E-Preservation of Old and Rare Books: A Structured Approach for creating a Digital Collection

Dr Sangeeta Chakravarty,
Directorate of Historical and
Antiquarian Studies,
Guwahati, Assam, India

Abstract

Antique books, old and rare documents are fragile and vulnerable to different hazards. Preserving them for an extended period is a real challenge. From ancient times people started expressing their knowledge by writing and keeping records and subsequently started collecting and storing these at later ages as antique materials. These can be seen in different museums, libraries, archives, individual households, and other places all over the world. Preserving and conserving these antique, old, and rare books, documents etc. in good condition is a challenge for librarians, conservators, preservation administrators or persons associated with storing these. In this paper, details of the digital preservation of such a collection available in the Directorate of Historical and Antiquarian Studies (DHAS), Guwahati, Assam, India, are discussed. DHAS is a Government of Assam wing and is mainly mandated to collect, preserve and research historical and antiquarian resources. The collection of DHAS is one of the oldest collections and has been serving as a study and research centre in Assam since 1928. A special drive has been taken for the digital preservation of an identified part of the collection, with grant support from the National Archive of India. This paper discusses the entire project process starting from the project proposal formulation to the structuring of the digital collection. The paper sequentially discusses the different steps of the entire work of digitization of a collection of 241 old and rare books from the main collection of DHAS.

Correspondence should be addressed to Dr Sangeeta Chakravarty, Directorate of Historical and Antiquarian Studies, Guwahati, Assam, India. Email: sangeeta0404@gmail.com

International Journal of Digital Curation
2023, Vol. 17, Iss. 1, 8 pp.

1

Introduction

Heritage Objects are priceless knowledge of history. They are century-old legacies and reflect our cultural identity. Preserving them is the responsibility of every society and nation. Such objects hold not only knowledge but also reflections on ancient history, oral history and traditions, and the culture of our society. This is our duty to keep the materials in proper condition and preserved for centuries ahead.

Historical and Antiquarian items are generally old and vulnerable to different hazards, including environmental, biological, natural, as well as man-made. These resources are properties of national importance, and useful for understanding the evolution of mankind and society. They are also significant in terms of their physical appearance, as well as for contents of historical, cultural, and traditional information. Since most of the old documents are delicate, it is primarily difficult to handle and preserve physically because of various factors. Physically preserving them is a real challenge, and therefore before the materials and information get erased, they need to be preserved digitally.

Digital technologies are the emerging trends and are being increasingly used by museums, cultural heritage centres and libraries etc. so that the collections can be digitally preserved, as well as accessed by a larger group of people who want to explore and study online sitting from any part of the world. Once digitization of documents is done, this will save both money and manpower since there is no requirement of regular handling or taking care, as the original copy can be kept in the library untouched and safe with proper monitoring and environmental control.

Review

In India, many discrete attempts and works have been seen for the digitization of old and rare books, manuscripts, and documents. Mostly these are project-based and targeted to a bunch of books or manuscripts stored in identified location/organization or a special collection with a particular identity. National Mission for Manuscripts is an organization established in 2003 playing a vital role in the collection and preservation of manuscripts in India. The organization is under the Ministry of Tourism and Culture, Government of India. In 2004, the Mission initiated a Pilot Project of Digitization, aiming at digitizing several caches of manuscripts across the country. In 2006, the Pilot Project was completed, with the Mission of setting standards and guidelines for digitization. New projects were taken up, targeting some of the most important manuscript collections of the country[1]. For example, Jayakar Library at the University of Pune had proposed to create a digital library and digitized more than 4000 manuscripts and 2500 handwritten letters (Londhe et al., 2011).

National Library of India, Kolkata has also started a programme for digitization, named 'Down the Memory Lane' for the digitization of manuscripts and rare books (Kaur, 2015).

The Asiatic Society of Mumbai's online platform, Granth Sanjeevani, hosts approximately 2000 rare volumes in Sanskrit, Arabic, Persian, and Prakrit from various regions of India and Nepal[2].

