20. Localization Case Implements-Banasthali Vidyapith under CoIL-Net

Dr. Aditya Shastri, E-mail: shastri@bv.emet.net

1. Introduction

Banasthali Vidyapith has been selected amongst a handful of Institutions to be a Resource Centre for IT localisation in Rajasthan by the Ministry of Communication & Information Technology (MCIT), Govt. of India. In its initiative on “IT localisation for socio-economic development for the Hindi speaking states” MCIT has started the project “Content Creation of IT Localisation Network (CoIL-Net)” under TDIL programme. The project aims at bridging the digital divide by improving IT penetration and awareness level using Hindi as a medium of delivery, in which academic Institutions from the states of Rajasthan, Uttar Pradesh, Uttarakhand, Madhya Pradesh, Jharkhand, Bihar & Chattisgarh alongwith research groups from C-DAC, Pune, IGNCA, New Delhi and IIT, Kanpur have been identified to develop core technologies for Hindi as well as IT localisation solutions.

The CoIL-Net project at Banasthali Vidyapith has focused on development of content and prototype localized applications in the areas of:

- **e-health**: (i) Health directory of Rajasthan  
  (ii) Ayurvedica
- **e-business**: Banking Information System
- **e-tourism**: Bilingual web portal on Rajasthan
- **e-learning**: (i) Vedic Literature  
  (ii) Distributed ITS for Hindi learning/teaching  
  (iii) Online course on IT Localization and disseminate them by way of holding IT localization clinics in the state. (Figure-1)

The strategy adapted for the development is:

1. To develop prototype applications in the above areas to demonstrate the development in above areas. The development follows the same approach all over which is
   - PC based  
   - Platform independent

2. Web based interactive systems  
3. Bilingual at every stage on the click of mouse  
4. Wherever possible Databases and interactivity both in Hindi and English  
5. XMLisation wherever found appropriate

II. To develop various tools needed for the application development such as

- Transliteration Tool  
- Font Converters  
- Testing and evaluation tools.

III. To disseminate the knowledge by holding trainer's training for educators in the field of **IT in the state**.

IV. To hold IT localisation clinics to provide solutions and to discuss on issues in localisation with the local vendors and entrepreneurs in the field.

The first IT localisation clinic was held in November 2002 in Udaipur where the focus was mainly to create awareness amongst the people involved in IT applications development and in masses in general on the localisation- meaning and issues involved.

Rajasthan, inspite of having designated IT as a priority thrust area, has not been able to let its presence felt in the IT sector. The major Government Departments involved in strategic IT development in the state are

- DoIT  
- Raj Comp  
- NIC-State Unit

All three have on their priority agenda computerization of public utility services such as

- Road Transportation  
- Electricity  
- Water supply  
- Medical & Health Information  
- Land Records

A number of projects are implemented but localization issue is not handled in the projects in the right
perspective resulting in low penetration of the applications in masses.

Banasthali Vidyapith interacted very closely with these departments and then held the second IT Localization in February, 2004 clinic exclusively for IT personnels in State Govt departments for proper deployment of IT localization solutions.

In this paper we present the case study of developing the Localization Applications with specific reference to problems faced in terms of Computing platforms, fonts and encoding related issues, terminology and transliteration.

2. Localization

The process of modifying (adapting) a system or product to be used in the local environment, language and culture of a particular region or locality is termed as Localization. In fact the concept of localization leads to the process of Globalization which is defined as

\[
\text{Globalization} = \text{Internationalization} + \text{N} \times \text{Localization}
\]

The steps to be followed in a software localization are:

(i) **Language enabling** - Encoding, rendering, input, cursor and input key controls.

(ii) **User Interface** - Menus, Dialogs, messages, documentation, Icons, layouts, Color Scheme, Interaction modes.

(iii) **Operational** - Interoperability issues, documentation standards practices.

(iv) **Cultural** - Behavioral models, rules.

