AUTOMATION OF INDIAN LANGUAGES THROUGH GIST: A STUDY OF PANJAB UNIVERSITY LIBRARY

VINOD KUMAR GARG   **JIVESH BANSAL**

Abstract

Indian languages literature other than English is very common in almost all the libraries in India. The present study explains the experiences of Panjab University Library in solving their problems of Indian languages literature for entering into the computer database with Techlib Plus software. Describes various features of Techlib plus software along with its implementation programme in the University Library under study. Discusses about GIST technology and GIST card developed by CDAC, Pune , India.

Keywords : Techlib Plus/ GIST/ Indian Languages Script

1. Panjab University Library: A Profile

The Panjab University Library System consists of Main Library along with 55 small departmental libraries including two major institutes; University Institute of Engineering and Technology (UIET) and Dr. HSJ Institute of Dental Sciences. The library system has approximately 6.4 lac collection including books, bound journals and other non book material. The library is fully computerized with an integrated library management system connected to the campus network on fiber optics backbone. The library has latest reference tools along with CD-ROM and online databases. Online Public Access Catalogue can be accessed through any terminal on the campus network. The library has recently launched its Web-OPAC on library website www.library.puchd.ac.in. The library subscribe to nearly 600 current periodicals in print form. Its holding of back volumes of periodicals goes back to 19th century. Near about 4000 full-text scholarly journals are available online at campus network through UGC-INFONET and INDEST Consortium.

2. Techlib Plus Software

Panjab University Library started its computerization programme in the year 1995 by the funds provided by INFLIBNET, Ahmedabad. After lot of discussions and meetings, finally Basis Plus/Techlib Plus software marketed by National Informatics Centre (NIC), New Delhi was chosen in the year 1996 due to initial cost, recurring expenses, provision for customization as well as being an open ended software.

Techlib Plus software is an integrated library management software developed by Information Dimension Incorporation (IDI), Dublin, Ohio, USA. The underlying relational database management system, Basis Plus is also being designed and developed by IDI, USA for specifically text retrieval applications like library application. Although Basis Plus/Techlib Plus was available under wide range of operating systems, but in India Basis Plus version L1G was available under Unixware operating system through NIC, New Delhi.

3.   GIST Technology and GIST Card

GIST technology developed by CDAC becomes standards for Indian language applications in information technology. The initial development took place in IIT, Kanpur in 1983 in the form of DOE (Department of Electronics) sponsored project for development of an integrated Devanagari Terminal. CDAC developed GIST-9001 Application Specific Integrated Circuit (ASIC) and transferred the rights to 25 companies for manufacturing of GIST cards and terminals. GIST–9001 chip provides a high resolution graphic screen for displaying very complex fonts. The powerful 68008 microprocessor provide the speed required for INTELLIGENT algorithms for perfect composition of scripts. The underlying architecture of all the GIST products uses the Indian Script Code for Information Interchange (ISCII), their representation on screen and printer (Intelligence Based Script Font Code), and INSCRIPT keyboard layout to cater to all Indian scripts in a uniform way. INSCRIPT keyboard facilitates instantaneous transliteration between different Indian scripts.

GIST Card is a PC adds on card. It allows the use of Indian and some foreign language scripts, with English in all existing text based application packages like dbase, FoxPro, Lotus 1-2-3, Qbasic and compilers like C, C++. The GIST Card can be used for multilingual database applications on MS-DOS. With the application of this card, the database can be entered, processed, sorted and displayed in any Indian Language with English. The Gist card is available in WordStar Compatible, customized multilingual processor called 'Apex Language Processor (ALP)' for GIST card. GIST is compatible with almost all the available dot-matrix printers, which have a graphics printing mode. GIST supports the following scripts: Devanagari (for Hindi, Sanskrit and Nepali), Bengali, Assamese, Oriya, Gurumukhi, Gujrati, Kannada, Telugu, Tamil and Malayalam. In other Indian scripts belonging to the Perso-Arabic family the following scripts are supported: Arabic, Persian, Urdu, Kashmiri and Sindhi. GIST also supports some foreign languages like Thai, Tibetan, Bhutanese, and Sinhalese.

4.   Strategies for Retro-conversion in University Library

Panjab University Library started using Techlib Plus software for library automation from January 1,1997. Initially English language books were taken for retro-conversion. University Library being an old and rich in collection, it was a gigantic job to convert this huge collection in electronic format. As a strategic planning, it was decided to start retro-conversion from the collection which is in circulation through shelf card. A sign was put on the book as well as on the shelf card which have been entered in computer database.

After some time when almost all the collection which was in circulation was converted into electronic format. Then retro-conversion from shelf list arranged in classified order in Technical Section (of course leaving those shelf list cards which were already entered into the database) was undertaken.

