3.24 Limited Domain Speech Synthesis System for Tourists

Technology Developer
Indian Institute of Information Technology
Language Technologies Research Centre
Gachibowli, Hyderabad-500 032
Tel: 3001412 or PABX 3001971
Fax: +(91) (40) 3001413
Prof. Rajeev Sangal, Director
sangal@iiit.net

Level: Product/ Subsystem
Prototype Product

Description
A user with a hand-held device will be able to select what he wants to say in English (using a stylus and menus), and the device will speak it out in Hindi or Telugu.

Encoding Standards
ISCII / Unicode

Portability/ Expandability
Prototype will be based on Linux on a hand held device like simputer.

Type of Technology
Human Machine Interface

Readiness for ToT
Documentation
Installation guidelines and User's guide, to use and configure the system, are under preparation.

IPRs
Not applied yet.

Testing
Under going.

Potential Beneficiaries
Tourists who do not know the local language will be able to communicate with the local people through this device.

User Tie-ups
None

4.1 ITERM - Indian Script Terminal for Unix X-windows

Technology Developer
Indian Institute of Technology
Department of Computer Science & Engineering
Kanpur-208 016
Tel: 0512-597652
Fax: 590725
Prof. Rajat Moona
moona@iitk.ac.in

Level: Product/ Subsystem
Software Subsystem

Description
Using this software it is possible to execute any text based utilities in Unix environment and use the Indian language support. It is possible to name the files and see the directory listing in Hindi. The Hindi files can be edited using terminal based standard Unix utilities such as vi, sed etc. The configuration files supplied with the ITERM are written to support ISCII files, Insiprit keyboard layout and Devanagari TeX fonts.

Encoding Standards
Characters can be coded using ISCII or any other standard. It is possible to support the ISFOC fonts or any other fonts. It supports inscript/ phonetic keyboard layout.

Portability/ Expandability
Few configuration files can be written so that it can be used with wide variety of scripts other then Devanagari.

Type of Technology
Tools

Readiness for ToT
Technology is open and is available in the source form under GNU license

Documentation
Installation guidelines and User's Guide to use and configure the system are being prepared.

IPRs
Not yet applied.

Testing
The system is tested on Linux and Solaris Operating System with X11R6.

Potential Beneficiaries
Users of Unix platform.

User Tie-ups
None
Section IV - Tools

4.2 Lekhan Patra - Bilingual Word Processor for LINUX

➤ **Technology Developer**
Electronics Research & Development Centre
C-56/1, Anusandhan Bhawan,
Sector-62,
Noida – 201301
Tel. 91-4587717-25 Fax: 91-4587726
Shri V.N. Shukla, Director (Special Applications)
shuklawn@hotmail.com

➤ **Level**: Product/Subsystem
Software Product

➤ **Description**
A multilingual communication tool to support Business/Academic activities. Lekhika has special features like Able to create and open multi lingal document, Able to edit and save them preserving their Font, Style and Language etc, Able to import and export various file formats and encoding, Can run concurrently multiple instance of itself, Native system's look and feel, Multilingual Help, Customizable, Scalable, Embedded utilities like calculator, Dictionaries, Spell Checker in local languages & calendar.

➤ **Encoding Standards**
It supports both ISCII & UNICODE storage code.

➤ **Portability/Expandability**
Package can be integrated with any platform as utility. Expandable to multilingual word processor.

➤ **Type of Technology**
Tool

➤ **Readiness for ToT**
Installation guidelines and User’s Guide, to use and configure the system, are being prepared.

➤ **IPRs**
It is being applied for.

➤ **Testing**
Beta testing in progress.

➤ **Potential beneficiaries**
Various organizations working in Hindi having LINUX / UNIX / Windows Platform.

➤ **User Tie-ups**
CTB, New Delhi

4.3 Linux ISFOC Script Manager (LISM)

➤ **Technology Developer**
Centre for Development of Advanced Computing,
GIST Group
Pune University Campus, Ganesh Khind Road
Pune-411007
Tel.: 91-20-5694000
Fax: 91-20-5651111
Shri M.D. Kulkarni
mdk@cdac.ernet.in

➤ **Level**: Product/Subsystem
Software Product

➤ **Description**
· Allows word processing on Linux systems- The user can create documents and text files in Indian Languages using Linux based word processor and text editors.
· Works in multiple documents simultaneously.
· Supports ten Indian scripts- Assamese, Bengali, Devnagari, Gujarati, Kannada, Malayalam, Oriya, Punjabi, Tamil and Telugu.
· Bilingual support- The user can create documents using two scripts simultaneously i.e. English script and any one of the Indian scripts.
· Keyboard Layout- Presently LISM supports Inscript Keyboard layout.

➤ **Encoding Standards**
8-bit Indian Standard Code For Information Interchange (ISCII), Indian Script Font Code (ISFOC).