In fact the complete process of Software Localization needs to follow a Software Engineering approach using an integrated tool set. The process involves:

- compilation
- translation
- localization
- testing and evaluation
- deployment and documentation

3. Case Study

In each of the applications developed at Banasthali Vidyapith in CoIL-Net project or even earlier than that in the Vidyapith’s Localization efforts (please see Table-1), the experience can be described as follows:

3.1 Compilation of the localized content:

Compilation involves issues such as:

- defining scope
- compiling authentic content (from literature, personally gathering information)
- copyright issues

2.2 Translation: Translation requirement arises for:

- if the content is not available in local language
- translation of terms on user interfaces

There are translation tools available for English to Hindi translation. However, Banasthali Vidyapith did not use any and all translation was done by the project team in consultation with the language experts. Translation raised terminology related issues, very basic ones such as:

- whether to use Hindi term which is uncommon or to use the more common English term and write it in Devnagari - e.g.: Operating System
- whether to use standard terminology (even if not agreeing with it)
- the standard terminology is not complete (possibly can not be) and requires updations.
- Moreover it needs to be available online facilitating search.

2.3 Localization : Localization involved handling issues related to encoding, font rendering and other edit related issues. It also involved handling all factors listed in section 2. Being from a Hindi speaking state the issues of cultural factors in User Interface design were easy to handle and this effect can be seen in the screen layouts developed (sample screens in the figure below). But the encoding related issues are continuously popping up. Using windows based application development requires reedit the whole content as soon as a new version Windows come in the market. Only recently Unicode seems to have resolved these issues but then encoding all content need to be converted to Unicode requiring Code Converter. The Code Convertors provided with the product are not fool proof and the application requires Manual post edit. In fact even fonts need to be available with the application developer in a manner that as soon as a locale is selected in an application the related fonts appear automatically in the font list.
Another important Localization issue is Transliteration, which involves both encoding and font. Not only that, the map table has to be complete, the search in the table needs to be efficient. Moreover, one may not want every input to be transliterated. For example, in databases, there are few fields having proper nouns and the localization means transliterating only those selected fields. How does one do that?

Many e-mail software provide ‘Transliteration’ facility (in fact some provide only that). But on cross server operation the results are not satisfactory (see example screens in figure-2).

Localizing web pages have to answer questions such as:
- Is the website multilingual?
- If yes, what is the strategy for correspondence both English-Hindi pages?
- Design of active pages (we found it hard to design bilingual pages).

2.4 Testing and Evaluation: Our experience finds this face the most difficult. In fact there are nometric present for Testing and Evaluation of localization efforts. We tried to develop - (see example screens in figure-3) for word processors.

One need to develop Benchmarker for evaluation of any localized product/Application launched and then system in place to have mandatory the marketing.

Table-1: Banasthali Vidyapith localization efforts

<table>
<thead>
<tr>
<th>(A) Book writing</th>
<th>DOEACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>कनकल में कंप्यूटर प्रोग्रामन</td>
<td>ओ. लेखाक पद्यमानपुसार हिंदी पुस्तकें</td>
</tr>
<tr>
<td>कंप्यूटर में संबंधित रोबोटिक्स</td>
<td>पी. जी. सोफ्टवेयर</td>
</tr>
<tr>
<td>डिस्क ऑपरेटिंग सिस्टम</td>
<td>बॉले एंड कंप्यूटर प्रोग्रामन</td>
</tr>
<tr>
<td>डीलिक्शन प्लास्टर</td>
<td>‘C’ में कंप्यूटर प्रोग्रामन</td>
</tr>
<tr>
<td>बैप्स एक पृष्ठमय</td>
<td>संकेत प्राकृतिक</td>
</tr>
<tr>
<td>कंप्यूटर घेति एवं सांधन</td>
<td>कंप्यूटर प्रोग्रामन</td>
</tr>
<tr>
<td>उपसंस्थान बैप्स के माध्यम से</td>
<td>प्रस्तुतिसंशोधन</td>
</tr>
</tbody>
</table>

2.5 Deployment & Documentation: There has to be a standard layout of Product Catalogue describing the localization issues and factors to be handled to a novice user.

4. Suggestion for tools development to be available in public domain:
- A subset of English-Hindi terminology available asH d l in a localized application. In fact a general shell architecture in needed which facilitates development of Application terminology glossary.
- A shell for Code-Convertor, where one enters the encoding tables of the pair of codes for which the convertor is to be designed.
- Font enable engine involved with selection of Hindi locale.
- Generalized transliteration engine.
- Content management system for bilingual website design.
- Benchmark for localization products.
- Indian Locale support - Calendar, Calculator, Diary, Currency Converter Games, Maps.

