

NEW ORDER OF KNOWLEDGE-BASED SOCIETY: Rise, Raise & Race
Overcoming Language barriers

Dr Om Vikas

Department of Information Technology
Ministry of Communication & Information Technology
6, CGO Complex, New Delhi -110003
Tel/Fax: +91(11) 24363076, Email : omvikas@mit.gov.in

Content

1. Catching up the Knowledge Wave
2. “Digital Divide” as They Behold
3. Pacing up the Digital Unite
 - 3.1 Riding over Individualism into collectivism
 - 3.2 Think globally and Act Locally
 - 3.3 Compete for Excellence & Collaborate for Innovation
 - 3.4 Raise to Rise & Race to Limits.
4. India’s Initiatives as a case study
 - 4.1 Technology Development for Indian Languages
 - 4.1.1 TDIL Mission Programe
 - 4.1.2 ZOPP Workshop for Consensus Building
 - 4.1.3 Industry Interface
 - 4.1.4 Innovation Management of Distributed Projects
 - 4.1.5 Media Lab Asia
 - 4.2 Majore Achievement
 - 4.3 Industry Picks up Indian Languages
 - 4.4 New Initiatives
 - 4.4.1 Speech-to-Speech Translation
 - 4.4.2 Intelligent Cognitive Systems
 - 4.4.3 Entrepreneurship Development
 - 4.4.4 Affordable PC
5. Summing up the Challenges Ahead

Abstract

Industrial revolution spurred the economic growth. Innovations turned into practices. With the advent of personal computers in 1970s, management of information about processes, products and people became important in every sector of economy. Convergence of computing, communication and content resulted into Digital Revolution - digital economy and knowledge based society. Processing power doubles every 18 months; storage doubles every 9 months and the bandwidth triples every 12 months. Thus, increasingly better technology at lower cost. Advanced communities began to use intelligent appliances which would gather information, process it and suggest alternatives for decision making.

Societies having requisite Information & Communication Technology infrastructure advanced rapidly reaping benefits of digital revolution. Other societies lagged behind. This resulted into sprawling “digital divide” which may either ruin the mankind or open up new vistas of opportunities to work together.

In this context it is essential to consider new order of knowledge-based society where collectivism will rein over individualism, local actions will be in global perspective, holistic approach will replace fragmented one; peer competitiveness will shift into collaboration for innovation aiming at attaining excellence. Communities will have to raise the neighbouring ones, and rise together to race for peace and prosperity. India's initiatives towards bridging digital divide in multilingual environment are cited as case study. These include development of technologies overcoming language barriers; their deployment for e-governance, e-health, e-learning applications, and promoting entrepreneurship in ITeS in local languages. To sum up, the knowledge based society, need to collectively follow the maxim of Rise, Raise & Race for sustainable development.

1. Catching up the Knowledge wave

Knowledge defies economic principle of scarcity. Knowledge is not scarce in traditional sense. The more you use it and pass it on, the more it proliferates. It is "infinitely expansible" or "non-rival in consumption". It can be replicated cheaply and consumed over and over again. Knowledge is more difficult to measure than traditional inputs such as steel or labour. Future prosperity of rich economies will depend both on their ability to innovate and on their ability to adjust to change.

20th century brought unprecedented erosion in knowledge of world communities. From an estimated 10,000 world languages in 1900, about 6,700 language surveyed in 2000. By the middle of 21st century, almost all of the world's many ecosystems will be occupied by people who have no indigenous language capable of describing, using, or conserving the diversity that remains. *Two percent of the world's languages are becoming extinct every year.* There is worldwide, unquantifiable erosion of cultural participation, knowledge and innovation. With the loss of a language, we lose art and ideas, scientific information and technological innovation capacity. World-level literacy is improving. More people can read than ever before, but fewer people create stories. We have moved from being **creators** to **consumers** at the time when technology could have amplified our creative capacities.

According to a UNESCO study (1999) of 65 languages for which data was available for both 1980 and 1994, 49 of the languages (75 percent) had experienced real decline in number of works translated from these languages into other languages. The proportion for English arose from 43 percent in 1980 to over 57 percent in 1994. The share held by top four translated languages (English, Spanish, French and German) rose from 65 percent in 1980 to 81 percent in 1994. The UNESCO study also shows that cultural erosion is not confined to the collapse in translated languages. There is also collapse in quality. According to an UNESCO study involving world's 140 most published authors; 90 out of 140 were English writers in 1994 compared to 64 out of 140 in 1980. There is collapse in authorship, translation and quality in other languages.

