digital library software like GSDL, DSpace and Eprints as well as the workflow of digitizing manuscripts. Londhe et al. (2011) have focused on the technical know-how required for digitization of manuscripts, discussed the digitization process of manuscripts adopted in the Jayakar Library, University of Pune in India and also evaluated the digitization software used in this project. Singh (2012) has depicted cultural heritage as the symbolic presence that integrates the history, traditions and culture of a country and examined the viability of preserving India's cultural heritage resources in a digital world to make it globally accessible.

### **Background of the study**

Digital Library India project was started in early 2000 with the vision to archive all the significant literary, artistic, scientific works of mankind and to preserve digitally and make them available freely for every one over Internet for education for education and research for future generations. As a first step in realizing this need, it is proposed to create the Digital Library with a free-to-read, searchable collection of one million books, predominantly in Indian languages of its different states. The Project was initiated by the Office of the Principal Scientific Advisor to the Government of India and subsequently taken over by the Department of Electronics and Information Technology, Ministry of Communications and Information Technology, Govt. of India. The idea was also to create a test bed for researchers to improve scanning techniques,

optical character recognition, intelligent indexing and in general to promote Indian Language Technology Research and finally to provide free access. Digital Library of India has currently 550,585 books with 191,657,791 pages (191.632 Million approx.) in Portable Document Format (PDF).

### **Methodology**

Online survey of Digital Library of India (DLI) was carried out. Websites of all centers, who are contributing to DLI, were evaluated thoroughly to know the actual number of contribution up to 30<sup>th</sup> Dec 2016.

### **Objectives**

1. To know the total number of digital documents contributed to DLI.
2. To examine digital documents language wise contributed to DLI.
3. To find centre wise report up to Dec 30<sup>th</sup> 2016.
4. To know contribution of top ten scanning centers to DLI.

### **Data Analysis**

There are various scanning centers across India to digitize rare books and manuscripts for digital library of India. These centers are funded by Govt. of India to carry out the digitization work. Table 1 shows the complete list of organizations involved in digitization task. It indicates the complete number of digitized books with total page numbers. A total number of 5, 50,585 digital books are available digital format with Digital Library of India. All are freely accessible in pdf format through Internet.

**Table -1 List of Institutions/organizations with digitized number of rare books and manuscripts up to 3/10/2016**

<b>Scanning Center</b>	<b>Number of Books</b>	<b>Number of Pages</b>	<b>No Pages In Millions(approx)</b>
AKCE, Universal Digital Library	7063	1026507	1.026
Allama Iqbal Library, University of Kashmir	17075	6102429	6.102
Anna University	231	106380	0.106
ASR, Melkote	1293	393852	0.393
Banasthali University	94022	37684591	37.684
Bharatiya Gyanpeeth, Delhi	18	7612	0.007
Bharatiya Jnanpith, Delhi	292	92357	0.092
C-DAC, Noida	101512	31529095	31.529
C-DAK, Kolkata	70691	28928701	28.928
Central Library, University of Delhi	52	16537	0.016
City Central Library, Hyderabad	17248	3757872	3.757
Digital Library, Mulund	492	160184	0.160
Dr. Zakir Husain Library Jamia Millia Islamia	4	613	0.000
Durga Shah Municipal Library, Nainital	465	159370	0.159
Goa University	1598	189936	0.189
Government Central Press, Mumbai	165	60743	0.060
Gujrat Vidyapith Library	1	43	0.000
Gurukul Kangri Vishwavidyalaya, Haridwar	28	5200	0.005
IGNCA, Delhi	10195	4205808	4.205
IIAp, Bangalore	1625	586047	0.586
IIIT, Allahabad	63740	20546267	20.546
IIIT, Hyderabad	2240	714886	0.714
IISc, Bangalore	4419	1617282	1.617
Kerala Sahitya Akademi, Kerala	999		
Maulana Azad Library, A.M.U	6	1832	0.001
Namgyal Institute of Tibetology, Gangtok	2	1802	0.001
NIT, Sikkim	1015	1666183	1.666
North Eastern States Libraries	30157	9595316	9.595

