

7. Resource Centres Technology Index

I. Indian Institute of Technology, Kanpur (Hindi, Nepali)

1. Machine Translation
2. Speech-to-Speech Translation
3. Lexical Knowledge-Base Development
4. Optical Character Recognition
5. Transliteration
6. Spell Checker Design
7. Knowledge Resources
 - 7.1 Gitasupersite
 - 7.2 Brahmasutra
 - 7.3 Complete Works of Adi Sankara
 - 7.4 Ramcharitmanas
 - 7.5 Upanishads
 - 7.6 Kavi Sammelan
 - 7.7 Bimari-Jankari
 - 7.8 Paramhans Ram Mangal Das Ji
 - 7.9 Nepali Texts
8. Technical Issues
9. Details of Texts/Commentaries for each of the Upanishads
10. The Team Members

II. M.S. University of Baroda, Vadodara (Gujarati)

1. Knowledge Resources
 - 1.1 Gujarati WordNet
2. Knowledge Tools
 - 2.1 Portal
 - 2.2 Multilingual Text Editor
 - 2.3 Code Converter
 - 2.4 Gujarati Spell Checker
3. Translation Support Systems
 - 3.1 Machine Translation
4. Human Machine Interface Systems
 - 4.1 OCR for Gujarati
 - 4.2 Text To Speech (TTS)
5. Localization
 - 5.1 Language Technology Human Resource Development
6. Standardization
 - 6.1 Unicode
7. Products which can be launched and the service which the Resource Centre can provide to the

State Government and the Industry in the region.

- 7.1 Technical Skills
- 7.2 Products Which can be Launched
- 7.3 IT Services

8. IT Services - Multimedia CDs
 - 8.1 Learn Gujarati Through English
 - 8.2 Adolescent Health
 - 8.3 Bibliography of books and journals
 - 8.4 Classic Knowledge-base
9. The Team Members

III. Indian Institute of Technology, Mumbai (Marathi, Konkani)

1. Lexicon And ontology
 - 1.1 Lexicon
 - 1.2 Ontology
2. The Hindi Wordnet
3. The Marathi Wordnet
4. Automatic Generation of Concept Dictionary & Word Sense Disambiguation
5. Hindi Analysis and Generation
6. Marathi Analysis
7. Speech Synthesis for Marathi Language
8. Project Tukaram
9. Automatic Language Identification of Documents using Devanagari Script
10. Object Oriented Parallel & Distributed Web Crawler
11. Designing Devanagari Fonts (Three Types)
12. Low Level Auto Corrector
13. Font Converters
14. Marathi Spell Checker
15. IT Localisation
16. Publications
17. The Team Members

IV. Thapar Institute of Engineering & Technology, Patiala (Punjabi)

1. Products Developed
 - 1.1 Spell Checker
 - 1.2 Font Converter
 - 1.3 Sorting Utility
 - 1.4 Bilingual Punjabi/English Word processor

- 1.5 Gurmukhi OCR
- 1.6 Gurmukhi to Shahmukhi Transliteration
2. Contents Uploaded on Internet
 - 2.1 Punjabi Classic Literature
 - 2.2 Bilingual Dictionaries and Glossary
 - Punjabi English On-line Dictionary
 - English Punjabi On-line Dictionary
 - Hindi Punjabi On-line Dictionary
 - Glossary of English-Punjabi Administrative Terms
 - 2.3 On Line Teaching of Punjabi
 - 2.4 On Line Font Conversion Utility
 - 2.5 Downloads
 - Punjabi Spell Checker
 - Punjabi Fonts
3. Interaction with Punjab State Govt.
4. Publications
5. The Team Members