Industry points to the communication democratisation offered by the Internet, yet around 80 percent of the information on the Internet is in English - even though only 8 percent of the world population speak English as first language. An estimated 85 percent of Internet revenue and 95 percent of Internet stock accrues to USA. [Development Dialogue 1999]

Over 25 Million pages are added per year as research in science & technology. Most of these are in English, few in European languages & Japanese, and negligible in the rest. On Internet, more than 2/3 content is in English alone. This is natural consequence of the fact that English has become lingua franca of Science & Technology, and research is conducted/ published in English speaking advanced countries. This has resulted into "Innovation Divide". Innovation traits set roots through mother tongue. Knowledge can be communicated from one language to another, and can grow

indigenously. Knowledge acquisition, absorption, communication, and generation are key processes in catching up the knowledge wave.

In the 20th century, we had the potential to use technology to liberate creativity and extend cultural participation. Instead, we used the **technologies to curtail participation and to control creativity**. [Development Dialogue 1999: 1-2]

In 1960, the world’s poorest countries (20 percent of world population) accounted for 4 percent of global exports; by 1990 their share slipped to barely 1 percent. Predictions that the ‘poor might not always be with us’ have not come true. By 1998, percentage of absolute poor in the world (income below US \$1 per day) was at 24 percent and the trendline had turned upward. Optimistic forecasts of gains of technology now seem illusory. Are we winning or losing? Is the world losing more knowledge than it is gaining? [Development Dialogue, 1999]

2. "Digital Divide" as They Behold

Analysis of use of world languages on Internet reveals that

- Latin Alphabet users (39%) enjoy 84% of access to the internet; Hanzi-users in CJK (22%) enjoy 13%
- Arabic script users (9%) enjoy 1.2% and
- Brahmi-origin scripts users in south east Asia and Indic Scripts users occupy 22% of world population and they have just 0.3% of internet access.

Digital Divide -- Difference in perceptions

Perception	Developed Countries	Developing Countries
Why discussed ?	Desire to capture larger markets	Fear of lagging behind in economic race
Policy	Information explosion	Localization
Results	Increasing use of English and thrust of western culture.	Preservation of local language and culture.
Consumer nature	“substitute the old”	“Upgrade the Old”
Technology development	IPR-Centric	Open preferred technology
Low cost PC	\$500	less than \$ 40
Reason: PPP : (15:1) GNP : (75:1)	34260 (USA) 24260	2400 (India) 460
Focus	Digital divide, Access to Information, Wider control	Digital Unite, Share the Knowledge, Small is beautiful.

Low affordability means low ICT penetration and larger digital divide.

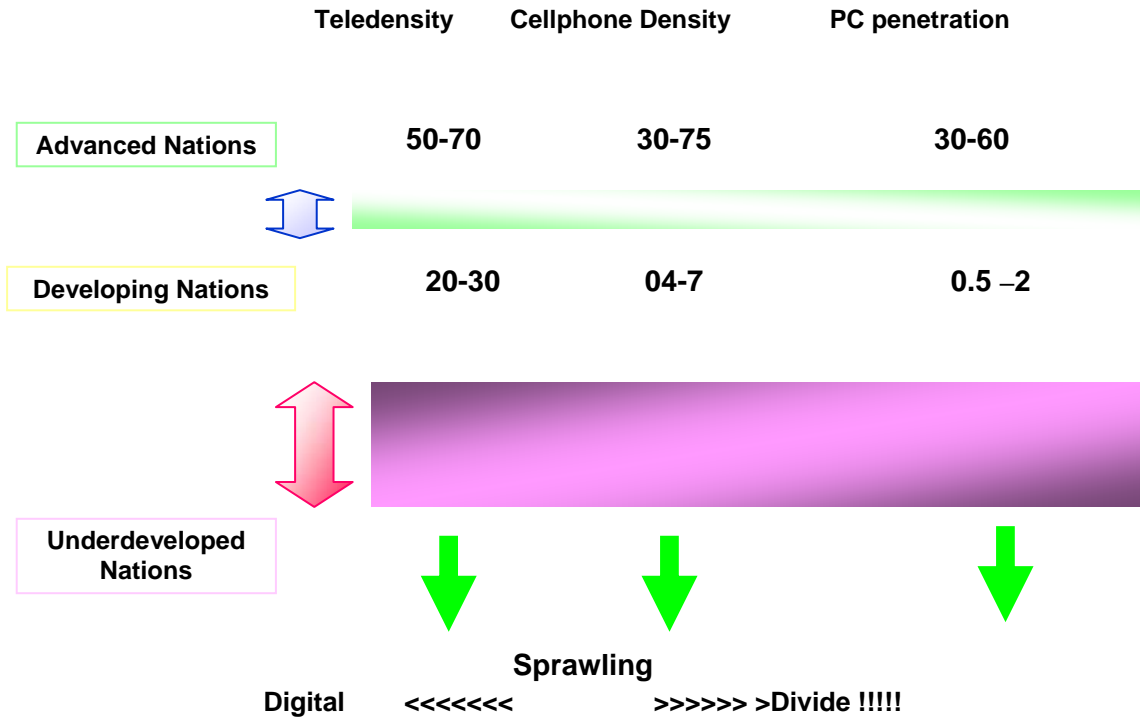
It is, therefore, necessary to evolve world-level policy to arrest the erupting “digital divide” and convert it into “digital unite”.

3. Pacing up the Digital Unite

India is large country with 1 Billion population and a large potential market. India ranks top among countries which harness IT most for economic development; interestingly Digital divide is also highest in India, according to UNESCO report. Comparison of some economic indicators as given below reveals the aspects of relative affordability and adaptability.

Table 1 . Comparison of ICT indicators between various knowledge based economies Source Global Information Technology Report 2001-2003

Country	Population (In Million)	GDP (per capita-PPP) in US \$	GCI (2001-02)	UNDP HDI (2001-02)	Teledensity	PC penetration	%PC connected to Internet	Cellular Phone Density	Internet User/ 100 inhabitants
Finland	5.176	24864	1	10	54.69	39.61	25.82	72.64	37.23
USA	275	33886	2	6	69.97	58.62	50.04	39.79	59.75
UK	59.8	23197	12	13	58.23	33.78	8.31	66.95	33.43
Germany	82.2	24931	17	14	60.11	33.64	7.38	58.58	29.21
France	58.8	24032	20	11	58.01	30.48	6.26	49.40	14.46
Canada	30.8	27783	3	2	67.65	39.02	19.70	28.45	41.30
Japan	127	25796	21	9	58.47	31.52	11.60	52.61	30.44
China	1300	3953	39	60	11.11	1.59	0.34	6.58	1.74
Korea	47.3	17311	23	25	46.36	19.03	4.42	56.69	40.25
Singapore	4.017	23000	4	24	48.44	48.31	9.06	68.38	46.05
Thailand	60.6	6469	33	49	8.66	2.43	4.31	5.04	1.98
Malaysia	23.30	8924	30	43	19.92	10.31	2.84	21.31	15.90
Taiwan	22.30	17223	7	NA	56.80	22.46	21.91	80.30	28.13
India	1010	2403	57(*48 in02-03)	71	3.20	0.45	0.78	0.35	0.9



New Order of Knowledge-based society will focus on: Collectivism , Localisation in globalisation, Compete for excellence, Raise to Rise, Race to Limits and Holistic Solution

3.1 Riding over individualism into collectivism

Hofstede provided individualism-collectivism scores for more than 50 countries in 1970s with average of 43 over the scores range from 6 to 91. High scores are for individualistic cultures while low scores are associated with collectivistic cultures”. USA has a score of 20 (highly collectivistic). In highly individualistic society, ties between people are loose and people focus on their own needs. People tend to put personal interests above team interests. People take their actions independently of what others think and base their self-understanding on these personal actions.

In the collectivistic society, people are integrated into cohesive groups , Which take care of them in exchange for unquestioning loyalty. People tend to put team interests above personal interests. They base their self- understanding on how others around them react because they consider themselves members of cohesive groups.

Collectivism: Sah veeryam karvaavahai (in sanskrit, it means Let us work together) is desirable to work out holistic solution to a real-life complex problem. Attitude of collectivism will make us sensitive to neighboring communities, nations and to the whole world. TEAM may also connote **T**ogether **E**veryone **A**chieves **M**ore. Team catalyses individual innovation, that magnifies into collectively directed achievements.

3.2. Think Globally and Act Locally

World is transforming into networked societies. Distances are shrinking. Concept of global village is being propounded. In the multi-cultural world, it is a challenge. Networking facilitates access to knowledge anywhere, anytime and thus enables to have global perspective of the problem and the associated solution. Hence enact localisation within globalisation. Localisation ensures relevance, efficacy and people's participation. Knowledge generated anywhere is influenced by local culture, that is, language, look & feel, beliefs and values. Technology is not a neutral agent, rather the technology reflects cultural specific beliefs and values of the place where it was developed. Technological culturisation in the process of localisation, ensures greater acceptance of a new technology.

3.3. Compete for Excellence & Collaborate for Innovation

Competitiveness is the slogan for success in business. Peer competition compels for hiding critical knowledge; focuses on increasing market-share. This approach will retard world-level progress of knowledge based society. In the context of globalisation, IPR-centric competitiveness between developed and underdeveloped nations needs to be discouraged. They should rather cooperate for innovations and compete for achieving excellence to suite to local environment.

Innovation is the most important source of competitive advantage in advanced economies. Building up innovative capacity has strong relationship to a country's overall competitiveness and level of prosperity. Countries vary significantly in their historical ability to produce global innovation. Those countries, such as Finland and Taiwan have proactively built innovative capability, have prospered.

Sarve Bhavantu Sukhinah(in Sanskrit), means 'happiness for all', should be our maxim. For example, India's innovation will include innovation in advanced technology adaptation and localisation during the *catch-up phase*. And innovation for newer products and services in the *competing phase*. We are in knowledge economy but our managerial and governance systems are stuck in the industrial era. It's time for a whole new model, and history offers prototypes based on individual freedom and collective action. It is essential to proactively build innovation capabilities.

An indicative comparison of technological innovations and diffusion is given in the following table:

Table : Technological Innovations & Diffusion
(Ref : The Global Competitiveness report 2001-02)

Indicator (7=best)	India	USA	Japan	Israel	Finland	China
Technology Sophistication	4.5	6.8	6.4	6.6	6.6	3.7
Quality of science Res. inst.	5.2	6.7	5.7	6.2	6.3	4.5
Univ.-Industry Res. Collaboration	3.7	5.3	4.4	5.5	6.1	4.2
Government Procurement of Advanced Technology products	3.8	4.6	4.8	4.5	5.3	4.6
Availability of Scientist/Engineers	6.4	6.2	6.1	6.4	6.4	4.4
# Our Guess #						
* Scope for Technology Adaptation	6.5	1	4	2	4	5.5

3.4. Raise to Rise & Race to Limits

Globalisation is a natural consequence of better telecommunication infrastructure and affordable & user friendly information appliances. Liberalisation is advice of advanced nations to the rest for creating conducive environment for technology acquisition and absorption and thus expanding their market. Mindset needs to be changed to help the underdeveloped nations to catch-up in technology absorption and participation in knowledge generation.

Following is an example of providing high-tech solution in low-tech environment. A group of engineer volunteers in USA designed and built a rugged and low-cost bicycle-powered computer and wireless network for villagers of phon kham in Laos which had no electricity or phone service. There was no way to call relatives living abroad or even in the next town. This is a project to bridge the digital divide. The machine had to be dirt cheap (the annual per capita income in Laos is US\$241), be easy for villagers to maintain and operate in their own language, draw power less than 20W, and withstand Laos' blistering heat (40C in the summer) and drenching monsoons. Hence the volunteers opted for ultra low-power Mach Z PC-on-a-Chip with IEEE 802.11b wireless LAN Card and Quicknet Voice-over-IP Card; a sturdy dot-matrix printer (with re-inkable ribbon), Lao-language Keyboard, a track ball, and Laonux – the localized Linux OS with Lao language support. [IEEE Spectrum, September 2003].

In India Community Information Centres (CICs) in 487 blocks in seven North-Eastern states and Sikkim have been established to reduce the digital divide by providing internet access and IT enabled services to the community and to facilitate citizen interface with government. The services provided by CIC are :

- Access to socio-economic databases
- E-learning , E-education
- E-medicine , E-consulting
- Citizen Centric services
- Weather Information
- IT awareness among the people

Innovation follows on Stretching our imagination to limits. As we noticed that constrained environment of a village in Lao led development of new operating system, cycle-powered PC, etc. Heterogeneity of communities opens up new opportunities for innovation and integration skills. Time is critical factor in the context of ICT. Let all the communities the world over catch up to the basic technology absorption capability and use it for improving quality of life of the people at large.

4. India's Initiatives as a case study

India is a multi-lingual multi-script country with officially recognized 18 Indian Languages and 10 scripts. Hindi in Devanagari script is the national language. 18 Indian languages include Hindi, Marathi, Konkani, Sanskrit, Nepali, Gujarati, Punjabi, Oriya, Bengali, Assamese, Manipuri, Kannada, