

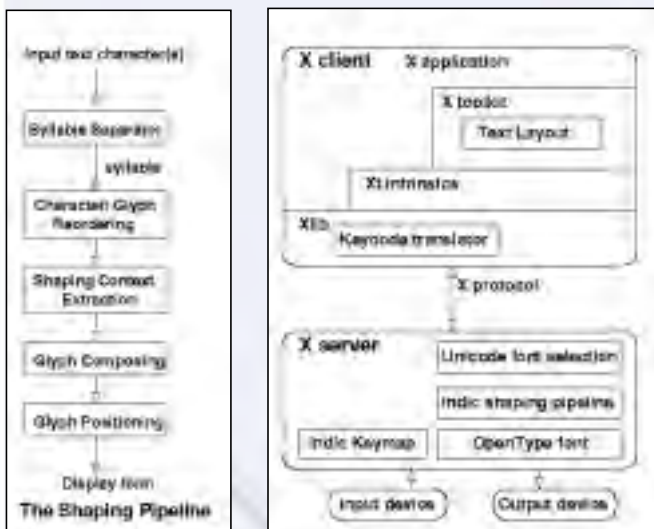
7. Localization of Operating System

Synopsis

<i>Sl No.</i>	<i>Name of The Technology</i>	<i>Nature of the Technology</i>	<i>Name of the Technology Developer</i>	<i>Whether Technology Handshake (TH) has been done</i>
1.	IndiX	Localization of Operating System	C-DAC Mumbai	Software has been put in the Open source domain
2.	Indian LINUX	Localization of Operating System	IIT Kanpur	Open source

7.1.1

- ▶ **Name of the Technology:** *Enabling Localization of Linux Open Source Software for Hindi: IndiX Phase 1*
- ▶ **Nature of Technology :** Technology development leading to production capability, Application-oriented research and development (R&D) having production potential
- ▶ **Level :** Technology and Sub-system
- ▶ **Technical Description of the Technology :** The goal of this project was to provide system level support for Hindi language in GUI of Linux operating system such that most of the existing applications can work with Hindi language without any modification or recompilation. An Indic script-shaping engine was built into the system and the solution was provided using the state of the art technologies like OpenType font and Unicode encoding (See Figure).



Following systems (sub-systems integrated in the X Window System on Linux) were the result of this R&D project:

- Hindi enabled windowing software
- Hindi support in text mode applications
- Basic text processing API
- Hindi locale in Linux format
- Testing and integration of locale aware packages

▶ **Representative Screenshots of the Technology:**



▶ **Scalability / Portability / Expandability:** The software sub-system developed focussed on the Linux open source platform. Due to the operating system-level development nature of the architecture, the result is very scalable but not portable. The developed architecture is very extensible, and accordingly the next version of the system will cover many more languages (specifically 11 more) in much shorter timeframe.

▶ **Readiness of Transfer of Technology (ToT):** The first phase of the system has been completed, and is ready for technology-transfer.

▶ **Availability of Documentation :** The source code and supplementary technical documentation is available on the IndiX website at <http://rohini.ncst.ernet.in/indix>. In addition, following published research papers document the findings and architecture of the system:

- S.P. Mudur, Keyur Shroff, Rekha Gaur, Vinod Kumar, Yogesh Ketkar, "Indic Script Processing Under X11", Symposium on Translation Support System (STRANS-2001), February 16-17, 2001, Kanpur, India.
- S. P. Mudur, Vinod Kumar, Keyur Shroff, "Evolution of Indian Language Localisation

Support in Linux”, Invited paper in International Conference on Software Maintenance (ICSM-2002) sponsored by IEEE, to be held on October 3-6, 2002, Montreal, Canada.

(c) Rekha Sharma, Keyur Shroff, S. P. Mudur, Parvati Rajan, “Development of a Bilingual Electronic Glossary for Automated Assistance in User Interface Localization”, Symposium on Translation Support System (STRANS-2002), March 15-17, 2002, Kanpur, India.

- ▶ **Testing of the Product / Technology :** Few select organizations and individuals participated in the beta-testing of the technology, including University of Hyderabad and VXL Systems. Many users have downloaded the software directly from the website, tried it and given valuable suggestions.

