

## Selecting Software Technology for Multilingual System and Issues Involved there after: A Strategic Plan

Gajula Venkata Nagaiah\*, Biswajit Sinha\*\*

### Abstract

In India, all apex literary libraries are gradually adopting to the changes that are taking place in the information world. This paper focuses primarily on those sectors that are nurturing multilingual resources, shows how changes in technology are impacting such libraries. Subsequently this paper describes the various steps in involved in selecting software/hardware technology for multi-lingual system, offering a broad strategic plan to go for automation, thus highlighting priorities of hardware and software selections. Also it takes it to considerations the integrated system and varied modules to replace all traditional library functions, at the same time capable of connecting to the world resources. Apart from identify various steps to embrace automation, makes an attempt to show how multi-lingual approach is not possible with earlier Indian script codes for information interchange. Eventually Unicode offers an excellent opportunity to Indian languages to store and publish the information in any Indian languages. Lastly it furnishes some softwares having Unicode support that helps in building and distributing multilingual library resources.

The intention of writing this article is to offer, an overview of how changes in technology are impacting today's libraries working in multilingual milieu and its user's expectation and discusses at length about selecting software technology for multilingual system, analyzing a strategic plan.

As we are all aware that The Sahitya Akademi Library which functions under aegis of Sahitya Akademi, the National Akademi of letters, occupies an important place not only in library map of the capital, truly this library is the premier library of letters in the cultural map of the Nation. It has an enriched collection of 24 languages in literature, literary theory and criticism, philosophy, history and cultural studies, which enjoys an active and appreciative readership. Apart from this apex body, other reputed multilingual libraries that deserve mention here are National Library, Tulsi Sadan Library, Delhi Public Library, and *Central Institute of Indian Languages* etc.

In the post independence face, particularly the last two decades we have witnessed that, with the acceptance of telecommunication systems and the strengthening of the computer technology infrastructure which have been key factor in revitalization of India's Science and technology, defense, world affairs, economy, art and culture and special research organizations and libraries bodings. The conventional thinking of a library as a store of information held locally has almost eradicated as the library of today is becoming a gateway to information resource worldwide. In India all prominent cultural and literary libraries are gradually adopting to this change with the advent of multilingual, multimedia concepts, networking and the information super highways such rapid changes in the world has placed the onus on India's library and information professionals to find ways of making the multilingual literary heritages in their custody, available electronically.

Recently the Sahitya Akademi selected a software for its multilingual library and gradually moving towards absolute library automation. This article therefore tries to highlight how to create a basic technology plan and develop a library profile in preparation of library automation. The paper also explains how such libraries access and identify

**Author's Affiliation:** \*Sr. Lib. & Info. Asst., \*\*Sr. Lib. & Info. Asst., Sahitya Akademi Library, Rabindra Bhawan, 35, Ferozeshah Road, New Delhi-110001.

**Reprint's request:** Gajula Venkata Nagaiah, Sr. Lib. & Info. Asst., Sahitya Akademi Library, Rabindra Bhawan, 35, Ferozeshah Road, New Delhi-110001.

E-mail:- venkatanagaiah@gmail.com

(Received on 21.05.2009, accepted on 23.04.2012)

institutional needs and priorities and then go about writing a technology plan. Now at the outset question arises as how to select and implement a new information system, which begins with system implementation checklist. In such situation, to commence library and information professionals first look for automation options. Specifications of software selection and subsequently the process of selecting a system, from preparing the invitation to the tenders through evaluating supplier responses to making the final decision. Pre-selection matters are negotiating a contract with the supplier of choice, testing and maintaining the system, training, and the special attention is paid to the multilingual aspect of the library system requirements and specifications for such a libraries.

#### *Library Automation*

Generally library automation involves, integrated systems on which the traditional library functions of Circulation, Cataloguing, the On-line Public Access Catalogue, Acquisitions and Serials control are computerized using the library's database as the foundation. While 'planning for automation' can still be defined as planning for integrated systems that 'Computerized a multiplicity of library functions using a common database'. However as the world has shrunk and rapid technological change, has forced for comprehensive re-examination of what automating the library really means. Subsequently, library automation *per takes* the following:

- access by users to library databases from home or office, with direct downloading of information and text on demand
- full-text storage of documents, complete with full-text keyword searching and on-demand printing
- vastly expanded storage of indexes
- storage of pictorial and graphic material
- The availability of 'intelligent systems' providing transparent, one-step searching and access to various library in-house and remote databases.

