Developing a Digital Archive on Cultural Heritage Information System of Murshidabad using GSDL Software

Dibakar Das

Librarian, Sripat Singh College, Jiaganj, Murshidabad

Shuvasish Sarkar*

Librarian, Sripat Singh College, Jiaganj, Murshidabad
(*Corresponding Author)

Abstract

This paper outlines the development of a digital archive using open-source digital library software, focusing on creating a Cultural Heritage Information System for Murshidabad using GSDL (Greenstone Digital Library) Software. The archive aims to benefit heritage enthusiasts and tourists by providing up-to-date information on Murshidabad's cultural heritage. The methodology involves key processes such as documentation, digitisation, metadata creation, software customisation, testing, and promotion. Materials such as books and journal articles related to Murshidabad are selected, scanned, or typed and saved in formats like PDF, DOC, and JPG to serve as digital objects. The paper aims to preserve, make accessible, raise awareness, and engage the community in heritage preservation. It underscores the significance of preserving and sharing heritage information for current and future generations. This article is intended to assist digital archive developers, heritage organisations, and library professionals in developing and customising digital collections using GSDL, thereby supporting the establishment and enhancement of digital libraries in their respective institutions.

Keywords: Cultural heritage, Cultural heritage information system, Digital archive, GSDL, Information system, Metadata

1. Introduction

India, known for its immense cultural diversity, boasts a rich repertoire of heritage sites that reflect its historical significance. Murshidabad district in West Bengal, specifically, holds numerous cultural treasures that serve as tangible memories of the country's past (Sinha & Murdia, 2018). The preservation of these historical properties is crucial, prompting the adoption of innovative digital preservation methods as advocated by researchers (Jenifar et al., 2024). The article emphasises the integration of such methods into heritage conservation efforts across India, highlighting the potential of platforms like GSDL (Sonkar et al., 2005). Digital archives play a pivotal role in this process, converting archival materials into machine-readable formats to enhance accessibility and ensure preservation, as exemplified by projects such as American Memory by the Library of Congress. However, despite the evident benefits, there remains a significant gap in implementing these methods effectively within India's cultural heritage sites, attributed to limited awareness, inadequate resources, and technological barriers. The diverse and complex nature of these sites further complicates implementation, necessitating

tailored solutions and interdisciplinary collaboration. This study focuses on developing a digital archive specifically for the cultural heritage of Murshidabad, underscoring the future potential to extend similar systems to other lesser-known sites across the district. The initiative aims to preserve, promote, and share valuable cultural information, paving the way for future research and technological advancements in heritage conservation and digital archiving within India.

2. Review of related literature

Significant numbers of research works have been conducted on the topic of Cultural Heritage Information System. Sehaiahand Veeraanjaneyulu (2009) developed an Institutional Repository using Greenstone Digital Library Software to gather, manage, preserve, and disseminate information output within and outside of an institution. Another study focused on the organisation and dissemination of digital objects through web and CD-ROM: A framework for Indian libraries was presented in the Proceedings of the International Conference on Digital Libraries (Mukhopadhyay, 2004). The author explained here how to create digital objects for Indian Libraries that are compatible with the web and CD-ROM. Gupta (2022) conducted a feasibility study of preservation metadata in Greenstone Digital Library Software. Sonar et al. (2005) conducted a study on setting up a digital archive on the application of Greenstone Digital Library (GSDL) Software in Newspapers clippings. Jenifar et al. (2024) also conducted a study on Digital Preservation Methods for Cultural Heritage Sites in India. Lata and Somvir (2019) made a study on the creation of the digital library for the assimilation and dissemination of information using GSDL, as referenced.

The review of literature identifies that there is no such work that shows the implication of the cultural heritage information system of Murshidabad. This study attempts to fill that research gap.

3. Objectivesof the present study

The main objectives of the study are to design and develop the mechanism to build the digital archive on cultural heritage through the application of open source software sub-objectives of the study are given below:

- To understand the overviews of the digital archive on cultural heritage monuments of historical place Murshidabad
- To implement the metadata standards in the design of the digital library
- To export the Cultural Heritage monuments on CD-ROM for offline access
- To help the learners and researchers for future study as well as the information seekers to get the information about the heritage monuments of Murshidabad
- To help the touristsget information about the cultural heritage of Murshidabad by providing remote access through the internet (Sinha & Murdia, 2018).

4. Research design and methodology

We have applied self-modified metadata element sets for design and methodology of this research paper. In this study, we have utilised the Greenstone Digital Library Software (GSDL) and Greenstone Librarian Interface (GLI)as a tool. These tools help to create a searchable information product. These are available on the most durable media. The following are the five

fundamental tasks needed to develop the product:

- Gather: Storage of the heritage monuments materials.
- Enrich: Encoding each through self modified schema.
- Design: Controlled indexing of digital materials.
- Format: Designing user interface for searching and browsing.
- Create: Building of collection on religious heritage monuments.

Following this method, we can implement successfully the aforementioned design.

Installation of the system:

- Installation of Apache server.
- Installation of Java language.
- Installation ofGSDL.
- Linking of Apache and GSDL
- Installation of PERL.

Data collection:

Developer can select the content of books and journal articles relating to Murshidabad, selection materials are scanned or typed and saved in a needed format like pdf, doc, jpg, etc, scanned or typed articles are used as digital objects, these methods are done in the background.

Collection development:

- Decision on scope and coverage.
- Create HTML template to store and display religious heritage monuments resources.
- Storage HTML formatted pages in a folder.
- Recording of religious heritage monuments, following a standard format self-modified as metadata elements sets.

Organising the collection:

- Uploading the HTML templates in local space through GLI.
- Incorporation of self-modified as metadata elements set in GSDL.
- Enriching of digital objects by using my self-created metadata schema.
- Controlled indexing.
- Design of the user interface.
- Designing of display format.
- Finally testing and debugging.

5. Organisation of cultural heritage monuments of Murshidabad

The development of Information and Communication Technology (ICT) created new phenomena for preserving and access of cultural heritage and develop a method of digitisation to build a digital archive. The latest technologies have endless advantages which reflect images of architects with metadata attached and can now be efficiently used for a variety of tasks, including archiving, conservation and promotion. The idea of digitising content for promotional purposes is relative. It helps to shorten the distance between the viewer and the architect and enables multidimensional object display. Digitisation is a procedure that requires detailed instructions on how to set up the process and a well-defined set of parameters to be implemented. A sophisticated technical procedure that is specifically tailored to every collection is needed to digitise them, and this difficulty increases with the number of materials collected. It appears to be lot easier, than digital museum collections (Mukhopadhyay, 2004).

In this research work, we have tried to build up different types of metadata schemes as standards across the work, such as library science, education, archiving and so on.

5.1 *Proforma of metadata schemas*

We selected Heritage Monuments of Murshidabad and the objectives of this research work are to provide offline service to the heritage lover's community. We have preferreda self-modified metadata element set, which is a standard format. The schema is designed to support structure of the HTML formatted web pages ina suitable form. First of all, we had prepared a domain specific metadata schema with the unique space (MDA) by the help of Greenstone Editor for Metadata Sets. As per this domain, we had to choose 12 fields to complete our metadata collection. The names paces of our metadata are

Table 1: Shows the namespace of special metadata schema

Sl. No.	Namespace
01.	MDA.Title
02.	MDA.Creator
03.	MDA.Measurement
04.	MDA.Materials and Techniques
05.	MDA.Additional Physical Characteristics
06.	MDA.Style
07.	MDA.Date
08.	MDA.Current Location
09.	MDA.Description

6. Developing steps of open-source framework work for web-integrated cultural heritage information system

The paper discusses not only the methods and steps of the open sources framework for web integrated cultural heritage information system of Murshidabad but also the procedures or techniques regarding the collections and their organised process, analysis and evaluation of data or facts. This would be executing the research works. To design of this research work of web integrated heritage place in Murshidabad information services, described in two angels:

I) Designing a working model of heritage places information service for Murshidabad.

II)Making searchable and browse able heritage monument information products on CD-ROM.

6.1 *Digital library environment development*

- Building a domain-specific collection of digital materials relating to my research area of heritage monuments of Murshidabad (Seshaiah & Veeraanjaneyulu, 2009).
- Incorporation of self-modified format in the digital objects.
- Installation and configuration of Apache web server (ver. 2.2.14) in Linux (Ubuntu) platform.
- Installation of Java Runtime Environment (ver. 1.6.0_26) in Linux platform.
- Installation of GSDL (Ver. 2.84) Linux platform.

6.2 *Web access mechanism development*

In order to provide access to the digital collection in the distributed information environment, configure the system as a server (Apache) and digital collection in GSDL by modifying the server configuration (Mukhopadhyay, 2004).