➤ **Portability/Expandability**
Portable

➤ **Type of Technology**
Tools

➤ **Readiness of ToT**
Installation guidelines and User’s Guide, to use and configure the system, are being prepared.

➤ **IPRs**
C-DAC Pune.

➤ **Testing**
Done

➤ **Potential beneficiaries**
Government & Industries

➤ **User Tie-ups**
None
Tools

4.4 Indica Multilingual Solution (SDK)

➤ Technology Developer
Summit Information Technologies Pvt. Ltd
816, Phase V, Udyog Vihar,
Gurgaon-122016
Tel.: 916-347380-84
Shri Sunil Khullar
sunilkhullar@indya.com
➤ Level: Product/Subsystem
Product/SDK
➤ Description
Indian Script Unicode for Windows 95/98/NT and above. First Indian language solution to support Unicode. Data processing solution in Indian languages supports Oracle 8i/9i and MS SQL 2000 server and MS-Access. Supports development platforms like D2K, VB, ASP, JSP, Java, HTML etc. Supports UNIX, Linux and NT/2000 servers. Delivered as an SDK (software development kit). Consists of Java classes, libraries and DLLs. Stores data in Unicode. Works even for Windows 95/98 and NT. This technology can be used both in Client Server and Web environment.
➤ Encoding Standards
UNICODE & ISCII
➤ Portability/Expandability
Supports Windows and Macintosh as front end. Server could be Unix, Linux, NT/2000, RISC or CISC.
➤ Type of Technology
Tools
➤ Readiness for ToT
Documentation
Available
Testing
A number of prototypes have been developed as proof of concept. Product is thoroughly tested.
➤ Potential beneficiaries
Software development companies developing eGovernance applications for various Govt. depts. State governments who would like to buy enterprise licenses which will meet all their text and data processing requirements.
➤ User Tie-ups
With Govt of Gujarat (rate contract), GNFC and TCIL.

4.5 Bilingual Punjabi-English Word Processor with Spell Checker

➤ Technology Developer
Thapar Institute of Engineering & Technology
Department of Computer Science & Engineering (Deemed University), Patiala 147 001
Tel.: 0175-214868, 393137
Fax: 0175-214498, 216391, 212012, 212002
Shri G.S. Lehal (Asst. Prof.)
Tel:393374(D),283502(R)
glehala@mailcity.com
➤ Level: Product/Subsystem
Software Product
➤ Description
It is a bilingual Punjabi-English word processor with spell checker support. It also supports both phonetic and typewriter keyboard layout for Punjabi and it has a spell checker which can operate on all documents typed in both Punjabi and English.
➤ Encoding Standards
The system encodes the output in ASCII encoding format.
➤ Portability/Expandability
The technology can be used with any Windows platform.
➤ Type of Technology
Tools
➤ Readiness for ToT
Documentation
Installation guidelines and User’s guide, to use and configure the system, are under preparation.
IPRs
Not yet applied
Testing
Self testing going on.
➤ Potential beneficiaries
Government sector, Typist, Educationist.
➤ User Tie-ups
None
4.6 Word Processor for Tamil

Technology Developer
Anna University
School of Computer Science & Engineering
Chennai - 600 025
Tel. 044-2351723 Fax: 044-2350397
Dr. T.V.Geetha/Ms. Ranjani Parthasarathi
rp@annauniv.edu

Level: Product/Subsystem
Software Product

Description
Suvadi is a word processor designed specifically for Tamil language. The menus and dialogues of this word processor are in Tamil. It also supports English. The initial version supports common facilities such as undo, redo, without limit, numbering, bullets, tables etc.

Encoding Standards
It uses TAM/TAB font set encoding, can be made to support ISCII and UNICODE.

Portability/Expandability
Suvadi is developed in a platform independent language using Java so it is highly portable and works with any platform. Since it is using Object Oriented concepts it can be expanded as per requirement. It can be integrated with modules such as Spell Checker, Grammar Checker, Text to Speech, etc.

Type of Technology
Tools

Readiness for ToT
Documentation
Installation guidelines and User’s guide, to use and configure the system, are under preparation.

IPRs
Not yet applied.

Testing
Testing in progress.

Potential beneficiaries
Tamil language users

User Tie-ups
None

4.7 Office Suite for Tamil

Technology Developer
Anna University
School of Computer Science & Engineering
Chennai - 600 025
Tel. 044-2351723 Fax: 044-2350397
Dr. T.V.Geetha/Ms. Ranjani Parthasarathi
rp@annauniv.edu

Level: Product/Subsystem
Software Product

Description
Office Suite in Tamil is a Package with a Word Processor, a Database Front end, a Presentation Tool and a Spread Sheet Tool for Tamil Text / Data processing. Functionality of this technology is similar to MS-Access, Navigation is made simple.