The Indira Gandhi National Centre for the Arts (IGNCA) is a premier institution in the area of the arts and culture, and has a large collection of cultural heritage resources in a wide range of formats, including manuscripts, microfilms, micro-fiches, negatives, photographs, audio-visuals, etc. IGNCA has adopted various methods of preservation, tools, and techniques for the digitisation of its heritage resources. It is estimated that IGNCA has digitised a significant

---

[1] National Mission on Manuscripts, https://www.namami.gov.in/digitization
[2] https://granthsanjeevani.com/jspui/

part of its heritage collection which has over 5,50,000 in the form of books, periodicals, manuscripts, microfilms, images, audio, and video recordings (Ahmad et al., 2020).

National Library of India, Kolkata, National Archives of India, Nehru Memorial Museum and Library (NMML), and Rampur Raza Library are a few of the prominent national heritage institutions, with storage of India's heritage collection. These institutions have also started the process of digitising rare collections to preserve and provide access to the general public.

There are a lot of International Projects/Initiatives for digital preservation of rare old books and volumes across the globe. The Manuscript Digitization Demonstration Project sponsored by the US Library of Congress Preservation Office in cooperation with the National Digital Library Program (NDLP), US was a major manuscript digitization project in USA already completed (US Library of Congress, 1998). Since the Library's digitization efforts are initially focused on its American holdings, this project emphasized the physical types of documents found in these collections. Through this Project, the US Library of Congress has developed its capabilities for providing computerized access to its collections.

The Digitized Manuscripts Repository is a significant resource much of which has been digitized as part of the British Library's digitization projects. Almost 900 Greek manuscripts and some of the most important papyri, ranging in date from the 1st to the 18th centuries have been digitized through the Greek Manuscripts Digitization Project, generously funded by the Stavros Niarchos Foundation[3]. The Harley Science Project, funded by William and Judith Bollinger, makes available images and descriptions of 150 medieval and modern scientific manuscripts from the British Library's Harley collection (British Library, 2011).

In 2014, Penn Libraries, University of Pennsylvania was awarded a grant from the Division of Preservation and Access of the National Endowment for the Humanities, USA intended for digitization and new cataloguing of South Asian manuscripts. Nearly the entire collection is now openly available online at Penn in Hand[4].

Digital Preservation: Significance

With frequent physical access and handling of a rare old collection, we increase the vulnerability of the items. This leads to the requirement for an alternative, to provide access to the resources without having to physically handle them, and this is facilitated by digitization. Digitization simply means- converting physical objects, analogue electronic files, or documents into digital images, electronic photographs, digital texts, digital files, or e-books by using digital devices like cameras, scanners, digital software etc. Using the Digitization process, hard copy pages of books, newspapers, maps, journals, catalogues, objects, coins, manuscripts, museum objects, or analogue files of audio-videos, microfilms, recording of oral histories, or any other records of tangible and intangible materials etc. can be copied, preserved, and accessed through an electronic system.

Some important benefits of Digitization are:

- Access: Information in the books and documents can be accessed globally by all interested users and may also be used/published without having physical access to the source.

- Easy to use: large maps, books, and newspapers may be formatted so that these can be viewed conveniently in smaller sizes, e.g. by zooming in to parts. Normally such original materials are difficult to handle as they may be very delicate, heavy, large, or very tiny.

---

[3] British Library: Digitized Manuscripts: https://www.bl.uk/manuscripts/
[4] South Asian Manuscript Collection, Penn Libraries:
https://www.library.upenn.edu/collections/notable/south-asian-manuscript

- Preservation: Information embodied in books and documents can be preserved before they are lost. An original hard copy can be kept in storage in a protected environment. Not only for the researchers or visitors but digital images/copies are also useful for conservation and restoration work in case of some damage, as minor changes, deterioration, colour changes in paint, cracks etc. can be observed minutely.

- Documentation and security: Digital preservation also helps in the documentation of each object digitally, crucially securing the original resources from the alarming rise of theft, looters, and antique hunters. Some materials and records are also damaged by natural hazards or deterioration, therefore before they are vanished or lost, these should be well documented for our future generations. Digital documentation is the best way to protect our resources/artworks for years.