- ▶ **IPR / Open-source :** From the Project goal itself, the software was to be put in the open source domain and correspondingly was built within large number of public domain components. We have not done any commercialisation of the software.

- ▶ **Potential Beneficiaries:**

(a) Offices, websites specialised in using Indian languages

(b) Developers of Hindi enabled popular packages like text editor and web browser. Software developers who can undertake conversion of non-locale enabled packages to locale enabled ones.

- ▶ **User-agency Tie-up:** None

- ▶ **Name and Address of the Resource Person:**

Dr. Alka Irani
Senior Research Scientist
C-DAC Mumbai (Formerly National Centre for Software Technology)
8th Floor, Air India Building, Nariman Point
Mumbai 400 021 Telephone: 91-22-22024641
Web : <http://www.ncst.ernet.in/projects/indix/>

7.1.2

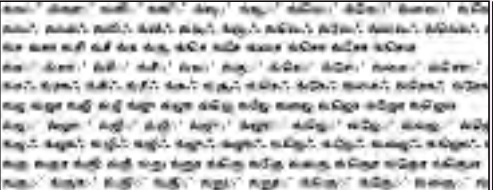

- ▶ **Name of the Technology :** *Localization of Linux and Java (Phase-II) and Development of OpenType fonts for Indian Languages: Assamese, Bengali, Gujarati, Kannada, Malayalam, Marathi, Oriya, Punjabi, Sanskrit, Tamil, and Telugu*

- ▶ **Nature of Technology :** Technology development leading to production capability, Application-oriented research and development (R&D) having production potential

- ▶ **Level:** Technology and Sub-system

- ▶ **Technical Description of the Technology :** The architectural framework of the system has changed in response to the changed technology scenario. Specifically the dependency of the IndiX system on specific aspects of the X Window System has been removed. The IndiX system now has been reorganized. This in effect provides more flexibility to the application developers to make use of the Indic-shaping sub-system. This will also enable regroupings of the IndiX components to adapt to emerging x118n frameworks, as and when they are standardized.

- ▶ **Representative Snapshot / screenshot of the Technology / Product:**

	Stress Testing of Tamil Shaping Engine and Font developed under IndiX Phase 2
	Stress Testing of Malayalam Shaping Engine and Font developed under IndiX Phase 2

- ▶ **Scalability / Portability / Expandability :** To improve the portability, usability and acceptability of the developed software sub-system, this phase focuses on re-organizing the IndiX Phase 1 components in a suitable manner. The development of various components is currently in progress. The planned architecture will suit not only the needs of specifically mentioned

languages, but also provide a base support for many phonetically based scripts in the Indian subcontinent.

- **Readiness of Transfer of Technology (ToT) :** The project is still ongoing. As per current state, support for two languages (Tamil and Malayalam), which includes high-quality OpenType font and stress-tested, robust shaping engine, has been completed. Sanskrit (classical) and Kannada are also near to their completion.
- **Availability of Documentation :** The documentation work is in progress.
- **Testing of the Product / Technology :** Some part is still under development, while some part is under beta-testing phase.
- **IPR / Open-source:** Open Source
- **Potential Beneficiaries:**
 - (a) Offices, websites specialised in using Indian languages
 - (b) Developers of Indian language enabled popular packages like word processors and web browsers.
 - (c) Software developers who can undertake conversion of non-locale enabled packages to locale enabled ones.
- **User-agency tie-up :** None
- **Name and address of the Resource Person:**
Dr. Alka Irani
Senior Research Scientist
C-DAC Mumbai (Formerly National Centre for Software Technology)
8th Floor, Air India Building, Nariman Point
Mumbai 400 021 Telephone: 91-22-22024641
Web : <http://www.ncst.ernet.in>

