Digitisation in an Academic Library: A Success Story at Tata Institute of Social Sciences

Muttayya M. Koganuramath and Mallikarjun Angadi

Tata Institute of Social Sciences, Deonar, Mumbai-400 088
E-mail: koganuramath@rediffmail.com

ABSTRACT

Digital libraries were viewed as systems providing a community of users with coherent access to a large, organised repository of information and knowledge. The ability of the user to access, reorganise, and utilise this repository is enriched by the capabilities of digital technology. Digitisation has become important in worldwide efforts to preserve, manage, and provide access to scholarly information. However, digitisation efforts require significant financial investments in infrastructure and technology as well as in human resources for project management. This paper discusses the practical experience of digitisation of two major projects carried out in the library. First one is that of digitisation of 55 volumes (1952-2006) of *Sociological Bulletin* published by the Indian Sociological Society, New Delhi and digitisation of 67 volumes (1940-2006) of *Indian Journal of Social Work* published by Tata Institute of Social Sciences, Mumbai. The paper highlights the objectives of the project proposals, justification and anticipated impact, plan for the digitisation, process methodologies for digitisation, volume of work, etc. The paper also describes various steps involved in the digitisation process, OCR software used and its advantages over other OCRs, search interface used, browsing and searching facilities, navigation facilities provided in the CD, etc. Some of the past and present digitisation projects that have been carried out in the library are also mentioned.

Keywords: Digitisation, developing world, Tata Institute of Social Sciences, *Indian Journal of Social Work*, S*ociological Bulletin*

1. INTRODUCTION

The information contained in traditional print materials like books, journals, reports, published works, minutes of the important meetings, manuscripts, cannot be preserved forever for a number of reasons. As years pass by, the information contained in them gets faded out, the medium becomes brittle and finally becomes unusable. Unless we have alternatives arrangement for recapturing and reproducing it another format important will be lost forever. Fortunately, technological advances have provided us with suitable alternatives for preserving such valuable information

Information technology has brought tremendous changes in the way of the life of the human being. Publication industry is no exception to it. To have wider access and long-term preservation policy of scholarly human knowledge many of the professional organisations, and publication houses are moving towards the electronic publication of their print resources.

Tata Institute of Social Sciences, being the premier institute in the field of social sciences in India, fulfils the mandate of producing knowledge and its dissemination in the field of social work and social development. Sir Dorabji Tata Memorial Library (SDTM Library), located in the main campus, is a unique Knowledge Center, which offers essential and specialised information resources and services to meet the growing information needs by developing user-based resources, organising information resources, providing human and technologically moderated access to information, and aiding and facilitating users to identify, locate, obtain and evaluate information.

In 2004, SDTM library set up a digital library (DL) with the financial assistance from Tata Education Trust, Mumbai to provide seamless access to social science information to users in a networked environment by creating digital contents. DL at the Sir Dorabji Tata Memorial Library is one of the finest on social sciences in the country as well as the South-Asian region. The

establishment of this digital repository will also keep the library on par with major social science libraries in the world. The DL facilities include state-of-the-art technologies like high-end desktop computers, digital cameras, OCR software, scanners, high speed internet connectivity, access to e-resources, etc. The library is carrying out many digitisation projects, both in-house and outside projects. The following are some of the digitisation activities that are being carried out in the library.

2. PROJECT-1: DIGITISATION OF *INDIAN JOURNAL OF SOCIAL WORK*

Institute of Social Sciences, a premier institute in the field of social sciences, fulfils the mandate of producing knowledge and its dissemination in the field of social work and social development through several publications. One of the major achievements of the institute is uninterrupted publication of its refereed journal *Indian Journal of Social Work* (*IJSW*) since 1940. The IJSW is published every January, April, July, and October, and has entered its 68th year of uninterrupted publishing, a record in journal publishing. The major emphasis of the first issue was on two aspects of child welfare—children in industry and juvenile delinquency. The *IJSW* is widely used by teachers, researchers, professionals in the field, and the students, and has played a facilitative role in social work education across the country. Research study findings and theoretical conclusions of research are regularly published in this journal.