Encoding Standards
The data is stored in .mdb format as in MS Access. Awt components first convert the Unicode characters to the host's native character, the TAB font set encoding. If the target character set does not have the needed Unicode character, a substitute character is often used to represent the original character.

Portability/Expandability
As the development work is in Java, all the code developed is portable. It consists of collection of modules hence supports portability. Since interface for this program is categorized through , new modules can be added very easily.

Type of Technology
Tools

Readiness for ToT
Documentation
Installation guidelines and User’s guide, to use and configure the system, are under preparation.

IPRs
Not yet applied.

Testing
Primary testing completed. Testing has been carried in two ways, i.e., Module testing and Integrated testing. Extensive testing needs to be done.

Potential beneficiaries
Any user who needs or is comfortable with a user interface in Tamil.

User Tie-ups
None
### 4.8 Bangla Spell Checker

**Technology Developer**
Indian Statistical Institute  
Computer Vision and Pattern Recognition Unit  
203, Barrackpore Trunk Road  
Kolkata-700035  
Tel: 033-5778085, 5777694 Fax: 5776680  
Prof. B.B. Chaudhary, Head  
bbc@isical.ac.in

**Level:** Product/Subsystem  
Software Subsystem

**Description**
This Utility programme works with the help of Bangla Editor. Spell checker can act as offline or online. Whenever user types Bangla, spell checks wrong spelling and gives suitable suggestions. On the other hand one can run this software on previously typed material. It has facility to add new words in the dictionary against which spellings are checked. One can delete words from this dictionary too.

**Encoding Standards**
The storage code is ISCII.

**Portability/Expandability**
The system is developed in Microsoft Visual C++ and hence, it runs on Windows/95/98/2000 or Windows NT.

**Type of Technology**
Tools

**Readiness for ToT**
Documentation  
Installation guidelines and User's guide to use and configure the system is ready.

**IPRs**
Not yet applied.

**Testing**
The results of the Alpha testing says: All misspelt words are captured. For a single-error word, intended word is found within top four suggested words in the suggestion list. This is true for all single error words. Words having more than one error are also captured and for most of them intended words are in the upper half of the suggestion list.

**Potential Beneficiaries**
Bangla Newspaper Houses, Publishing Houses, Libraries, Offices looking for Office Automation, Linguistic Community, Writers, Poets, Blind people, etc.

**User Tie-ups**
None

### 4.9 ANUSAARAKA- A Language Accessor among Indian Languages to Hindi

**Technology Developer**
University of Hyderabad  
Department of Computer and Information Sciences  
University of Hyderabad, P.O. Central University  
Hyderabad-500 046  
Tel: 040-3010500(O), 3010846(R)  
Fax: 040-3010145/3010120  
Dr. G. Uma Maheshwar Rao  
guraosh@uohyd.ernet.in

**Level:** Product/Subsystem  
Subsystem

**Description**
Anusaaraka is a computer software which renders text from one Indian language into another. It produces output which is comprehensible to the reader, although at times it might not be grammatical. For example, a Telugu to Hindi anusaaraka can take a Telugu text and produce output in Hindi which can be understood by a Hindi reader, but which is not fully grammatical. Therefore, the reader will require some amount of training for reading the output. Anusaarakas have been built from Telugu, Kannada, Bengali, Marathi, and Punjabi to Hindi. Beta versions of all of these have been released for use over the internet as e-mail servers.

**Encoding Standards**
The storage code is ISCII.

**Portability/Expandability**
The system is developed in Microsoft Visual C++ and hence, it runs on Windows/95/98/2000 or Windows NT.

**Type of Technology**
Tools

**Readiness for ToT**
Documentation  
Installation guidelines and User's guide to use and configure the system is ready.

**IPRs**
Not yet applied.

**Testing**
The results of the Alpha testing says: All misspelt words are captured. For a single-error word, intended word is found within top four suggested words in the suggestion list. This is true for all single error words. Words having more than one error are also captured and for most of them intended words are in the upper half of the suggestion list.

**Potential Beneficiaries**
Bangla Newspaper Houses, Publishing Houses, Libraries, Offices looking for Office Automation, Linguistic Community, Writers, Poets, Blind people, etc.

**User Tie-ups**
None
4.10 Multilingual Text Editor

➤ Technology Developer
University of Hyderabad
Dept. of Computer & Information Sciences
P.O. Central University, Hyderabad-500 046 (AP)
Tel. 040-3010500 Extn. 4056 (D), 3010064(O), 3010374(R)
Fax: 040-3010120, 3010145
Prof. K. Narayan Murthy (CI)
knmcu@uohyd.ernet.in

➤ Level: Product/ Subsystem
Software Product

➤ Description
This is a Platform Independent Robust Editor for Indian Languages. It works under Microsoft Windows as also under Linux. Supports INSCRIPT, Roman as also On-Screen keyboards. It can effectively deal with large files, long lines, etc. Syllable based design - ideal for Indian languages. Supports all major Indian languages and English-swift transliteration between various scripts-allows documents in any language to be read in a script that one knows.