6.3 *Organisation the digital collection through GSDL*

- Gather: Storage of cultural heritage monuments materials.
- Enrich: Encoding each Architect by the self-modified schema.
- Design: Controlled indexing of digital monuments.
- Format: Designing user interface for searching heritage monuments.
- Create: Creation of a collection of cultural heritage monuments.

6.4 *Offline Access Mechanism Developments*

The Primary goal of my research work is to create a digital archive that can be searched or accessed from any PC on any platform, and it can be compatible with CD-ROM.

6.5 *Apache web server Installation*

In this research work, Apache has been used as a web server because it is free and open-source software. Installing of a web server is necessary in order to deliver your web pages online. Because of its stability and flexibility, the Apache web server is used in many production scenarios.

6.6 *Installation of Java Runtime Environment*

It is advised you that disable your Internet firewall before Installing Java online. The following manner is followed for Installation of Java:

Step 1-Go to Java download page by Google search then click on Windows online,

Step 2- You see the download dialog box then click on download, after download you can run or save the downloaded file.

Step 3- To Installation you may choose file from your system and double click on save file.

The installation of the downloaded file begins when click on Install and accept the terms and conditions, at the end of the Installation you can click on finish.

6.7 *Installation of PERL (Programming Environment)*

Perl is a Programming language. The Perl language was originally developed by Larry Wall. The Greenstone Digital Library Software uses Perl programmes, Perl compiler should be loaded.

6.8 *Installation of GSDL*

In this article, the installation process of Windows version GSDL software from CD-ROM has been discussed. First of all you can place the CD-ROM into the drive then restart your system if the installation process does not begin immediately. Then click on setup.exe file to launch the software then click on next button after accepting the default each time to begin the software installation. After installation is finished, you may click on start button, open the programme menu and select Greenstone Digital Library Software. A dialogue box in the Greenstone Digital Library Software is located there, you can simply click on enter library, this may automatically start your internet browser and load the GSDL home page (Seshaiah & Veeraanjaneyulu, 2009) which should look something like the below:

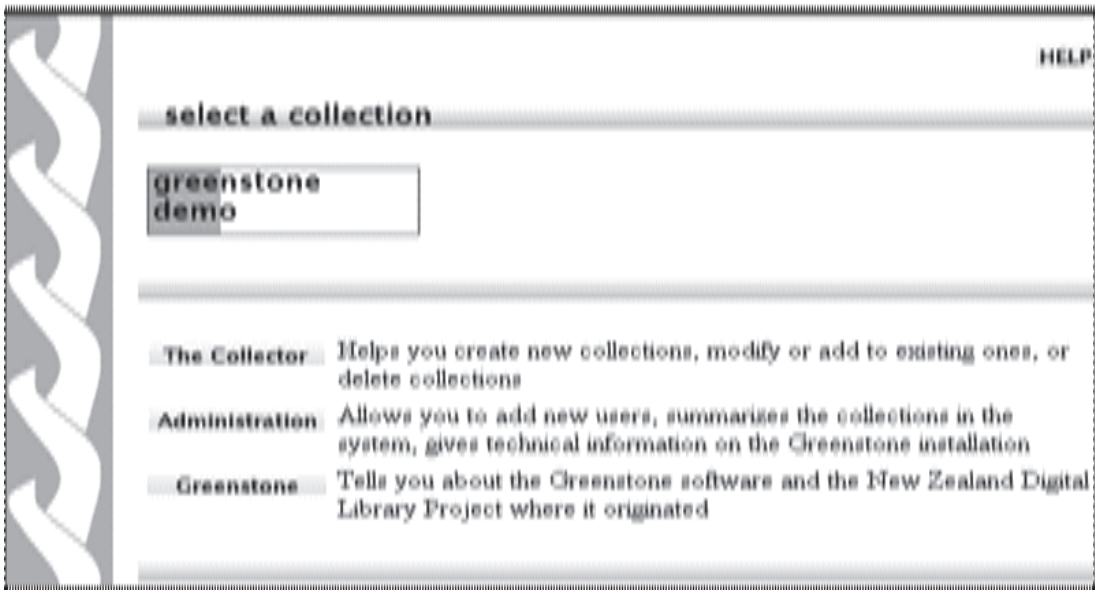

Figure 1: Greenstone digital library software homepage

Basically Local Library and Web Library are the two distinct programmes that are included on the CD-ROM. The URL http://localhost/gsdl/..... type into address bar of the web browser to use GSDL home page.