2.1 Need for Digitisation

The need for digitisation of back volumes of *IJSW* stemmed from the following objectives:

- To develop a world-class web ready searchable database of *IJSW* contents.

- To expose to the world, the untapped and under utilised contents of *IJSW*.

- To facilitate quick retrieval of the contents of *IJSW* with user friendliness.

- To permit other world scholars to use the resources for collaborative practical experimentation.

2.2 Targets

- All universities and colleges with a course in social work.

- All NGOs and non-profit organisations involved in social work research education and practices.

- All social work institutes across the world.

2.3 Justification and Anticipated Impact

Presently, the reach of *IJSW* is limited to few who have access to its hard copy. Surely there are thousands of researchers and social scientists who will find this source useful. Once the contents of *IJSW* are digitised, it can be accessed from all across the world. Digitisation of *IJSW* will have significant impact because:

- It will provide a structured way to search *IJSW* contents using keywords, author, title, etc.

- *IJSW* will have seamless access across the world. Its reach will be far wider and faster.

- The contents of *IJSW* can be preserved forever.

2.4 Volume of Work

IJSW has over 67 back volumes. Each volume has on an average 700 pages totalling to 46,900 pages (approx). All the bound volumes are available with library. Old volumes have become yellowish because of wear and tear and natural aging process. These had to be photocopied and then scanned.

2.5 Funding of the Project

The project was funded by the institute under the BRS grants. A library professional with computer skills was hired to carry out the project for 18 months.

2.6 Stages of Digitisation

Scanning: ABBY Fine Reader ver.8.0 software was used for scanning. The following were the steps involved while scanning:

- Scanning of the printed documents using HP flatbed scanner.

- Running OCR for the scanned documents.

- Editing-removing the unnecessary characters, redundant data.

- Spell-check using ABBYY fine reader.

- Proofreading the documents page by page and line by line for any missed character and removal of redundant characters.

- Save the final documents in RTF (Rich Text Format), PDF (Portable Document Format), and batch mode files–as a source file

GSDL Search Interface: Greenstone Digital Library Software (GSDL) was used to provide searching facility.

Publishing on the CD ROM: Entire GSDL database along with the software was published on CD using the GSDL publishing facility.

2.7 Browsing *Indian Journal of Social Work* CD

Browsing from Main Page: The *IJSW* CD opens in browser (tested for all popular browsers compatibility) and contains information about guidelines for the contributors, subscription/renewal, contact details, brief about the journal, etc. On the first page links have been provided to browse and search the journal (Fig. 1).

Browsing facility provides navigation from volume to volume and issue to issue (Fig. 2). Links to full text has been provided for all articles and book reviews from the content page.

Browsing from GSDL: One can also browse the entire content of the journal from GSDL.. Following are the two ways of browsing through GSDL:

- Browse articles by titles A-Z by clicking the title button (Fig. 3).

- Browse articles by author A-Z by clicking the author button (Fig. 4).

2.8 Searching *Indian Journal of Social Work* on CD

To search *IJSW* journal on CD, a search interface has been provided using GSDL ver. 2.72 that works only on Windows OS (Windows Vista/XP/2000/NT/98). Greenstone is a suite of software for building and distributing digital library collections. It provides a new way of organising information and publishing on the Internet or on CD-ROM. The steps to search *IJSW* are:

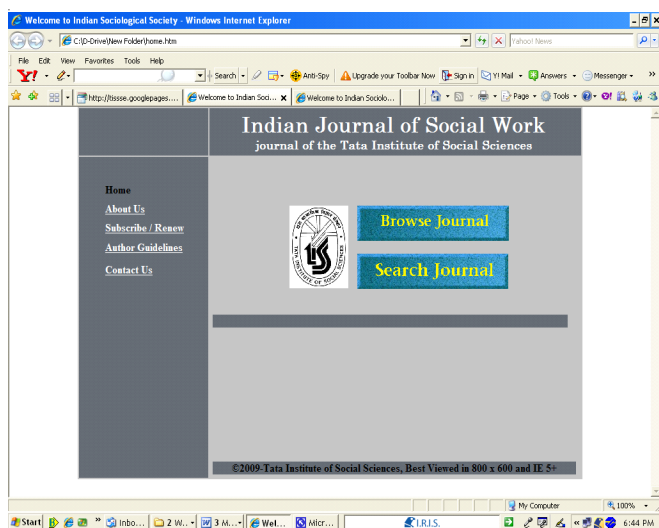- Go to My Compute' and open the CD by right clicking the CD Drive and select Open.