➤ Encoding Standards:
It supports ISCII as well as Unicode.

➤ Portability/ Expandability
It will be integrated with our dictionaries, thesaurus, morphological analyzers and generators, spell checkers, OCR system, Website development tools and our Web Based Education System. It will become a comprehensive platform for a variety of work in Indian Languages. The code is developed in Java and is fully portable - works under MS Windows as well as Linux. It can be easily ported to other Operating Systems as well. The current version is a simple text editor and a number of enhancements are being carried out to improve the functionality.

➤ Type of Technology
Tools

➤ Readiness for ToT
Documentation
Installation guidelines and User's guide, to use and configure the system, are being prepared
IPRs
Not yet applied
Testing
Alpha Testing: Being used in our labs. on a daily basis - regular improvements and enhancements being done.

➤ Potential beneficiaries
Anybody who works with Indian languages, i.e., Students, Teachers, Writers, Publishers, Government Offices, etc.

➤ User Tie-ups
Computer Literacy Centre, Hyderabad.

4.11 VARTALAAP : A multilingual Communication Tool

➤ Technology Developer
National Centre for Software Technology
Gulmohar Cross Road No. 9, Juhu,
Bombay - 400 049
Tel. 022-6201606/6201574/6249817
022-7579935/7812120(D)
Fax: 022-6232195/6210139
Shri Durgesh Rao
durgesh@ncst.ernet.in

➤ Level: Product/ Subsystem
Software Product

➤ Description
A multilingual communication tool designed to support Business activities including meetings and conferences and Academic activities like virtual class rooms. Ideally suited to host discussion forums with a high level of interactivity. The technology uses a unique concept of Object Oriented User Interface where every co-user is represented by his/her own window space. This has the facility to send isolated files as well as files attached with discussion messages. It is capable of handling one-to-one, one-to-many communication in five Indian languages.

➤ Portability/ Expandability
This technology is based on Windows 2000/NT and above. The technology has Generic and Plug-In architecture to allow extensions, also Linux portability is being added. This technology has the potential to work with Knowledge Management Systems.

➤ Type of Technology
Tools

➤ Readiness for ToT
Documentation
Installation guidelines and User's guide, to use and configure the system, are ready.
IPRs
Not applicable (being developed under public domain GPL Model.)
Testing
Being tested by diploma students at Bangalore.

➤ Potential Beneficiaries
Educational Institutions involved in long distance and online training. Small scale organizations requiring decentralized communication.

➤ User Tie-ups
None
4.12 Bilingual Address Management System

➤ Technology Developer
Electronics Research & Development Centre
C-56/1, Anusandhan Bhawan,
Sector-62,
Noida - 201301
Tel. 91-4587717-25
Fax: 91-4587726
Shri V.N. Shukla, Director (Special Applications)
shuklavn@hotmail.com

➤ Level: Product/Subsystem
Software Product

➤ Description
It is a multilingual communication tool to Support Business/ Academic activities. It has front-end in Hindi/ English with Remington Keyboard support. The package generates the Labels/ Reports in Hindi. It has letter drafting facility using MS-Word. The package supports Mail/Merge facility in Hindi/ English and Data Back-Up Facility.

➤ Encoding Standards
The storage code is ISCII.

➤ Portability/Expandability
Expandable to Document Management System. The system can be integrated with MS-Word or any other windows based Word Processor

➤ Type of Technology
Tool

➤ Readiness for ToT
Documentation
Installation guidelines and User’s guide, to use and configure the system, is prepared.

➤ IPRs
It is being applied for.

➤ Testing
Working successfully at various organisations like ICMR, New Delhi & Pondichery, CHD New Delhi and N I H Roorkee, etc.

➤ Potential beneficiaries
Offices Using Address Databases

➤ User Tie-ups
ICMR, New Delhi
Elitex Group, MIT

4.13 Morphological Analyzer and Generators for Tamil

➤ Technology Developer
Anna University
School of Computer Science & Engineering
Chennai - 600 025
Tel. 044-2351723 Fax: 044-2350397
Dr. T.V. Geetha/Ms. Ranjani Parthasarathi
rp@annauniv.edu

➤ Level: Product/Subsystem
Software Subsystem

➤ Description
This software is a complete morphological Generator & Analyzer for Tamil nouns, verbs, adjectives and Postposition. Morphological Generators is one of the two important tools for language processing. Morphological Analyzer is another basic tool required for language processing. Given any word or a group of words, it analyzes the word and determines the root and all the other add ones that it has taken.

