

For each of these outputs specific Indicators and Means of verification were also worked out. With regard to centre specific activities, it was agreed that each centre would do a similar exercise and submit a PPM.

The role of C-DAC in transferring the technology already developed was recognized. It was agreed that a technology workshop would be conducted with C-DAC in Pune by August 2000 to enable the Resource Centres for a speedy take off in the respective activity. It was also felt that the experience of IIT, Kanpur would need to be shared with other Resource Centres for successful execution of the project.

Discussions also took place on the review mechanism of the project. It was agreed that the indicators evolved from the Project Planning Matrix would be utilized for monitoring the progress. Six monthly reviews were recommended and each centre giving the progress based on the indicators would make a report.

GIST Technology Workshop

A technology workshop for Resource Centres was organized by C-DAC during the period 17-7-2000 to 21-7-2000. The need for this workshop was suggested during the ZOPP workshop held in Bangalore. Main idea of this workshop was to provide training to all the Resource Centres on the tools and technologies developed by C-DAC, and also to evolve a collaborative development approach among the Resource Centres.

The main topics that were covered are

- Spell checker architecture.
- How to build spell checkers/dictionaries.
- Issues related to Indian language keyboard management.
- Introduction to GIST SDK.
- Using GIST SDK for VC/VB/Web applications.
- Methodology of development of search engines.
- Development and rendering of Indian Language Fonts.
- Development of web applications using iPlugin.

OCR Technology Presentation

On 21-7-2000 Presentations were made on the development of OCR technology for Devnagari script by C-DAC, Pune and ISI, Kolkata. ISI demonstrated the product that was developed. It was felt the ISI product is capable of delivering 98% accuracy and exhibit font invariance. The means of making it a commercial product need to be explored.

Indian Languages Technology Vision 2010

Tentative Guidelines evolved during TDIL WG Meeting on 19th October 2000

Vision statement

Digital Unite and Knowledge for all

Mission Statement

Communicating and Moving up the knowledge chain overcoming Language barrier.

Intermediate Goals

- ▶ Standardization of code, font, keyboard, and Indian Script Roman Code.

- ▶ Coexistence of Unicode and ISCII. Code conversion utilities be available.
- ▶ 3-4 fonts should be standardized and be in public domain. The font and content development should be encouraged.
- ▶ IT Localization Information Services Cell should be setup that should provide information on phone and through e-mail of Indian Language Technology products and services.

Short-term Goals

- Corpus creation (about 10 million words) and analysis: Unilingual corpus, parallel corpus automated tagging and statistical dictionary.
- Web content creation.
- Applications: land records, voters list, reservation chart, cooperative stores, telephone dictionaries on web.
- Language Technology should be integrated into IT curricula and the projects. Curricula for functional languages and linguistics should cover the issues related to language technologies.
- Collaborative development of Indian language lexical resources.
- Development of web enabled courseware.
- Writing aids (Spell checkers, grammar checkers and text summarization utilities).
- Sharing of standardized lexware (written text and speech).
- Development of lexware tools.
- 1-4 weeks training programmes on ILT awareness, lexware development.

Medium-term Goals

- ★ Indian language speech database
- ★ Multilingual, multimedia, content development with semantic indexing, classical and multi font and decorative fonts, off-line/on-line OCR.
- ★ Cross lingual information retrieval (CLIR) tools.
- ★ Human speech encoding
- ★ Speech Engine : Speech recognition, specific speech I/O.
- ★ Indian language support on Internet appliances.
- ★ Understanding and Acquisition of languages, knowledge representation, gisting and interfacing.
- ★ Distinguished achievement awards for M.Tech/MCA/Ph.D. level projects/dissertations in the field of Indian Language Technology.
- ★ Machine aided translation: English to Indian languages, among Indian languages, Indian languages to English and other foreign languages.
- ★ Web based translation.

Long-term Goals

- ◆ Speech to Speech Translation.
- ◆ Human Inspiring Systems.

Readers are requested to give their feedback on Indian Language Technology Vision 2010

E-mail : info@tdil.gov.in