- Double click the application Server which opens GSDL server in a separate pop-up window.

- Click Restart Library button to start a default browser. (Don't close Greenstone Digital Library version 2.72 pop-up window until the search is complete).

- Click Sociological Bulletin icon to open Search Window.

A step-by-step instruction manual with screen shots has been provided in the CD on how to open GSDL Search Window. There are three ways to search the journal's contents:

- Search the journal for a particular word(s) in the text.

- Search the journal author-wise.

- Search the journal title-wise.

3. PROJECT 2: DIGITISATION OF *SOCIOLOGICAL BULLETIN*

3.1 About *Sociological Bulletin*

Sociological Bulletin, started in the year 1952, is the official journal of the Indian Sociological Society,. It is now published thrice a year: January-April, May-August, and September-December. It carries reports on research conducted in India and abroad by professional sociologists and anthropologists. Over the last five decades, the *Bulletin* has been advancing the frontiers of knowledge about Indian society, its social institutions and culture, and their structure and dynamics. It has been contributing to our understanding of other societies as well. The *Bulletin* meets the needs of higher education in India and provides for specialists and non-specialists alike. It also serves as an important source material for students and researchers. Digitisation of the entire
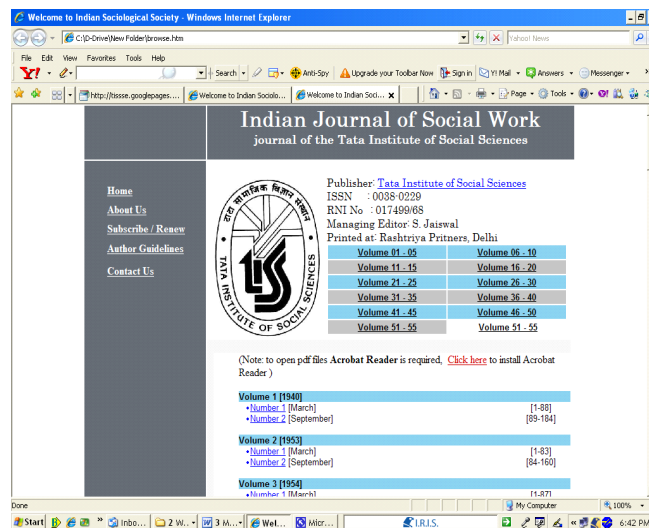

Figure 1. Main page.

Figure 2. Browsing page.

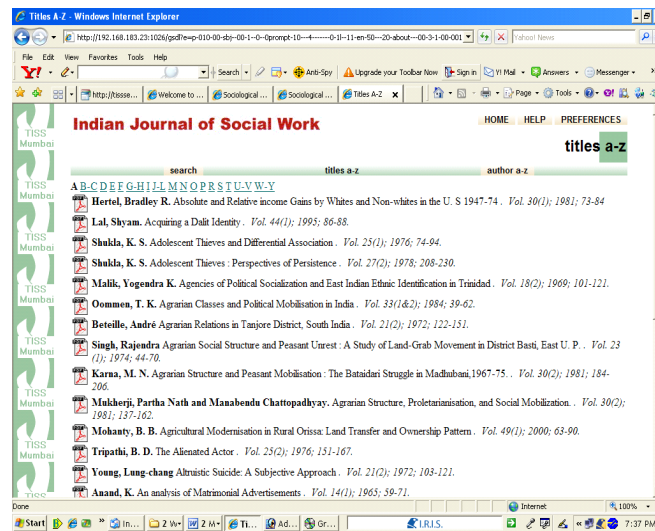**Figure 3. Browse by title.**

Figure 4. Browse by author.

collection of *Sociological Bulletin* (Vol. 1-55) has been completed by the SDTM.