One should also take into consideration that today's integrated system should not only must provide module automating the traditional library functions but also must be capable of connecting through the local system into systems of other suppliers, databases – bibliographic and full content, online and CD-ROM and the internet.

But in the process of converting traditional library to automated one in a multi-lingual setup, some additional issues come up like software should be able to process multilingual script materials; the specific MARC code have to be followed for all language etc. thus the library management software should have MARC format / support.

**MARC 21:** MARC short for Machine Readable Cataloguing, constitutes a group of communications formats that conforms to the ISO 2709/NISO

Z 39.50 standards. It is the vehicle for converting bibliographic paper files to machine readable ones.

The important segment of a MARC bibliographic record is the 008 field, fixed-length data elements, also known as 'fixed field codes'. It provides such information as whether the item being described in the record is a monograph or serial, in which country it was published, etc. While cataloguing, the language code has to be selected from 008 tag subfield along with other delimiters for entering the data.

008	090701e2001 ii 000 f tel d
-----	----------------------------

Source: Sahitya Akademi Library Indian language MARC sheet

*Unicode: Encoding of Indian Language characters*

Indian Script Code for Information Interchange, acronym of ISCII refers to a standard devised for Indian character representation by the Bureau of Indian Standard (BIS) which is 8 – bit code. It touches

**Indian language codes as per ISO standard.**

*Table language code:*

ISO 639-2 Code	Name of the language
asm	Assamese
ben	Bengali
doi	Dogri
eng	English
guj	Gujarati
hin	Hindi
kan	Kannada
kas	Kashmiri
kok	Konkani
mai	Maithili
mal	Malayalam
mni	Manipuri
mar	Marathi
nep	Nepali
ori	Oriya
pan	Punjabi
raj	Rajasthani
san	Sanskrit
snd	Sindhi
tam	Tamil
tel	Telugu
urd	Urdu

Source: <http://library.igcar.gov.in/readit-2005>

upon 10 Indic script derived out of ancient Brahmi Script.

ISCII uses extended ASCII (American Standard Code for Information and Interchange) that means it uses 128 characters positions for character representation of Indic scripts.

Multilingual approach is not possible with ASCII. ASCII was built around the Latin alphabet. As such, they are restricted in their abilities to provide data representation for the

non-Latin alphabets used by the majority of world's population.

As all countries began using computers, each was devising codes that would most effectively represent their native languages. None of these were necessarily compatible with any others, paving yet another barrier in the way of the emerging global economy. In 1991, before things got too far out of hand, a consortium of industry and public leader was formed to establish a new international information exchange code called Unicode.

As the basic code of Unicode is 16 bit, it has the capacity to encode the majority of characters used in every language of the world. It provides code points for more than 6500 characters. The data can be entered, stored, and indexed in the original language itself. Unicode has therefore, given wonderful opportunity to the Indian languages, as we can now create, store and publish the information in our own mother tongue.

*Roman Transliteration*

Transliteration between Indian languages is simple, unambiguous and phonetically similar. However, there is need for a Scheme for transliteration from Indian Scripts to Roman Script. In order to maintain the uniformity, the books are entered in the Accession register in Roman script only. Traditionally, the transliteration was phonetic based and had many diacritical marks in order to represent letters in unique way. Later came the computer era wherein it was still more difficult to manage the diacritical marks. The phonetic symbols were created as symbols, saved as pictures and inserted within the text wherever it was required, because those representations were not in the computer keyboard.

Let's see as to what kind of development has taken place in our country in the direction of storage of information, say in Hindi language which is spoken by about 30% of the population. Currently propriety Hindi fonts of different standards are using which are not compatible with each other that's why causing problems in information exchange. In

order to facilitate free exchange of information, Department of Information Technology, Government of India has accepted Unicode encoding for fonts as Indian standard in this regard. By using Unicode compliant fonts for Hindi, the problem relating to the exchange of information with Hindi content is completely solved. Majority of users are now using Windows 2000 operating systems or later version which are fully capable of working in Hindi by using any Unicode compliant font. Therefore, without installing any additional software these computers can be straightway used for working in any Unicode compliant Hindi fonts for ensuring inter compatibility of the fonts. Even today Hindi is rendered properly only on Window XP and beyond. Linux has very little support for other languages.