➤ Encoding Standards
The code has been developed using Java. It uses a specially designed internal code instead of storing string. The data storage is Unicode compatible.

➤ Portability/Expandability
The system is developed in Java hence fully portable. The system can be integrated with MS-Word or any other windows based Word Processor

➤ Type of Technology
Tool

➤ Readiness for ToT
Documentation
Installation guidelines and User’s guide, to use and configure the system, are under preparation.

➤ IPRs
Not yet applied.

➤ Testing
The system is being tested by linguists.

➤ Potential beneficiaries
All the agencies requiring Tamil Language processing capabilities. It can be integrated into Tamil Language Software products.

➤ User Tie-ups
None
4.14 Morphology for Telugu and Kannada

- **Technology Developer**
  University of Hyderabad
  Dept. of Computer & Information Sciences
  P.O. Central University, Hyderabad 500 046 (AP)
  Tel. 040-3010500 Extn. 4017, 3010064(D), 3010374(R)
  Fax: 040-3010120, 3010145
  Prof. K. Narayan Murthy (CI)
  knmcs@uohyd.ernet.in

- **Level:** Product/Subsystem
  Software Product

- **Description**
  Morphological Analyzers and Generators to analyze the internal structure of words and for generating complete word forms from its parts. Required for language teaching/learning, spell checking, syntactic parsing, translation, etc.

- **Encoding Standards**
  The storage code is ISCII.

- **Portability/Expandability**
  Will be integrated with our editor, spell checker, OCR system, UCSG parsing systems and MAT translation systems.

- **Type of Technology**
  Tools

- **Readiness for ToT**
  Documentation
  Installation guidelines and User's guide to use and configure the system are being prepared.

- **IPRs**
  Not yet applied

- **Testing**
  Being tested against CIIL, Mysore corpora.

- **Potential beneficiaries**
  Language Teachers/Learners, Researchers and Developers working on Spell Checking, Syntactic Parsing, Translation, etc.

- **User Tie-ups**
  None

---

4.15 Sanskrit Authoring System (SAAS)

- **Technology Developer**
  Centre for Development of Advanced Computing
  Ramanshree Plaza, 2/1, Brunton Road,
  Bangalore - 560 025
  Tel. 080-5584205(O) 080-3309215(R)
  Fax:080-5247724
  Prof.V.V.P. Ramanujan
  rama@cdacb.ernet.in

- **Level:** Product/Subsystem
  Software Subsystem

- **Description**
  Editors, Databases and utilities for Sanskrit authors – including index, concordance, data source, hyperlink, etc., quotations extraction, sources.

- **Encoding Standards**
  ISCII, G1ST-SDK

- **Portability/Expandability**
  Compatible to all window platforms.

- **Type of Technology**
  Tools

- **Readiness for ToT**
  Documentation
  Installation guidelines and User's guide to use and configure the system is under preparation and will be ready in a month.

- **IPRs**
  Copyright to be Processed.

- **Testing**
  To be organised.

- **Potential beneficiaries**
  Sanskrit Authors, Researchers, Linguists.

- **User Tie-ups**
  To be established.
4.16 DESIKA-Parser for Sanskrit

- **Technology Developer**
  Centre for Development of Advanced Computing
  Ramanshree Plaza, 2/1, Brunton Road,
  Bangalore - 560 025
  Tel: 080-5584205(O) 080-3309215(R)
  Fax: 080-5247724
  Prof. V.V.P. Ramanujan
  rama@cdacb.ernet.in

- **Level: Product/ Subsystem**
  Software Subsystem

- **Description**
  Word/Sentence analysis of Sanskrit input using Paninian grammar etc.

- **Encoding Standards**
  ISCII

- **Portability/Expandability**
  Gist card required

- **Type of Technology**
  Tools

- **Readiness for ToT Documentation**
  Installation guidelines and User's guide to use and configure the system is under preparation and will be ready in a month.

- **IPRs**
  Copyright by C-DAC

- **Testing**
  To be organised

- **Potential Beneficiaries**
  Any Sanskrit Enthusiast/Learner/Students

- **User Tie-ups**
  To be worked out (Rashtriya Sanskrit Sansthan to begin with).

4.17 Publishing Websites on CDROM’s

- **Technology Developer**
  Indian Institute of Technology
  Department of Computer Science & Engineering
  Kanpur- 208016
  Phone: 0512-597618
  Fax : 0512-590725
  Professor T.V. Prabhakar
  tvp@iitk.ac.in

- **Level: Product/ Subsystem**
  Software Product

- **Description**
  In this work it is proposed to publish websites on CD ROM’s. The problem of mirroring a website is addressed. A tool is developed for creating a mirror of website and then solve the associated problems. Browsing on the CD ROM has the same look and feel of original website. A applet is also developed for synonym-based search of such CDROM content with provision for pluggable dictionaries, which contain synonym definitions. It can run on Linux and Windows platforms.