V. Indian Statistical Institute, Kolkata (Bangla)

1. Core Activities
 - 1.1 Website Development and the Language Design Guide
 - 1.2 Training Programmes
2. Services
 - 2.1 Corpus Development
 - Printed Bangla Document Images & Ground Truth for OCR and Related Research
 - Development of Bangla Text Corpus in Electronic Form (including a Bangla Dictionary, and Several Bangla Classics)
 - Electronic Corpus of Speech Data
 - 2.2 Font Generation and Associated Tools
 - Public Domain Bangla Font Generation
 - Converter-Font File to ISCII File and Vice-Versa.
 - Bangla Text Editor
 - Bangla Spell Checker
3. Products
 - 3.1 OCR system for Oriya
 - 3.2 Adaptation of Bangla OCR to Assamese

- 3.3 Information Retrieval System for Bangla Documents
- 3.4 Script Identification and Separation from Indian Multi-Script Documents
4. Research & Development
 - 4.1 Automatic Processing of Hand-Printed Table-Form Documents
 - 4.2 Research and Development of Neural Network Based Tools for Printed Document (in eastern regional scripts) Processing
5. The Team Members

VI. Utkal University (Oriya)

1. Intelligent Document Processing (OCR for Oriya)
2. Natural Language Processing (Oriya)
 - 2.1 Oriya Machine Translation System (OMTrans)
 - 2.2 Oriya Word Processor (OWP) (Multilingual)
 - 2.3 Oriya Morphological Analyser (OMA)
 - 2.4 Oriya Spell Checker (OSC)
 - 2.5 Oriya Grammar Checker (OGC)
 - 2.6 Oriya Semantic Analyzer (OSA)
 - 2.7 E-Dictionary (Oriya 1 English)
 - 2.8 OriNet (WordNet for Oriya)
 - 2.9 SanskritNet (Word Net in Sanskrit)
3. Speech Processing System for Indian Languages
 - 3.1 Text to Speech (TTS)
 - 3.2 Speech To Text (STT)
4. The Team Members

Orissa Computer Application Centre (Oriya)

1. Products Developed at OCAC
 - 1.1 Oriya Spell Checker
 - 1.2 Thesaurus in Oriya
 - 1.3 Bilingual Electronic Lexicon
 - 1.4 Corpus in Oriya
 - 1.5 Bilingual Chat Server
 - 1.6 Net Education
 - 1.7 XML Document Creation and Manipulation
 - 1.8 Oriya OCR

- 1.9 Oriya E-Mail
- 1.10 Oriya Word Processor with Spell Checker under Linux
- 1.11 Computer Based Training (in Oriya)
- 1.12 Oriya Language Based E-Governance Applications
2. ToT done for various projects so far.
3. Training Programmes run for the officials for the state Govt.
4. Core Activities
 - 4.1 Web Hosting of Oriya Languages Classics
 - 4.2 Hosting of Web Sites of Govt. Colleges in Orissa.
5. Products Proposed to be Developed
6. The Team Members

VII. Indian Institute of Technology, Guwahati (Assamese & Manipuri)

1. Knowledge Resources
 - 1.1 Corpora
 - Assamese Corpora
 - Manipuri Corpora
 - 1.2 Dictionaries
 - 1.3 Design Guides
 - 1.4 Phonetic Guides
2. Knowledge Tools
 - 2.1 The RCILTS, IIT Guwahati Website
 - 2.2 Fonts
 - 2.3 Spell Checker
 - 2.4 Assamese Language Support for Microsoft Word
 - 2.5 Morphological Analyzers
3. Translation Support System
4. Human Machine Interface Systems
 - 4.1 Optical Character Recognition for Assamese and Manipuri
 - 4.2 Speech Recognition System
 - 4.3 Interface for e-Dictionary
5. Language Technology Human Resource Development
6. Standardization
7. Publications
8. Manika Newsletter
9. The Team Members