5. A FSM model for Software Localization:

Any developer, whether developing an abinitio localized application or localizing an existing product must first prepare a checklist for compilation, Translation, Localization and testing and the can follow the following FSM model to achieve the desired output.
(B) Web based content Development:

<table>
<thead>
<tr>
<th>श्रेणी</th>
<th>व्याख्या</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>सूचना लेखनीयी</td>
</tr>
<tr>
<td>2.</td>
<td>पी.सी. लेखनीयी</td>
</tr>
<tr>
<td>3.</td>
<td>कंप्यूटर द्वारा प्रोग्रामांने</td>
</tr>
<tr>
<td>4.</td>
<td>'C' द्वारा प्रोग्रामांने</td>
</tr>
</tbody>
</table>

कार्यालय - नेट पायोजना म्हणजेच निर्माणाधीन अनुमोळण

- ई-व्यवस्थापन
- ई-व्यवसाय
- ई-परीक्षण
- ई-सांस्कृतिक एवं दूरस्थ विविधता
Localisation Case Implements Under CoIL-Net Project

Figure 1: CoIL-Net Project

Figure 2: epatra mail sent to yahoo mail server
Metrics Selection

In the context of the Testing of Word Processors we have sorted the Metrics, on the basis of which Testing can be done.

These Metrics are:

1. Bilingual/Multilingual
2. Proper Typing
3. Formatting Features
4. Search and Replace
5. Simple Editing Functions
6. Import/Export
7. Language Sensitive Multilingual Editor
8. Transliteration
9. Keyboard Shortcuts
10. Number to word conversion
11. Table Manipulation
12. Sorting
13. Application Integration
15. HTML Converter
16. Mail Merge
17. Mail Facility
18. Font Converters
19. Proper Text Rendering in the Browser
20. Web Page Development Facility
21. Spell Checker
22. Check File Format
23. Shabdavali
24. Platform Independent
25. Encoding Conversion
26. Various Keyboard Layouts
27. Graphics Display of Keyboard Overlay
28. Self Explanatory User Interface
29. User Configurable Keyboard
30. Keyclick on Screen Keyboard

Marks assigned to each Tool

Following is the result of Evaluation after Test Executions.
Each Tool is Tested on various Metrics and marks assigned to each metric are shown in the Table.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Shabdavali (Glossary)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mail Merge</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Mail Facility</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Facility for Checking File Format</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Proper Typing</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Formatting Features</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Search and Replace</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Simple Editing Functions</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Table Manipulation</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Font Converters</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Proper Text Rendering in the Browser</td>
<td>2</td>
<td>2</td>
<td>*</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Web Page Development Facility</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Multilingual/Bilingual Spelling Checker</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Import/Export</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Platform Independent</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Encoding Conversion</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Keyclick on Screen Keyboard</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Configurable Keyboard</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Graphic Display of Keyboard Overlay</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Various Keyboard Layouts</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Self Explanatory User Interface</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>59</strong></td>
<td><strong>77</strong></td>
<td><strong>68</strong></td>
<td><strong>83</strong></td>
<td><strong>84</strong></td>
</tr>
</tbody>
</table>

*Not confirmed (Because Evaluation copy does not allow Export in HTML Format as well as Pasting to other applications was not supported therefore creating an HTML page was not possible and thus its rendering could not be checked)

**Result after Evaluation**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Measured Value</th>
<th>Rated Level</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISM Version 3.04</td>
<td>59</td>
<td>Fair</td>
<td>Acceptable/Satisfactory</td>
</tr>
<tr>
<td>ISM 2000 Office Version 4.0</td>
<td>77</td>
<td>Good</td>
<td>Acceptable/Satisfactory</td>
</tr>
<tr>
<td>iLeap Version 2.07d</td>
<td>68</td>
<td>Fair</td>
<td>Acceptable/Satisfactory</td>
</tr>
<tr>
<td>Akruti Version 6.0</td>
<td>83</td>
<td>Good</td>
<td>Acceptable/Satisfactory</td>
</tr>
<tr>
<td>Shree Lipi Ankur Patrika Version 1.2</td>
<td>84</td>
<td>Good</td>
<td>Acceptable/Satisfactory</td>
</tr>
</tbody>
</table>