- **Encoding Standards**
  Not relevant since existing format of the site will be used

- **Portability/Expandability**
  Portable and expandable- uses WebL and Java

- **Type of Technology**
  Tools

- **Readiness for ToT Documentation**
  Installation guidelines and User's guide to use and configure the system is under preparation and will be ready in a month.

- **IPRs**
  Technology uses published standards. Nothing is proprietary

- **Testing**
  To be organised

- **Potential Beneficiaries**
  CD-ROM publishers

- **User Tie-ups**
  None
## 4.18 eXtensible Chhand Markup Language

**Technology Developer**
Indian Institute of Technology  
Department of Computer Science & Engineering  
Kanpur- 208016  
Phone: 0512-597618  
Fax : 0512-590725  
Prof. T.V.Prabhakar  
tvp@iitk.ac.in

**Level:** Product/Subsystem  
Software Subsystem

**Description**  
Metric prosody is a form of poetry which is characterised by the rhythm and pattern. This is a context sensitive grammar. Standard tools like Lex and YaCC cannot be used to capture the grammar of meters. The objective of this project is to create a user friendly tool which takes bicomponential input:- the rules of the meter and the metric text, and recognises the meters in the text according to the rules. Here XML is used to define the specification interface for the rules. The specification language is capable of defining metric prosody in many natural languages. XML was chosen because of its strength as a meta-language and availability of parsers like DOM.

**Encoding Standards**  
ASCII, ISCII, font based

**Portability/Expandability**  
Portable and expandable since it uses XML

**Type of Technology**  
Tools

**Readiness for ToT**  
Documentation  
Installation guidelines and User's guide, to use and configure the system, are being prepared.

**IPRs**  
None

**Testing**  
No alpha, beta testing done yet.

**Potential Beneficiaries**  
Literary scholars

**User Tie-ups**  
None

## 4.19 A Devanagari Search Engine for Unicode Documents with Compression

**Technology Developer**
Indian Institute of Technology  
Department of Computer Science & Engineering  
Kanpur- 208016  
Phone: 0512-597618  
Fax : 0512-590725  
Professor T.V.Prabhakar  
tvp@iitk.ac.in

**Level:** Product/Subsystem  
Software Subsystem

**Description**  
A search engine capable of searching the web documents written in languages other than English, is highly needed, especially when more and more sites are coming up with localized content in multiple languages. In this thesis, a web based Search Engine for Indian Languages has been implemented. This software allows full-text indexing and searching of a database of UTF-8 documents written in Brahmi-based Indian Language. The Search engine gathers the HTML documents from the web, indexes and compresses the documents, then searches for the given keywords. The main features of the search engine are phonetic tolerance, compression and indexing, leading and trailing substring matches for keyboards, search through compressed documents and WYSIWYG front end for Indian Languages.

**Encoding Standards**  
It is Unicode based.

**Portability/Expandability**  
It is portable and expandable.

**Type of Technology**  
Tools

**Readiness for ToT**  
Documentation  
Installation guidelines and User's guide, to use and configure the system, are being prepared.

**IPRs**  
Not yet applied

**Testing**  
Alpha Beta testing not done yet.

**Potential Beneficiaries**  
Browser users, those who visit Indian languages sites.

**User Tie-ups**  
None
4.20 Webdunia Language Tools

➤ Technology Developer
Webdunia.com (India) Pvt. Ltd.
Uttore Building, 7/2 New Palasia
Indore (M P) - 452001 IN D I A
Tel : 0731-548550/1/2/3/4
Fax : 0731-548555
Sh. Vinay Chhajlani, M D
shishir@webdunia.com

➤ Level: Product/ Subsystem
Both Software Products & Subsystem

➤ Description of Tools

1. Webdunia Browser - enables you to connect to the Internet to gain fast access to data, information, news and opinions. Its menu bar and dialog window can be viewed in all Indian languages. It is developed in VC++ on windows platform.

2. Webdunia Composer helps in composing mails in 11 different Indian languages in conjunction with outlook express. It is developed in VC++ on windows platform.

3. Webdunia Pad is a full-fledged multilingual word processing application. The menus and dialogs interface can be set to the language of user's choice. Documents can be prepared in eleven Indian languages along with English. It is developed in VC++ on windows platform.

4. Webdunia Type-n-cast is a utility which helps you to type and paste Indian language text in an active application, where it is needed. Text can be pasted in 11 Indian languages. It is developed in VC++ on windows platform.