After entering nearly 80% collection of English language, we started to look for solutions for other Indian languages in general and Hindi and Punjabi in particular. After lot of survey and discussion with different solution providers for Indian language retro-conversion Graphic and Intelligence based Script Technology (GIST) solution developed by Centre for Development of Advance Computing (CDAC), Pune was chosen to implement in the library. There were two options under existing hardware/software infrastructure, either a GIST Terminal or GIST card for PC. As we already have PC's to enter English language data, the GIST card was chosen to install in one of the PC. In the beginning, GIST card was installed in the Acquisition Section of the library for entering new books in Hindi language (Devanagari Script) for placing orders to vendors. With the successful trial, second GIST card was installed in Technical Section for retro-conversion of Hindi language books. As on date approximately 38,000 collection of Hindi language are entered in database through GIST card. In order to utilize this large database, another GIST card was installed in one of the OPAC terminals. Now library users frequently search for Hindi language books through OPAC instead of printed catalogue cards. Being connected through campus wide network OPAC can be accessed from any system in the whole campus, but to use this OPAC for Indian languages collection GIST card has to be installed. To take advantage of this facility, 'Hindi Department' of the University has also installed GIST card and now they are searching OPAC from their department itself for all the Hindi language collection along with English language collection.

Recently, the library has started retro-conversion for Punjabi language collection through same GIST Card, next rich collection after Hindi in University Library. Although entry of new books of Punjabi language were started at the very beginning of GIST card in the library.In new electronic environment, the library is still maintaining its printed catalogue for old collection as well as newly added collection along with OPAC. For new books catalogue cards (Main card, shelf card, added cards etc.) are generated through Techlib Plus software for English language. The Techlib plus being a text based application do not support printing on laser printer. The files generated through Techlib plus software is transferred through FTP in Windows based environment. Then after editing and setting work in MS-Word, finally prints are taken through laser printer. The same process when done for catalogue generation of Indian languages, the MS-WORD did not support Indian languages script entered through GIST. In order to retain the Indian language scripts in catalogue cards, a 'MS-WORD' like application 'ISM Office' developed by same agency CDAC is now being used by the library.

5.   Analysis and Status of Computer Database

Panjab University library has major collection in English language. Apart from this, library is rich in Hindi, Punjabi, Sanskrit and Urdu language collections which are taught under separate well established departments. Other than this, some minor collection is in Bengali, Marathi, Gujarati, Tamil, etc.

The language wise approximate distribution of the total library collection along with total records entered into the computer database of Techlib plus software is shown in the table given below:

Language	Total Records	Records in the Computer Database	Percentage wise in Computer Database
English	5,50,000	5,21,000	94.72%
Hindi	45,000	38,000	84.44%
Punjabi	22,000	2,000	9.09%
Sanskrit	11,000	-	-
Urdu	4,000	-	-
Bengali	1,000	-	-
Marathi	400	-	-
Gujarati	200	-	-
Tamil	200	-	-
Others	4,000	-	-
Total	**6,37,800**	**561,000**	**87.95%**

The table shows that the approximately 86% of the total collection is in English language and 95% of this total English language collection is entered in the computer database. The second major language Hindi forms approx. 7% of the total collection and 85% of this are entered into the computer database. The Punjabi collection forms approximately 3.5% of the total collection. The data entry work of Punjabi language collection is recently started by the University Library and approximately 9% work has been done. Another collections which have significant existence in the library are Sanskrit and Urdu. The retro-conversion of these languages will be taken up subsequently. Other Indian languages collection is very minor in Panjab University. In totality nearly 88% of the total collection has been entered into the computer database.
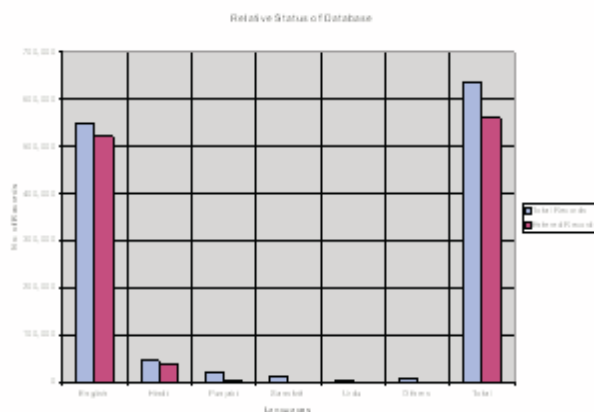
The above graph shows the relative status with regard to total records in the library over the records entered into the computer database. The last column in the graph shows the position of total collection into the library.

6. Conclusion

Libraries using text based library management software can implement GIST technology for Indian languages solution easily and successfully.

References

1. Apex Language Processor: user's manual. Pune: Centre for Development of Advanced Computing, 1995.

2. Basis Plus Training manuals. Ohio: Information Dimensions Inc., 1994.

3. Centre for Development of Advanced Computing. http://www.cdac.in/ [Accessed on 30.12.2006]

4. GIST Card user's manual. 2nd ed. Pune: Centre for Development of Advanced Computing, 1996.

5. Techlib Plus Training manuals. Ohio: Information Dimensions Inc., 1994.

BIOGRAPHY OF AUTHORS

Mr. Vinod Kumar Garg works as System Manager at Panjab University Library, Chandigarh - 160014. He holds MSc, PGDCSA, M.Tech. His areas of interest are Database Management Systems, System Administration, Open Source Software, Automation of Libraries.

Mr. Jivesh Bansal works as Assistant Librarian with Panjab University Library Chandigarh -160014. He holds MLIS, PGDCA. His areas of interest are Information and communication Technology (ICT) and its applications in libraries.

Email : jiveshbansal@gmail.com