VIII. Indian Institute of Science, Bangalore, (Kannada)

1. Web Sites and Support to Instruction
 - 1.1 Kannudi
 - 1.2 LT-IISc
 - 1.3 Bodhana Bharathi : Multimedia Educational CDs for 7th, 8th and 10th Standards
 - 1.4 Bilingual Instructional Aid for Learning German through Hindi
 - 1.5 Information Base in Hindi Pertaining to German History
2. Knowledge Bases
 - 2.1 Sudarshana a Web Knowledge base on Darshana Shastras
 - 2.2 Indian Logic Systems
 - 2.3 Indian Aesthetics
3. Technologies and Language Resources
 - 3.1 Brahmi : Kannada Indic Input Method, Word processor
 - 3.2 Open Type Fonts : Sampige, Mallige, Kedage
 - 3.3 Kannada Wordnet
 - 3.4 OCR for Tamil
 - 3.5 OCR of Printed Text Document in Kannada
4. Research
 - 4.1 Automatic Classification of Languages Using Speech Signals
 - 4.2 Algorithms for Kannada Speech Synthesis
5. Publications

IX. University of Hyderabad (Telugu)

1. RC for ILTS
2. Products
 - 2.1 DRISHTI: OCR
 - 2.2 Tel-Spell : Spell Checker
 - 2.3 AKSHARA : Advanced Multi-Lingual Text Processor
 - 2.4 E-mail in Indian Languages
 - 2.5 WILIO : Interactive Web Page in Indian Languages
 - 2.6 Telugu Corpus
 - 2.7 Dictionaries, Thesauri & other Lexical Resources

- 2.8 Morphology
- 2.9 Stemmer
- 2.10 Part of Speech Tagging
- 2.11 VIDTA : Comprehensive Toolkit for Web-Based Education
- 2.12 Grammars & Syntactic Parsers
- 2.13 Machine Aided Translation
- 2.14 Tools
 - 2.14.1 Font Decoding
 - 2.14.2 Web Crawler for Search Engine
 - 2.14.3 PSA : A Meta Search Engine
 - 2.14.4 Corpus Analysis Tools
 - 2.14.5 Website Development Tools
 - 2.14.6 Character to Font Mapping Tools
 - 2.14.7 Dictionary to Thesaurus Tools
 - 2.14.8 Dictionary Indexing Tools
 - 2.14.9 Text Processing Tools
 - 2.14.10 Finite State Technologies Toolkit
- 3. Service & Knowledge Bases
 - 3.1 On-line Literature
 - 3.2 History-Society-Culture Portal
 - 3.3 On-Line Searchable Directory
 - 3.4 Character Encoding Standards, Roman Transliteration Schemes, Tools
 - 3.5 Research Portal
 - 3.6 VAANI : A Text to Speech System for Telugu
 - 3.7 Manpower Development
- 4. Epilogue
 - 4.1 Strengths and Opportunities
 - 4.2 Outreach
- 5. Publications
- 6. The Team Members
- X. Centre for Development of Advanced Computing – Thiruvananthapuram (Malayalam)
 - 1. Human Machine Interface Systems
 - 1.1 NAYANA™ – Optical Character Recognition System for Malayalam
 - 1.2 SUBHASHINI™ – Malayalam Text to Speech System (TTS)
 - 2. Knowledge Tools
 - 2.1 NERPADAM™ – Malayalam Spell Checker

- 2.2 AKSHARAMAALA™ – Malayalam Font Package and Script Manager
- 2.3 Text Editor
- 2.4 Code Converters
- 2.5 ANWESHANAM™ – Malayalam Web Based Search Engine
- 2.6 Malayalam Portal
- 3. Services
 - 3.1 SANDESAM™ – Malayalam E-mail Server
 - 3.2 Malayalam E-Com Application
- 4. Knowledge Resources
 - 4.1 Trilingual Dictionary
- 5. Language Tutors
 - 5.1 Ezhuthachan – The Malayalam Tutor
 - 5.2 English Tutor
- 6. Other Activities
 - 6.1 Providing Technology Solutions
 - 6.2 Interaction with State Government
 - 6.3 Training Activities
 - 6.4 COWMAC (Consortium of Industries working in the area of Malayalam Computing)
- 7. Publications
- 8. Expertise gained
- 9. Future Plans
- 10. The Team Members
- XI. School of Computer Science & Eng., Anna University Chennai (Tamil)
 - 1. Knowledge Resources
 - 1.1 Online Dictionary
 - 1.2 Corpus Collection Tools
 - 1.3 Contents Authoring Tools
 - 1.4 Tamil Picture Dictionary
 - 1.5 Flash Tutorial
 - 1.6 e-Handbook of Tamil
 - 1.7 Karpanaikkatchi- Scenes from Sangam Literature.
 - 1.8 District Information
 - 1.9 Educational Resources
 - 2. Knowledge Tools
 - 2.1 Language Processing Tools
 - Morphological Analyser
 - Morphological Generator