The Historical and Antiquarian Collections

Organisation

The Directorate of Historical and Antiquarian Studies, Guwahati, Assam, India, is one of the oldest centres established by the Government of Assam for the research and study of historical and antiquarian resources like documents, records, literature etc. The primary objective of the department is to collect and preserve all kinds of historical and antiquarian resources like old and rare books, manuscripts, documents, old coins, copper plates, maps etc. and make them available for users for studies and research. Replication, Translation, and Publication of old and rare books and manuscripts etc. are also the other objectives of this department.

These rich collections of DHAS, which are the heritage of the region, and evidence of our past, are to be preserved in a stable state without further damage by minimising the use of the original copy. It is of urgent necessity to adopt a modern method of Digital Preservation and hence requires urgent attention for scientific care with suitable methodologies, machines, and digitization of all the rare books, newspapers, manuscripts etc. Such an initiative would help in preserving the knowledge and information for the future generation, as longevity and useable condition of such old rare resources could not be guaranteed for long because of various external factors and also the delicate physical condition of these old collections.

The Collection

This directorate (DHAS) is full of historical and antiquarian resources and is housed on four floors of the building, with rare books, manuscripts, old newspapers, journals, magazines, maps, copper plate inscriptions, transcripts, etc. Manuscripts stored in DHAS are of various subjects like Mantra puthi (Mantra book), Ankar puthi (Mathematical book), religious texts, mythological texts, medicinal or Ayurveda, and most importantly the Buranjis or chronicles. Most of the manuscripts are written in Xaachi paat.[5] DHAS also has a separate section of *Ahom* Manuscripts collections, which has written literature on religion, customs, traditions, and beliefs of Tai ancestors. Ahom section has Tai and Other-Tai puthis in different materials like Xaachi paat or bark of Agor wood, Tula paat or cotton pulp, palm leaves and paper pulp manuscripts. This section is even visited by scholars and tourists from Australia, Thailand, China, and other parts of the world.

All books and documents in DHAS are very important, rare, and valuable resources with immense research and information value. These resources have lots of potential for new knowledge and fact findings and searching for hidden knowledge. They all carry lots of scopes for research and decoding. Thus, the collection is a treasure-house and source of research for research scholars, historians, antiquarians, ethnographical scholars, academicians, students etc.,

---

[5] http://www.environmentportal.in/content/38553/sanchi-paat-assam/

and requires proper facilitation for easy access. A few important parts of the entire collection, to be mentioned here are- Ancient Manuscripts (2819 numbers), Valuable Books (20,680 numbers), Very old newspapers (158 volumes), *Sadiniya Assamiya*-News Paper from 1923 to 1956- (55 vol.), *Tinidiniya Assamiya*-From 1930 to 1994 (29 Vol.), *Dainik Assamiya* -From 1946 to 1949 (18 Vol.), *Natun Assamiya* -   From 1949 to 1981 (62 Vol.), Copper Plates- 16 Nos, Transcripts (222 Nos.), Ahom & Tai Manuscripts (240 Nos).

Project Funding

There are only a few funding agencies where such a proposal for financial grants could be placed in India. National Archive of India (NAI) is one among these, undertaking regular activities in promoting digital preservation through financial support to the government as well as non-governmental organizations. NAI is the central government archive under the Ministry of Culture, Government of India, for collection and preservation of   Historical records of India. It helps in scientific management, administration, preservation, and conservation of records all over the country. We prepared a proposal against NAI's call for proposals which they publish every year. The proposed project was intended to digitally preserve an identified sub-collection of rare books from the central collection in DHAS. The Project was proposed for two years duration, and in this period, we proposed to digitally preserve at least 40,000 pages of documents/rare books available in the DHAS library. Then such copies are intended to be put on the local computer server, and visitors can have local access to this content server, which is linked with networking with the other desktop PCs. In future, the server may also be connected to the internet, so that remote access becomes possible for the users.

Project Preparation

Project Justification

The major justifications for the projects are:

1. This contributes to enhancing the longevity of the documents stored in DHAS because these physically vulnerable items shall have less physical handling when people can access their digital copies.

2. A digitally preserved book begins with the conversion of the hard content into a computerized soft format. Such a digital copy becomes searchable over the internet or on a computer. Thus, DHAS can have a digital collection of rare books which if we put it on the website at a later stage, can be accessible by the users over the internet through keyword searching.