5. Webdunia Software Development Kit has been provided mainly for developers. This is a full-fledged Active X control made to be embedded in any application on any development platform. It is an editor window, which facilitates typing in Indian Languages using different keyboards besides all text formatting operations. It is developed in VC++ on windows platform.

6. Webdunia’s W M S (Web-Mail-Service) has a complete multilingual email service and infrastructure set up, giving our partners the ability to experience a superior, feature-rich and comprehensive email service. CORPORATE / GOVT. / SEMI-GOVT. / INDIVIDUAL can provide email under their own brand name as Webdunia works seamlessly in the background to ensure uninterrupted service. Email enhancements, upgrades and new features are seamlessly introduced at the server level, making adding functionality easy and avoiding costly delays caused by shipping, installing, testing and debugging of software.

1. Features of the Mail: The most feature rich mail Web based as well as POP3 Multilingual Mail Service
1. Multilingual Mailing System. (11 Languages)
2. Multilingual Mail Sending And Receiving In Any Email Client Or From Other Site.
3. Secured authentication Of Users Using SSL.
4. Send, receive, forward and reply to email messages
5. Store sent and received mail
6. Send and receive attachments (size varies depending on amount of storage)
7. Create, delete, and manage folders.
10. Multiple POP retrieval And Storing Facility.
11. Settings: Signatures, Return address, Multiple Personalities, change password, auto-response, save copies of outgoing mail
12. Send and receive HTML-formatted messages
13. Sorting functions
14. Vacation Reply
15. Mail Blocking for Specific Mail Address And Specific Domain.

➤ Encoding Standards
ASCII, ISCII, Unicode Compl is Unicode Compliant.

➤ Portability/Expandability
Available for Windows Platform & work under progress for Linux.

➤ Type of Technology
Tools

➤ Readiness for ToT Documentation
Help/user documents available.

IPRs
Webdunia.com (India) Pvt. Ltd.

Testing
Tested & in use.

➤ Potential Beneficiaries

➤ User Tie-ups
None
4.21 Fonts and Keyboard Handler

➤ Technology Developer
Modular Infotech Private Limited
26A, Electronic Co-operative Estate
Pune-Satara Road, Pune-411009
Tel. : 91-20-4223342
Fax : 91-20-4225896
Shri M.N. Cooper
mncoooper@modular-infotech.com

➤ Level: Product/Subsystem
Software Subsystem

➤ Description
Unlike English, Indian language requires composition. Here it is required to substitute one string of key code sequence with another to form a conjunct character. Different matras and exceptions should be properly displayed as per language rules. The fonts are designed to support a very large set of conjuncts. The fonts are available in TrueType format or Adobe Type1 format.

➤ Encoding Standards
There are as yet no acceptable standards on Indian language fonts. A variety of keyboard layouts are supported for each language.

➤ Portability/Expandability
The keyboard handler and fonts can be installed in Windows 95 to Windows 2000. It works with most of the applications.

➤ Type of Technology
Font Technology

➤ Readiness for ToT
Documentation
Help files and user manual available
IPRs
Belongs to Modular Infotech Pvt. Ltd.
Testing
Has undergone extensive testing

➤ Potential beneficiaries
Indian community wanting to use Indian language on PC.

➤ User Tie-ups
We are marketing Keyboard handlers and fonts under the name of Shree-Lipi. It has a user base over 40,000.

4.22 Indica Multilingual Solution

➤ Technology Developer
Summit Information Technologies Pvt. Ltd
816, Phase V, Udyog Vihar,
Gurgaon-122016
Tel. : 916-347380-84
Fax : 916-347136
Shri Sunil Khullar
sunilkhullar@indya.com

➤ Level: Product/Subsystem
Software Subsystem

➤ Description
Fonts: A library of over 250 very high quality fonts in all Indian languages.

➤ Encoding Standards
N.A.

➤ Portability/Expandability
Fonts are available as Postscript Type1, TrueType, OpenType or PFR for Web.

➤ Type of Technology
Tools

➤ Readiness for ToT
Documentation
Ready
IPRs
Fonts like Chanakya, Geetanjali, Panchari etc.

➤ Testing
The Font are being used with Indica in an installed base of over 20,000 users.

➤ Potential beneficiaries
Software development companies developing eGovernance applications for various Govt. deptts. State governments who would like to buy enterprise license which will meet all their text and data processing requirements.

➤ User Tie-ups
Custom development for end-users and Software developers.
4.23 Kannada Fonts

➤ **Technology Developer**
Indian Institute of Science
Department of Electrical Engineering
Bangalore-560012
Tel.: 080-3092222, 3466022
Fax: 080-3600683
Prof. V. Vinay, Prof. N. J. Rao
njrao@mgmt.iisc.ernet.in

➤ **Level:** Product/ Subsystem
Software Subsystem

➤ **Description**
Developed high quality glyphs for Kannada in Metafont. This can be integrated with the Latest, text composing system to generate high quality output.