Osmania University	7511	2660451	2.660
Others	13	1914	0.001
Pune University	1681	496638	0.496
Rashtrapati Bhavan Library	4831	2022258	2.022
RMSC, IIT-H	54995	18947375	18.947
Salar Jung Museum	6127	2338824	2.338
SASTRA University, Thanjavur	1867	575329	0.575
Sringeri Sharada Peetham	899	282431	0.282
State Central Library, Hyderabad	8802	3080062	3.080
S.V. Digital Library, Tirupati	15096	3860044	3.860
The Record And Persion Office , War Department	4	280	0.000
University of Hyderabad	75	35140	0.035
University of Kashmir	76	19725	0.019
University of Pune	4	854	0.000
Unknown	11801	4169140	4.169
UOD, Delhi	6374	2622452	2.622
U.S. Department Of Agriculture	2	303	0.000
U.S. Department Of Education	1	1074	0.001
Veer Sewa Mandir, Dariya Ganj, Delhi	3783	1409476	1.409
<b>Total :</b>	550585	191657791	191.632

It is evident from table-1 that CDAC-Noida with 101512 is number one contributor to DLI, followed by Banasthali University with 94022 uploads and IIT, C-DAK, Kolkata is on 3<sup>rd</sup> spot which has contributed 70691 digital books. IIT, Allahabad has submitted 63740 electronic books, RMSC, IIT-H has forwarded 54995 digital books and North Eastern States Libraries have contributed 30157 digital books to digital library of India. Similarly City Central Library, Hyderabad with 17248 uploads, Allama Iqbal Library, University of Kashmir has 17075 digital contributions and S.V. Digital Library, Tirupati with 15096 digital books. IGNCA, Delhi has contributed 10195 digital books to Digital Library of India. These are all the top ten contributors to DLI.

Table-1 also indicates the least contributors to DLI, Gujrat Vidyapith Library and U.S. Department of Education have contributed only 1 digital book each, similarly Namgyal Institute of Tibetology, Gangtok and U.S. Department of Agriculture both have uploaded 2 books. The

Record And Persion Office, War Department and Dr. Zakir Husain Library Jamia Millia Islamia both have forwarded 4 digital books to DLI. Maulana Azad Library, A.M.U has contributed only 6 books to digital library of India.

Language wise break up of digital books are given in Table-2 with all statistical detail.

**Table-2 Digitization of language wise report up to 30/12/2016**

Language	Number of Books	Number of Pages	Pages in Millions
English	287715	115285335	115.285
Hindi	51385	15848226	15.848
Kannada	3335	720013	0.720
Arabic	7579	2460519	2.460
Persian	7621	2685095	2.685
Others	247	107650	0.107
Urdu	32219	9901054	9.901
Kashmiri	479	58146	0.058
Punjabi	401	92391	0.092
Sanskrit	35083	13879278	13.879
Gujarati	39457	6607737	6.607
Farsi	3	3424	0.003
Awadhi	2	323	0.000
Bhojpuri	1	160	0.000
Bengali	30389	9233793	9.233
Braj	5	2818	0.002
Danish	2	268	0.000
Dutch	7	4826	0.004
French	1012	468853	0.468
Garo	116	13746	0.013
German	405	187109	0.187
Greek	24	11502	0.011
Irish	2	1226	0.001
Italian	55	29666	0.029
Khasi	600	62512	0.062
Kokborma	1	28	0.000
Kokborok	14	3258	0.003
Malayalam	581	15653	0.015
Manipuri	2714	373734	0.373
Marathi	3904	1044999	1.044

Mizo	280	47528	0.047
Nagari	1	64	0.000
Nepali	28	4952	0.004
Norwigen	3	882	0.000
Oriya	49	18454	0.018
Pali	92	31808	0.031
Pnar	4	404	0.000
Polish	4	3312	0.003
Portuguese	1447	167332	0.167
Prakrit	2	268	0.000
Rajasthani	4	756	0.000
Roman	2	506	0.000
Russian	7	2385	0.002
Sindhi	86	32819	0.032
Spanish	17	9937	0.009
Swedish	12	5907	0.005
Tamil	5371	1286844	1.286
Telugu	23257	4461203	4.461
Tibetan	1036	1673129	1.673
Turkish	133	25004	0.025
Multilingual	1932	495998	0.495
Unknown	9307	3751250	3.751
Assamese	2152	533328	0.533
Mongolian	2	352	0.000
Nagamese	1	27	0.000
Total	550585	191657791	191.632

It is evident from Table-2 that majority of digital books contributed to DLI are in English language with 287715 digital books, followed by Hindi language which comprises of 51385 digital books and 39457 digital books are in Gujarati language. Similarly Arabic language digital books are 7579 and Persian language with 7621 digital books. In Sanskrit language there are 35083 digital books, Urdu language uploads are 32219 and Bengali language electronic books are 30389 in DLI. 23257 digital books are in Telugu language, 5371 are in Tamil language and 3335 are in Kannada language.