3. To provide DHAS with the possibility to reprint the books from digital copies whenever planned.

The NAI's scheme for which the proposal was submitted was "The Grant-in-aid under the Scheme of Financial Assistance to State Government/Union territory Archival Repositories, Government Libraries, Museums, during 2018-19".
As part of the proposal, along with mandatory organizational details, the following writeup and information were also included:

2. Nature and number of manuscripts/rare books, old and rare documents, records of history, and holdings in the Directorate. Also, a list was included with the title of books, authors, inclusive years/period, subject, and language to which they belong in the custody of the DHAS.

4. Details of the project for which assistance is sought along with details of experts to be engaged.

5. Estimated cost of the Project with item-wise details of financial assistance sought.

6. Source from where 25% cost of the project would be met, was also required to be submitted to NAI. Here, as DHAS is a Government of Assam agency, it was mentioned that this part of the grant shall be borne by the state government.

List of Rare Books Proposed

Rare books were identified at first, by checking the books in the library which are very old, deteriorating and also have only single copies. Books which are with damaged covers and pages separated or torn are first fixed in the binding department before documenting. Then the books are checked and documented with information like the name of the book, the number of pages, year of publication, and the author. For the proposal of the project, a list of 100 such rare books from the library is prepared. Few such books are British Enactments in Force In Native States, Kachari Folk-Tales and Rhymes, Ao Naga, History Of Assam Rifles, Bahi-Compiled volume, Outline Grammar of Angami Naga Language, An outline Grammar of the Deori Chutiya language, Noakhali (Eastern Bengal and Assam District Gazetteers), Short Accounts of Assam, etc.

The resources for digitization have been meticulously selected considering a few defined criteria. First, the materials must not have been digitized through any previous efforts or projects. Second, the copies are really rare, and with contents related to substantial historical and antiquarian values. Another parameter based on which the final 241 numbers of resources have been shortlisted for digitization, has been the first publication year. Half of the total items were published before 1950, and the rest were also published during the 5th, 6th, and 7th decades of the last century. The last parameter of selection has been the diversity of content themes. Based on these selection parameters, the 241 numbers of volumes have been shortlisted and digitized.

Project Implementation

For digitizing the materials different devices/machines can be used according to the size and dimension of the material. In Libraries, digitization is done by different types of scanners like flatbed scanners, V- cradle scanners, and overhead scanners. Other devices required for such work are - adequate digital storage, copiers, cameras, microfilms etc. DHAS has proposed overhead book scanners, copiers etc.

Project Progress

The project was accepted by the National Archive of India and the amount was sanctioned in December 2018. As part of the Project setup, we purchased work station computer with storage of 10 TB, software licenses for the operating system-MS Windows and document processing package-MS Office, Laptop, Overhead Scanner, Colour Laser Printer, Online 5 KVA UPS, and Photocopier. We also purchased Bookshelf, Tables for the project. An external Hard Drive of 1 TB, a pen drive of 64GB capacity, Antivirus software and consumable items like Tonner, Pages etc. have been purchased for the project.

Books/Documents were cleaned and treated before sending them for digitization. The works were performed through hiring of temporary contractual staff, for digitization of 40,000 pages of rare books through scanning and generating the soft-copies book/document wise. They were first trained for scanning with the book scanner, and how to generate the book/document-wise soft files. Once a book is scanned, the scanned pages are checked one by one, whether the

pages are scanned as it is in the book by observing the page numbers. Some pages might be missing or sometimes pages are not scanned properly or are blurred. These are detected at this stage, and necessary corrective measures are taken. Similarly, sometimes it may require opening a folded page before scanning and resizing the page. Some editing was also required for removing stains/marks at the sides of pages. Sometimes it is required to change the background colour (particularly in the case of old newspapers) of the scanned pages, to make the text distinct on a white background.