Managing Kashmiri script materials is still a difficult task as no established keyboard support is still available. Perso-Arabic fonts were used for creating Kashmiri collection.

Urdu script has been successfully written using open type font Nafees Pakistani Naskh v. 2.0 and Nafees Tahreer Naskh v. 1.0

At some centres, the task of building digital collections in the Indian languages have been taken, such as the Kannada and Tamil interfaces have been completed. Efforts are also being made to develop interfaces and test digital collections in a few other South Asian languages, such as Malayalam, Bengali, Marathi and Nepali.

#### *Suite of Softwares with Unicode support*

Following are few such suite of softwares for building and distributing multilingual library collections:-

Library management software, Ahmedbad is software which has all the International Standards like MARC21, Z39.50, AACR-II, etc., manages Learning Resource Centre of any type and any size. It also has all the functional modules like Acquisition, Cataloguing, Circulation, Article Indexing, OPAC/ WebOPAC, Administration, etc. it is

multiuser, multilingual and multi-currencies using Unicode technology. Bar-code and RFID options are also available.

A new feature of SOUL is its Unicode based multilingual catalogue module. The automation software has the feature of cataloguing in Indian languages.

Even Libsys which handles Indian languages/ Scripts using ISM publisher and GIST of C-DAC. Libsys also adheres to the Standards such as MARC and Z39.50, suitable for co-operative networking and resource sharing. It runs on various platforms such as Windows (95/98/NT/2000/XP) Unix (Various flavors), Linuxm etc., has various models from acquisition to article indexing system.

VTLS is the first library automation vender that has received ISO certificate in 1997. We, thus, see that VTLS is also an integrated library systems, as said earlier, known as 'Virtua' with thousands of libraries as its users covering major countries of the globe. Its support of Unicode, multilingual database, access to external tools such as OCLC, RLIN, unique data entry templates, adherence to a variety of international MARC formats are few striking elements of its cataloguing module.

*The Centre for Indian Institute of Languages (CIIL) library and its regional centres in India, as we all know are unique in the sense that their information resources are on Indian linguistics languages. This is already the second digital library in India which earlier for its library automation had chosen an international software packages, i.e VIRTUA Integrated Management Software developed by M/s. VTLS Inc., Virginia, USA. Certified by ISO 9001.*

With a main feature of UNICODE support to Indian languages. It all included all the library housekeeping operations such as

Acquisition – Books & serials

Circulation

Reference service

Library management - Planning and budgeting

Stock verification

Networking and Resources Sharing of RLC's libraries

Every package will have its own merits and limitations. It is the individual library to decide which software will be best suited to their requirements in the best way.

*Identifying priorities of Hardware and Software Selections:*

Let's us again come back to our central issue of how to create a technology plan for library automation. Once we finalize the need of assessment and identified possible approaches to meeting these needs, next logical step will be to determine which library functions should be automated and in what order of priority. It must be bore in mind that planning and consulting costs include the direct and indirect costs associated with getting started. The cost of this process may not be immediately apparent. So the planner has to be very cautious and realize that:-

- *Hardware would cover* the computer itself, disk drives, workstations, printers and other machine peripherals.
- *Software would cover* the function-specific modules that he buy, such as acquisition, cataloguing, circulation, serials, articles, WEBOPAC, and also Unicode support.
- *Network-specific hardware, software and cabling* would obviously require the design and implementation of local area network (LAN) architecture on which the system will run, i.e. the selection of appropriate wiring, network architecture, and a network operating system compatible with the system selected.

Once done, next major considerations would be:-

- *Site preparation will* indulge identifying space for the equipment and assuring proper

room ventilation and, as necessary, air conditioning.

- *Staff Training* is utmost, subsequently costs to be considered when the system is first installed.

### **Selection (Purchase) process**

*Technical adviser committee* the consultant works with library planners to provide specific information on the technical aspects of computerization, system capabilities or evaluation of tenders, thereby supporting the decision making process rather than making decision.

The software need into financial and technical specifications, you are ready to begin the actual process of procuring a system. If you are in a Government related institution, you are subject to a procedure of Government rules and regulations are required to purchases of this type to be offered publicly in a competitive manner.

There are basically two procedures, the 'open' procedure and the 'restricted' procedure. The open procedure requires you to place an advertisement in the National newspapers to inviting system suppliers to tender within a 30 days. The disadvantage of the open procedure is that you could be flooded with tenders, all of which must be considered during the evaluation process. Its advantages are that it is more competitive and may drive down costs. The second method is restricted procedure you place an advertisement in the National newspapers inviting suppliers to express an interest in tendering. After a short listing process you then formally invite selected suppliers to demonstrate their system function before the library technical committee.

### *Post-Installation Phase*

Once the suitable software is procured, with close consultation and supervision of the team of computer specialists and library professionals, language specialists, each areas of library service should be customized

according to indigenous need. This means, all housekeeping jobs ought to be tailored-made. It is desirable that the Bar-Code technology may be used for circulation transactions.

We all know that cataloging module is, the most important module as it caters to the needs of database creation of library holdings. The major challenge for any library will be its retro -conversion work. This can be assigned on contract basis to the software vendor/ supplier or any other appropriate agency. What is ultimately required is a team of library professionals. One may categorizes them as one set of team working for the preparation of worksheets and eventual selection of tags for record; other team will do the entry of worksheets; an expert team will work for the verification and validation of the worksheets. A team leader should be appointed for each language from within the organization for the supervision of the work.

Once the cataloguing of the retrospective collection is over i.e information recorded for the final record is complete, the data is input in the client system to prepare the library's database and directly transported to the server. After this, the cataloguing of current resources can be undertaken by the cataloguer in the language concerned.

As said before also, at the data input stage, software must provide languages code support such as Bengali, Dogri, Kannada, Konkani etc. For language display, the user could change the language on the screen at all levels including menu tool bars. The combo box may furnish various Indian languages.

## Conclusion

Every library has to embrace automation. Need of the hour is to go beyond the activities of traditional libraries. It is seen that in the present library scenario, a number of foreign as well as indigenous library automation software packages are being used in India.

India is a multilingual country and in the domain of library and information services,

as the day progresses, the question of multilingual access and multilingual information retrieval is becoming a necessity.

So we hope that in ensuing few years, all library functions operate simultaneously in as many languages as needed and whose search and retrieval functions are language independent.

## References

1. Nadim Akhtar Khan, Rosy Jan and Sheikh Shazia. *Designing digital library of Perso-Arabic script: an experiment*. In 'International Conference on Digital Libraries ICDL 2010: shaping the information paradigm, New Delhi, 23-26 February 2010; 2:778-787.
2. Raghavan, KS, Neelameghan and Lalitha BS. *Digital Libraries in South Asian Languages: some issues and developments*. In 'International Conference on Digital Libraries ICDL 2010: shaping the information paradigm, New Delhi, 23-26 February 2010; 2: 788-794.
3. Sengar DS and Carg RG. *Digital information storage in Hindi language*. In 'International Conference on Digital Libraries ICDL 2010: shaping the information paradigm, New Delhi, 23-26 February 2010; 2: 795-800.
4. Shabahat Hussain and Ansari, Mehtab Alam. *Library automation software packages in India: a study of the cataloguing modules of Alice for Windows, Libsys and Virtua*. *Annals of Library & Info. Studies* 2007; 54: 146-151.
5. Suman Kumar and Sharada BA. *Central Institute of Indian Languages Library*. (available from <http://library.igcar.gov.in/readit-2005>).
6. Manjunath GK. *Library Automation: Why and How?* Available from <http://www.igidr.ac.in/lib/paper1.htm>
7. Cohn John M, Kelsey Ann L and Fiels, Keith Michael. *Planning for library automation: a practical handbook*, London: Library Association pub, 1998.
8. Indian language codes as per ISO standard. Available from [http://www.loc.gov/standards/iso639-2/php/code\\_list.php](http://www.loc.gov/standards/iso639-2/php/code_list.php)
9. Sahitya Akademi library Indian language's MARC Sheet.