3.2 Objectives of the Digitisation

The need for digitisation of back volumes of *Sociological Bulletin* stemed from the following objectives:

� To expose to the world, the untapped and under utilised contents of *Sociological Bulletin*.

✿ Quick retrieval of the contents of *Sociological Bulletin* user friendly.

✿ To allow other world scholars to use the resources for collaborative practical experimentation.

3.3 Justification and Anticipated Impact

The digitisation achieves the preservation of important sources forever, searching for information becomes easy, accurate and quick, and digitised information can be updated easily. Digitisation of *Sociological Bulletin* will have significant impact because:

✿ It will provide a structured way to browse its contents.

✿ It will have seamless access around the world. Its reach will be far wider and faster.

3.4 Volume of Work

Sociological Bulletin has 55 back volumes. Around 15,000 pages have already been digitised. The entire digitisation work was completed in 10 months.

3.5 Funding of the Project

The project was funded by the Indian Sociological Society, New Delhi. A library professional with computer skills was hired to carry out the project for 10 months.

3.6 Browsing *Sociological Bulletin* CD

The *Sociological Bulletin* CD (tested for all popular browsers compatibility) contains information about guidelines for the contributors, subscription/renewal, contact details, brief about the journal, etc. On the first page links have been provided to browse and search the journal in alphabetically arranged author index and title index (Fig. 5). Browsing facility provides navigation from volume to volume and issue to issue (Fig. 6). Links to full text has been provided for all articles and book reviews from the contents page. The journal can also be browsed from GSDL application.

3.7 Searching *Sociological Bulletin* CD

To search *Sociological Bulletin* on CD, a search interface has been provided using GSDL software ver. 2.72, which works only on Windows OS (Windows Vista/ XP/2000/NT/98). GSDL provides a new way of organising information and publishing it on the Internet or on CD-ROM. The steps to search the journal are:

✿ Go to My Computer and open the CD by right clicking the CD Drive and select Open.

✿ Double click the application Server which opens GSDL server in a separate pop-up window.

✿ Click Restart Library button to start a default browser (do not close Greenstone Digital Library version 2.72' pop-up window until the search is complete).

✿ Click *Sociological Bulletin* icon to open Search Window.

A step-by-step instruction manual with screen shots has been provided in the CD on how to open GSDL Search Window. The journal can be searched by a particular word(s), author-wise and title-wise.
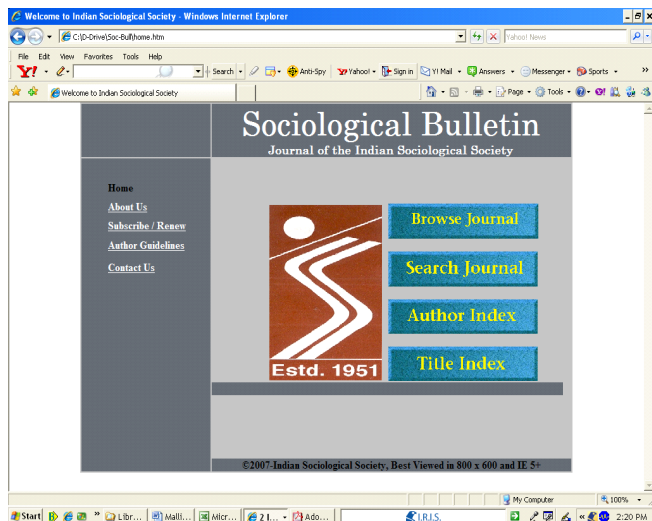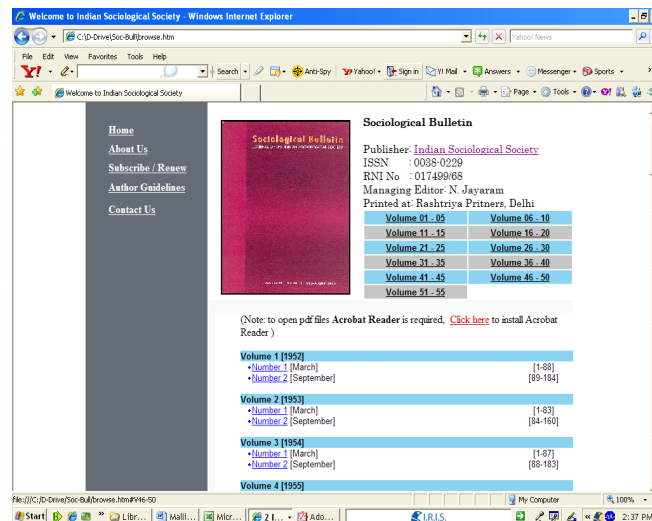
Figure 5. Main page.