Cleaning and mending were done before digitization process. Required repairing, taking care of fold-free scanning, removing insects and fungus etc. were the mandatory stages adopted before the actual process of digitization.
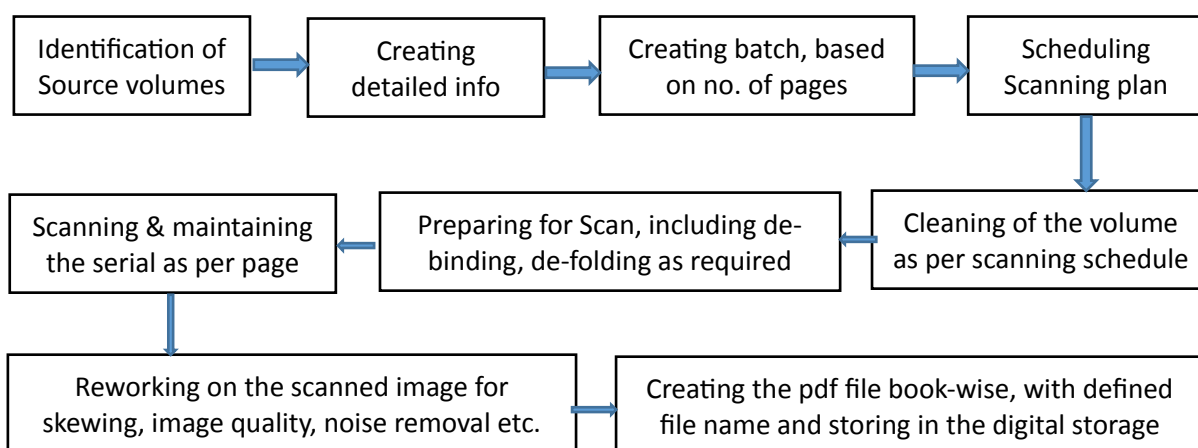

Figure 1. Process flow

The digitization of 40,664 pages of 241 rare and old books was completed by the end of 2019. The digitized books are stored in the workstation storage, after checking each digitized book and entering the details into a database using an excel sheet. All softcopies are stored in pdf format.

As NMI facilitates digital access to its repositories, supplemented by concise metadata, this digitization effort shall create new opportunities for a wider reader base with exposure to this vast knowledge base which was otherwise hidden in the DHAS treasure house. And thereby creating an expanded interest to the readers for having physical access and visits to this treasure house of rare and antiquarian learned materials of this part of the world.

Conclusion

Digitization has helped us to preserve our heritage collection and its valuable information. Advancement of imaging technologies and easy availability of large digital storage, supplemented by widely expanded digital reach has created a very conducive environment for creating large digital repositories facilitating a wider reach. Access problems had been a major bottleneck for knowledge mining, thereby restricting knowledge expansion and newer knowledge creation. And digitization projects of old rare collections have done a tremendous contribution to overcoming this. This digital preservation of DHAS rare collection has not created digital copies for preservation but also has contributed to training the staff of the department, building their capacity for similar future works, and facilitating easy access for students/researchers.

Acknowledgements

References

Ahmad, A., & Sharma, S. (2020). Sustainable Digital Preservation and Access of Heritage Knowledge in India. *DESIDOC Journal of Library & Information Technology, 40*, 321-325. DOI:10.14429/djlit.40.05.15822
https://www.proquest.com/openview/ea24e71ffded4eb75bbb077c81fc1b92/1

British Library (n.d.) Medieval manuscripts blog: https://blogs.bl.uk/digitisedmanuscripts /science/ page/9/

Kaur, H. (2015) "Digital preservation of manuscripts: An Indian perspective with special reference to Punjab," *2015 4th International Symposium on Emerging Trends and Technologies in Libraries and Information Services*, Noida, India, 2015, pp. 271-274, doi: 10.1109/ETTLIS.2015.7048210. https://ieeexplore.ieee.org/document/7048210

Londhe, N.L., Desale, S.K. and Patil, S.K. (2011), "Development of a digital library of manuscripts: A case study at the University of Pune, India", Program: electronic library and information systems, Vol. 45 No. 2, pp. 135-148.
https://doi.org/10.1108/00330331111129697

US Library of Congress (1998) Manuscript Digitization Demonstration Project: Final Report October 1998: https://memory.loc.gov/ammem/pictel/pictel.pdf