There is only one digital book in Nagamese , Kokborma, Nagari and Bhojpuri languages in DLI. Five languages Awadhi, Irish, Roman, Prakrit and Mongolian are having same number of digital books (i.e 2) in DLI. Similarly Pnar, Polish and Rajasthani languages have 4 digital books each in digital library of India.

## **Advantages of digital collection**

- Digital library can be accessed anywhere through www.
- Digital library provides full text search.
- Digital Library is a protected information source.
- Digital content provides efficient & qualitative services by collecting, organizing, storing, disseminating, retrieving and preserving the information.
- Digital preservation benefits besides making information retrieval & delivery more comfortable with easier access.
- Digital library provides online access to historical and cultural documents whose existence is endangered due to physical decay.
- Digital libraries necessarily include a strong focus on the management of digital content, just as traditional libraries have focused for long on the management of content in physical formats.

## **Conclusion**

Rare books and manuscripts are decaying with the passage of time. Some of them are very useful, but their damage is beyond restoration. In order to preserve their information contents it is urgently needed to digitize them as soon as possible. Digital preservation of recourses is finding ways to maintain our cultural heritage and history. Digital preservation is ensuring the long term maintenance of a byte stream sufficient to reproduce the document and provide continued accessibility of the contents over time and through evolving computer technology. If the manuscripts and rare books are digitized and made available on Internet, then they are easily accessible from every nook and corner of the entire Globe; the manuscripts and rare book users will not come to those places where the manuscripts are physically preserved in libraries. The Digital Library of India is playing a pivotal role in providing access to digitized books and manuscripts. DLI presently is providing access to 550583 digital rare books to users throughout the World through web. But thousands of rare books and manuscripts are still there in Indian libraries without digitization. DLI needs to reach every nook and corner of India to achieve its goal of maximum digitization of Indian heritage and historical record.

## References:

1. KumarMand Sharma N (2007) Digitization of manuscripts and rare literature: Initiatives of Archival Cell, Panjab University, Chandigarh (INDIA). In: 5th International CALIBER 2007, Punjab University, Chandiharh, India, 8–10 February 2007, pp. 12–24.
2. Kumar S and Shah L (2004) Digital preservation of manuscripts. In: 2nd Convention PLANNER – 2004, Manipur University, Impha, India, 4–5 November 2004, pp. 27–36. Ahmedabad:IFLIBNET.
3. Majumdar S (2005) Preservation and conservation of literary heritage: A case study of India. *International Information & Library Review* 37(3): 179–187.
4. Mazumdar NR (2009) Digital preservation of rare manuscripts in Assam. In: 7th International CALIBER, Pondichery University, Puducherry, India, 25–27 February 2009, pp. 14–19.
5. Nair RR (2006) Digitization of indigenous materials: Problems and solutions in the context of Kerala University. In: Rajan GD (ed.) *Library and Information Studies in the Digital Age: Prof. K. A. Isaac Commemoration Volume*. New Delhi: Ess Publications. pp. 1–21.
6. Ramana YV (2005) Digital preservation of Indian manuscripts – An overview. In: 3rd International CALIBER 2005, Cochin, India, 8–10 February 2007, pp. 370–376. Ahmedabad: INFLIBNET.
7. Maltesh M, Lakhar N and Gajakose S (2007) Digitization of culture. In: 5th Convention PLANNER – 2007, Gauhati University, India, 7–8 December 2007, pp. 245–253. Ahmedabad: INFLIBNET.
8. Saikia RR and Kalita B (2011) Prospects of digitizing manuscript collection in KKH Library: A model. In: 8th International CALIBER – 2011, Goa University, Goa, India, 2–4 March 2011, pp. 312–319. Ahmedabad: INFLIBNET
9. Gaur RC and Chakraborty M (2009) Preservation and access to Indian manuscripts: A knowledge base of Indian cultural heritage resources for academic libraries. In: ICAL 2009 – Vision and roles of the future academic libraries, New Delhi, India, 5–8 October 2009, pp.90–98. New Delhi: University of Delhi.

10. Singh A (2012) Digital preservation of cultural heritage resources and manuscripts: An Indian Government initiative. *IFLA Journal* 38(4): 289–296.
11. Londhe NL, Sanjay KD and Suresh KP (2011) Development of a digital library of manuscripts: A case study at the University of Pune, India. *Program* 45(2): 135–148.
12. National Library of India, Kolkata. Available at: [www. nationallibrary.gov.in](http://www.nationallibrary.gov.in) (accessed 13 January 2015).
13. National Mission for Manuscripts. Available at: [http:// www.namami.nic.in](http://www.namami.nic.in) (accessed 5 November 2014).
14. National Digital Library of India (available at <https://www.ndl.iitkgp.ac.in/>)