- 2.2 Text Editor
 - 2.3 Spell Checker
 - 2.4 Document Visualization Tool
 - 2.5 Utilities
 - Code Conversion Utility
 - Tamil Typing Utility
 - 3 Translation Support System
 - 3.1 Tamil Parser
 - 3.2 Universal Networking Language (UNL) for Tamil
 - 3.3 Heuristic Rule Based Automatic Tagger
 4. Human Machine Interface System
 - 4.1 Text-to-Speech (Ethioli)
 - 4.2 Poonguzhali (A Chatterbot)
 5. Localization
 - 5.1 Tamil Office Suite (Aluval Pezhai)
 - Tamil word Processor (palagai)
 - Presentation Tools for Tamil (Arangam)
 - Tamil Database (Thaenkoodu)
 - Tamil Spread Sheets (chathurangam)
 - 5.2 Tamil Search Engine (Bavani)
 6. Language Technology Human Resources Development
 - 6.1 Creating awareness among the general public
 - 6.2 Language Technology Training
 - 6.3 Co-ordination with Linguistic & Subject Exports
 - 6.4 Co-ordination with Govt. & Industry
 7. Standardization
 8. Research Activities
 - 8.1 Text Summarization
 - 8.2 Latent Semantic Indexing
 - 8.3 Knowledge Representation System based on Nyaya Shastra
 9. Publications
 10. The Team Members
- XII. Perso-Arabic Resources Centre, CDAC, GIST, Pune (Urdu, Sindhi, Kashmiri)**
1. The PASCII Standard
 - 1.1 Characteristics of Proposed Standard
 - 1.2 Standardization of Urdu Keyboard
 - 1.3 Standardization of Fonts
2. Script Core Technology & Rule Engine
 - 2.1 Glyph Property Editor
 - 2.2 Rule Engine
 - 2.3 Testing Version of Urdu Editor
 - 2.4 Font Glyph Editor
 - 2.5 The Font Editor Utility
 3. Devanagari/Gurmukhi to Urdu Transliteration "UTRANS"
 - 3.1 Following is a Sample of Hindi Text Transliterated into Urdu (UTRANS)
 - 3.2 Following is a Sample of Punjabi Text Transliterated into Urdu (UTRANS)
 4. Dictionary Development Tools
 5. Nashir – Word processor
 6. The Urdu SDK Controls
 7. Pocket Translator
 8. Multipromter
 9. LIPS Advance Creation Station & LIPS for DVD
 10. Perso-Arabic Web Site Development
- XIII. Jawaharlal Nehru University, Delhi (Sanskrit, Japanese, Chinese)**
1. Sanskrit Lessons and Exercises
 2. Devanagari Script Learning and Character Recognition System
 3. Sanskrit-English Lexicon
 4. Dictionary of Nyaya Term
 5. Panini Astadhyayi
 6. Dhatu to DhatuRoop
 7. DhatuRoop to Dhatu
 8. Sandhi
 9. Japanese Lessons
 10. Japanese Script Learning and Character Recognition System
 11. Japanese-Hindi Lexicon
 12. HTML Content Converter
 - 12.1 Other Code Converters
 13. Current/Future Tasks