➤ **Encoding Standards**
Trying to port to the standards imposed by the Karnataka Government

➤ **Portability/ Expandability**
The font can be used on all unix/linux machines.

➤ **Type of Technology**
It is currently written in the Metafont programming language.

➤ **Readiness for ToT**
Documentation: Minimal.

IPRs: None

Testing: The main aim of developing these fonts was to throw the fonts open to public domain. We are in a position to share our design with any organization.

➤ **Potential beneficiaries**
Anybody who wants to use Kannada fonts.
Anybody who want to use Latex for high quality composing, especially of scientific/ mathematics articles or books.

➤ **User Tie-ups**
None envisaged currently.

4.24 Font Converters and Other Tools

➤ **Technology Developer**
Indian Institute of Information Technology
Language Technologies Research Centre
Gachibowli, Hyderabad-500 032
Tel: 3001412 or PABX 3001971
Fax: +(91) (40) 3001413
Prof. Rajeev Sangal, Director
sangal@iiit.net

➤ **Level:** Product/ Subsystem
Software Subsystem

➤ **Description**
1. **Font Converters**: Converters among various fonts/glyphs. Between ISCII and a large number of Devanagari fonts. Technology developed for building converters quickly.

2. **Indian language support on Linux**: Indian language TTS support on Linux alpha version `xpt-idev` which works for all applications running under `xpt` in `x-windows`.

3. **Indian language support for browser**: (a) Server is in ISCII (character-based) (b) client side may have any font installed in the browser (c) font independent (with the help of converters from/to ISCII) (d) platform independent.

4. **Corpus and other tools**: (a) language detector which figures out the language of an unknown text (b) corpora organisation and search tools (c) word and n-grams.

➤ **Encoding Standards**
ISCII

➤ **Portability/ Expandability**
Open source code under GPL therefore users can easily adapt it for their use.

➤ **Type of Technology**
Tools

➤ **Readiness for ToT**
Documentation: To be prepared.

IPRs: Yes. (Available for free under GPL license).

Testing: Not done yet.

➤ **Potential beneficiaries**
Anybody having access to the internet and wanting to use Indian languages.

➤ **User Tie-ups**
None
### 4.25 Aakanksha

| Developer | Banasthali Vidyapeeth  
Banasthali (Rajasthan)  
Tel. 01438-27847/51(O)  
Fax: 01438-27850  
Prof. Rekha Govil  
rekha.govil@yahoo.com |
|-----------|------------------------------------------------------------------------------------------------------------------|
| Level     | Product/Subsystem  
Software Product |
| Description | Aakanksha is a complete visual environment in Hindi for children having three modules:  
Ankan-for learning Numbers  
Akshar-for learning alphabets  
Aakar-for drawing and coloring  
It is a play environment web enabled. Children should enjoy it even if they do not know Hindi, in fact they can learn Hindi using Aakanksha. |
| Encoding Standards | Uses ISCII fonts. |
| Portability/Expandability | Web enabled, needs a browser on any platform. |
| Type of Technology | Tools |
| Readiness for TOT | Documentation  
Online help plus a manual available. |
| IPRs | Not yet applied |
| Testing | Educational testing remains. |
| Potential beneficiaries | Children of any kind, age and breed. |
| User Ties-ups | None |

### 4.26 Gist-II ASIC

| Developer | Centre for Development of Advanced Computing,  
GIST Group  
Pune University Campus, Ganesh Khind Road  
Pune-411007  
Tel.: 91-20-5694000  
Fax: 91-20-5651111  
Shri M.D. Kulkarni  
mdk@cdac.ernet.in |
|-----------|------------------------------------------------------------------------------------------------------------------|
| Level     | Product/Subsystem  
Engineering sample of GIST-II chip. Prototype board based on Gist-II chip. |
| Description | CRT Controller.  
Color Support.  
Support up to 2 MB DRAM.  
Support up to 2 MB SRAM.  
Support up to 2 MB ROM.  
16 bit peripheral interface.  
Printer interface.  
AT Keyboard Interface.  
Gen lock.  
Operates in PC mode and Terminal mode.  
Max. operating Freq – 40 MHz. |
| Encoding Standards | Not Applicable. |
| Portability/Expandability | Web enabled, needs a browser on any platform. |
| Type of Technology | Tools |
| Readiness for TOT | Documentation  
Internal Diagrams for Gist-II  
Gist Terminal prototype board.  
All related documents |
| IPRs | Owned by C-DAC Pune. (Sponsored by DSIR)  
Full fledged Gist Terminal VT-52, VT-100/200 compatible  
Gist Terminal has been developed. |
| Potential beneficiaries | M ultilingual support to Various Hardware products and Video products. |
| User Ties-ups | Manufacturer |