7.1.3

- **Name of the Technology:** *Indian Linux*
- **Nature of Technology:** Linux distribution having support for application software in Indian languages. It provides a framework for developing software for Indian languages.
- **Level: (Product / Technology / Sub-system)** We support various character coding standards for Indian scripts such as ISCII, Unicode, UTF-8. The ISCII code standard specifies a 8-bit ISO compatible environment. It allows English and Indian script alphabets to be used simultaneously. The ISCII code table is a super-set of all the characters required in ten Brahmi-based Indian scripts.
- **Technical Description of the Technology / Product including Basic Block diagram, Algorithm used, O/S used, Front-end / user interface, and Specification of the Technology / Product :** Various tools have been developed by us that are contributed to Indian Linux, these include –
 - ixterm:** Xwindow based terminal emulator that displays characters in all Indian Scripts as well as in English. Compatible with the rxvt program. The Indian Languages characters can be typed on the keyboard using INSCRIPT or roman phonetic layout. Works with all text based Unix utilities like vi, cat, more, ls etc.
 - iscii2ps :** This tool takes an Indian language text file (coded using character coding standards such as ISCII-8, Unicode, UTF-8) and generates the corresponding postscript output, which can be printed by any printer supported by ghostscript.
 - isciiilib:** A C-library to handle ISCII coded text.
 - iconverter:** Converts one code space to another code space.
 - itransform:** Converts multi-lingual script texts from one code space to another.
 - isciispell :** Spell Checker for the ISCII-coded texts.
 - Xiscii :** Additions to the X-lib to support display and entry of the ISCII text.

XKB : Provide keyboard support for Indian languages

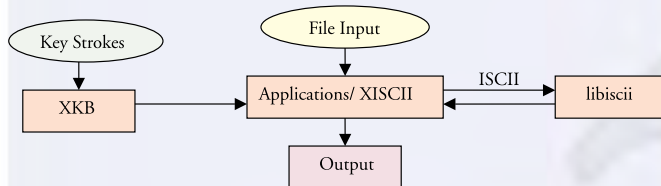
iLaTeX : Convert Indian Languages .tex file to .dvi file

utf2unicode : Converts utf8 coded text file to corresponding unicode.

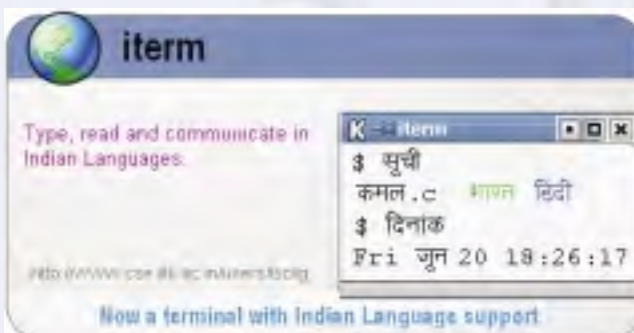
locales : locales are specific to the language. The concept is introduced in ISO C. On setting the locale environment to Indian locales, locale dependent applications(such as date, cal etc) gives output in corresponding Indian language.

Linux-ISCII Distribution: Linux distribution with support for application software in Indian languages

The implementation is materialized on the Linux Platform using C Programming Language.



▶ **Representative Snapshot / screenshot of the Technology / Product:**



▶ **Scalability / Portability / Expandability:** In this project, the ultimate aim is to come up with Linux for Indian Languages. Under our approach, various existing applications run without any modifications. The approach is generic enough to be implemented on other UNIX platforms even though the Linux is being used to develop the support for Indian Languages.

▶ **Readiness of Transfer of Technology (ToT) :** The Indian Linux is ready for use and can be downloaded from the URL~

<http://www.cse.iitk.ac.in/users/isciig>

▶ **Availability of documentation :** Documentation can be found at the URL mentioned above; apart from this man pages and the user manuals etc. also get installed with the corresponding RPMs.

▶ **Testing of the Product / Technology :** Testing of the various tools for Indian Linux has been done.

▶ **IPR / Open-source :**Open Source (Source RPMs are available on the above mentioned URL under downloads link).

▶ **Potential beneficiaries:**

▶ **User-agency tie-up:**

▶ **Name and address of the Resource Person:**

*Dr. Rajat Moona, Professor
CSE Department, I.I.T. Kanpur
Email: moona@cse.iitk.ac.in*

*Dr. Deepak Gupta, Asst. Professor
CSE Department, I.I.T. Kanpur
Email: deepak@cse.iitk.ac.in*