Figure 6. Browsing page.

4. OTHER DIGITISATION PROJECTS

Parsee Literature Digitisation: 150 years old Parsee literature of 10,000 pages has been digitised and made available for the reference in CD-ROM format. The literature was in Gujarati script. Entire 32 volumes have been accommodated in three CDs.

Digitization of Theses and Dissertations: Apart from the general collections, the library also holds a good number of theses and dissertations submitted to TISS covering diverse subjects of social sciences. It is this rich valuable collection that has made the library unique in its own way. TISS library has approximately 2500 Master of Arts dissertations and around 500 doctoral theses.

Project Reports/Monographs/Technical Reports brought out by Faculty: Digitisation of all the project reports, monographs, course material brought out by the faculty members is under progress. There are approx. 100 monographs with 15,000 pages that require digitisation.

Books: The Publication Unit at TISS has published over 100 books since 1940. All theses can be digitised to convert them into e-books. Each book is of an average 300 pages.

Proceedings Volumes: Many national and international conferences/seminars/workshops are being conducted by departments, research units and library every year. Proceedings volumes of these conferences/seminars/ workshops are a good source of information. Hence the library has taken up the job of digitising all the proceedings of the institute.

5. CONCLUSION

Digitisation of contents, preservation, and digital asset management has become the driving process. This case study describes very small digitisation projects but the lessons learnt can easily be expanded to larger scale projects. The authors plan to continue to work on bringing more and more social sciences journals in the digital format and look forward to streamline the research value in the field of social science in India.

ACKNOWLEDGEMENTS

REFERENCES

1. Kuny, T. An introduction to digitisation technologies and issues. *Network Notes,* 14. http://www.nlc-bnc.ca/pubs/netnotes/notes14.htm (Accessed on 25 December 2008).

2. Berger, Marilyn. Digitisation for preservation and access: A case study. *Lib. Hi Tech.,* 1999, **17**(2), 146-151.

3. Whitt, Jennifer. & Hastings, S.K. The best of both worlds: A digital preservation project. *Lib. Hi Tech. News*, 2005, **10**, 13-16.

4. Eden, Brad. Getting started with library digitisation projects: Funding your first (and subsequent) digital projects. *The Bottom Line Mag. Lib. Fin.,* 2001, **14**(2), 53-55.

5. De Ridder, Jody L. Choosing Software for a digital library. *Lib. Hi Tech. News,* **9**, 19-21.

6. Witten, Ian H. & Bainbridge, David. Creating digital library collections with Greenstone. *Lib. Hi Tech.*, 2005, **23**(4), 541-60.

7. Krishnamurthy, M. Open access, open source and digital libraries: A current trend in university libraries around the world. *Program: Electro. Lib. Inf. Sys.,* 2008, **42**(1), 48-55.

8. Lopatin, Laurie. Library digitisation projects, issue and guidelines: A survey of the literature. *Lib. Hi Tech.,* 2006*,* **24**(2), 273-89.

9. Graham, Theresa. Electronic access to and the preservation of heritage materials. *The Electro. Lib.*, 2003, **21**(3), 223-26.

10. Witten, Ian. H.; Bainbridge, David & Boddie, Stefan. F Greenstone: Open source digital library software with end-user collection building. *Online Inf. Rev.*, 2001, **25** (5), 288-98.

11. Rowley, James Daniel & Sherman, Herbert. The special challenges of academic leadership. *Management Decision*, 2003, **41**(10), 1058-63.

12. Schaffler, Hildegard. How to organise the digital library: Reengineering and change management in the Bayerische staatsbibliothek, Munich. *Lib. Hi Tech*., 2004, **22**(4), 340-46.