Process to creation of metadata schema in Greenstone Editor for Metadata Sets

- Open the Metadata set editor (GEMS). Then click on File and then click on New. Then put the metadata set title ("MDA"), metadata set name space ("MDA"), metadata set description ("Digital Library of Murshidabad") and click on ok.
- Right click on given metadata set title and click on "Add Element" and enter a name for the new element and so on.

6.9 *Collection building with GLI under Linux (Ubuntu)*

The GLI is an extremely user-friendly interface for creating digital library collections that is based on Java. There are four ways to create a digital library that the librarian can be used, this is Librarians, Assistant Librarians, Library specialists and users. The GLI facilitates basic six types of activities, such Gather, Enrich, Design, Create, Build and Preview. The collection, which may be consider as modules for making Greenstone collections (Mukhopadhyay, 2004). The Major tasks of GLI are:

- Gather: Storage of heritage monuments of Murshidabad.
- Enrich: Encoding each monument by self-modified metadata schema.
- Design: Controlled indexing of digital collections.
- Format: Designing user interface for searching and browsing heritage monuments.
- Create: Creation of the collection on cultural heritage monuments of Murshidabad.

- Convert: Digital library converted into a CD-ROM.

Gather: Greenstone Librarians Interface allows us to gather heritage collections of Murshidabad in different ways. These consist of i. Documents in GSDL collection, ii. Local file space, iii. Home folder (root), and iv. Download files. Here, I gather heritage monument collections from Local file space. So, the processes are Gather→Local file→Home→Desktop→ Related file drag.

Figure 2: Gathering procedure of documents

Enrich: To prepare metadata for every source of the document in the collections, enrich involves entering the source document data in the metadata fields and assigning metadata to each source of document.

Design: To setup of a collection is specified in terms of format, index, classifier, document plugin and display, among other things. It is able to automatically exact basic metadata like Time, Date, and so forth. It is necessary to exact explicit metadata using 'Classifiers' such as Subject, Author, Organisation, and so on. If we select design then different options appears, which provide different functions. These consist of(I) Document plugins, (ii) Search index, (iii) Partition indexes and (iv)Browsing classifiers.

- **Document Plugins:** Document plugins are necessary in order to convert the document format that Greenstone requires.

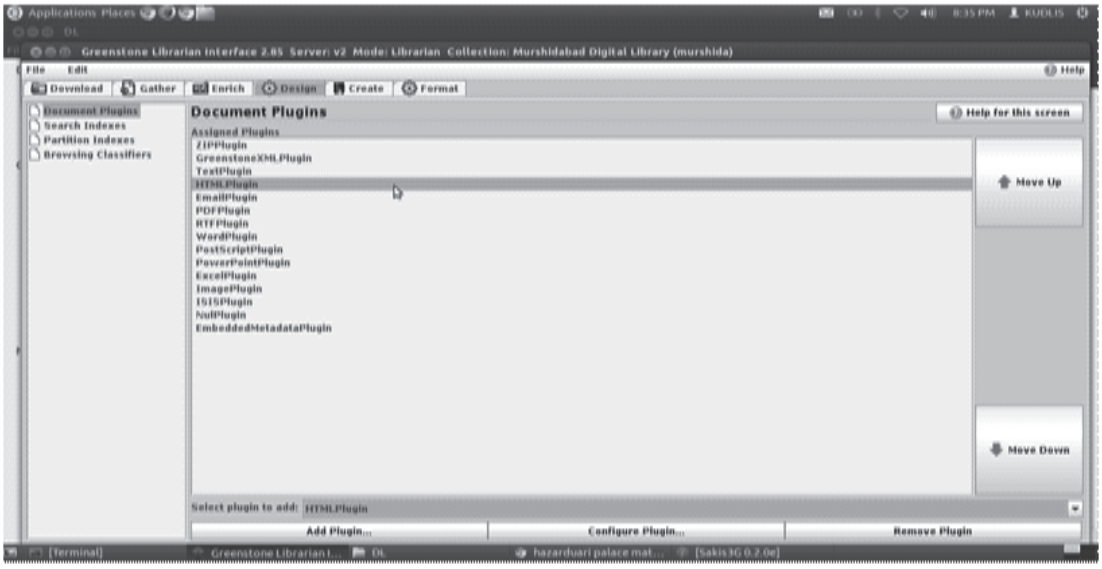

Figure 3: Document plugins

- **Search indexes:** The steps of search index are - Design then Search index → all files delete → New index → Full text and finally Add index.
- **Browsing Classifiers:** The process of Browsing classifiers are- Design then Browsing classifiers → all file remove → select classifiers to add → A to Z compact list → Add classifiers → Metadata then Sort → Bottom name and finally OK.

Format: The next step for collection building is format. When we select this option then here appear different portions, which provide different functions. These consist of (i) General, (ii) Search, (iii) Format Features, (iv) Translate Text, (v) Cross collection search, (vi) Collection specific macros and (vii) Depositor metadata. Here I discuss some of them

i. General: To give general information about the collection of digital content, we can select a photo that will appear as an icon of the collection in the home page.

ii. Search: The steps are - Format → Search → Index text → Full text

iii. **Format Feature:** The steps are - Format → Format features → Choose features → HTML format string

Create: Create is the initial step to starting the building operations, and build a collection is the collection of digital materials and in the compressed form of the source documents and coordinating all the activities done in the previous models into one place in the GSDL system. A connection to the user interface that allows one to view the current collection is called a preview.

7. Cultural heritage information system on CD- ROM

The method of creating cultural heritage information system of Murshidabad on CD-ROM is described in the previous chapter. Here, this paper describes the searchable access mechanism and how to write digital content into a CD-ROM.

Converting the collection into a CD-ROM

The primary goal of my research work is to create a digital archive that can be searched or accessed from any PC on any platform, and that is compatible with CD-ROMs. The following are the steps to convert the digital content into a CD-ROM:

Go to File then write CD-ROM → Check box to select your collection → write CD/DVD → Close.

The Greenstone Digital Library software facilitates both basic and advanced searches, as well as convenient browsing of digital content. The homepage of Information system is shown below:

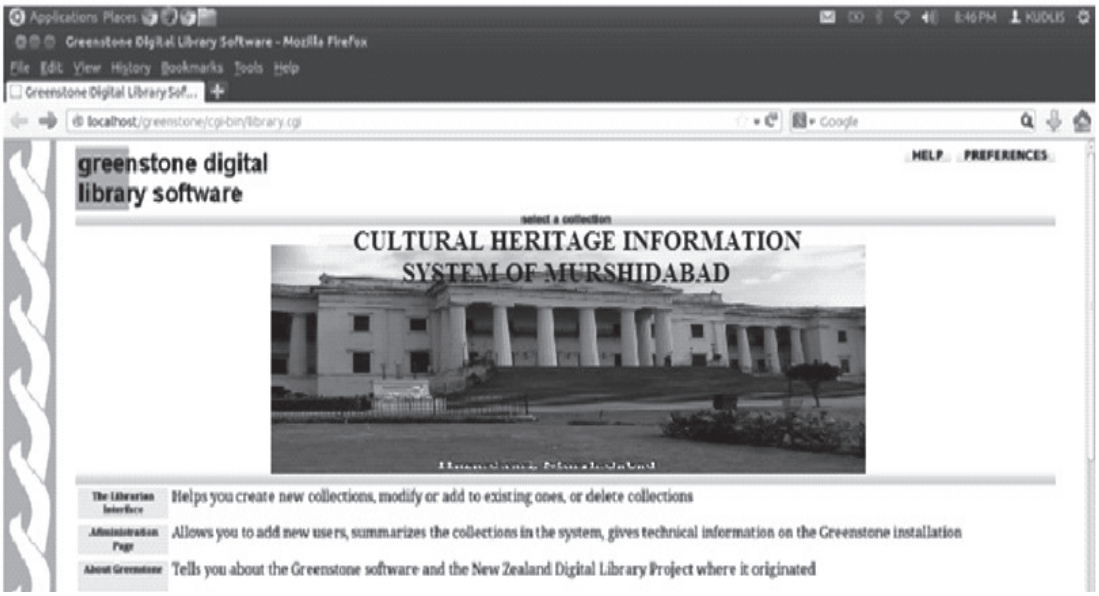

Figure 4: Homepage of information system

Users can choose some keywords and type in into the search box, so they can get needed information. Users can search the collection through the word 'about' and the search results relating to about are shown.

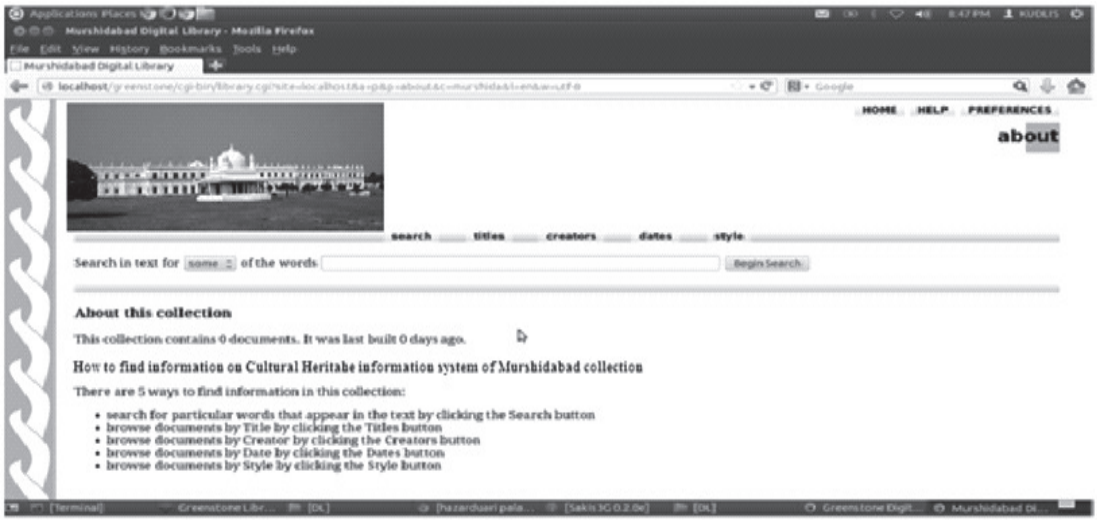

Figure 5: Main user interface

8. Conclusion

This research paper gives information regarding the heritage places of Murshidabad and how to build a digital archive. It also provides an overview of processes such as documentation, digitisation, metadata creation, software selection, customisation, testing, and promotion. The concept discussed in this article will help create a web-based information service for the cultural information system in Greenstone Digital Library Software. From this concept, one can expand the method of creating value-added e-information services that are available on CD-ROM as well as the Web. It is possible to implement the idea for all districts of West Bengal, and one can improve the same. This is the future scope of work. This work will also help researchers in the future build a digital archive for preserving, promoting, and sharing information. It will assist the specialists, tourists and heritage lovers by giving them the latest information about the different famous cultural heritages of Murshidabad.

References

Das, Shyamal (1996). *Murshudabad Parikrama.* Khagra: Murshidabad Sandesh.

Gupta, K. P. (2022). Feasibility study of preservation metadata in Greenstone Digital *Library Software. Library Waves,* 8(1), 36-45. Retrieved from https://www.librarywaves.com/index.php/lw/article/view/119/109

Jenifar, L, et al. (2024). A study on digital preservation methods for cultural heritage sites in India. *IPE Journal of Management,* 14(1), 163-170. Retrieved from https://www.researchgate.net/publication/378907574

Mukhopadhyay, P. S. (2004). Organisation and dissemination of digital objects through web and CD-ROM: a framework for Indian libraries. *Proceedings of the International Conference on Digital Libraries,* 470-478. Retrieved from https://www.researchgate.net/publication/263046446

Parvez, A., et al. (2012). Building digital collection with Greenstone: development and customization. *International Journal of*

Library Science, 5(1), 14-29. Retrieved from https://www.academia.edu/55375078

Reitz, J. M.(Ed.). (2013). *Online Dictionary of Library and Information Science.* Danbury: Western Connecticut State University.

Roy, Nikhilnath (2008). *Murshidabad Kahini.* Kolkata: Punthipatra.

Seshaiah, O. & Veeraanjaneyulu (2009). Developing an institutional repository using Greenstone Digital Library (GSDL) Software. In *7th International CALIBER: E-Content Management: Challenges and Strategies*. Puducherry: Pondicherry University, 161-166.

Sinha, R. & Murdia, M. (2018). Heritage festivals: an important factor to promote cultural tourism, a study of Murshidabad, West Bengal. *Annals of Art, Culture & Humanities,* 3(2), 09-19. Retrieved from https://www.researchgate.net/publication/326930561

Sonkar, S. K., et al., (2005). Application of Greenstone Digital Library (GSDL) Software in newspapers clippings. *DESIDOC Journal of Library & Information Technology,* 25(3), 9-17. Retrieved from https://publications.drdo.gov.in/ojs/index.php/djlit/article/view/3655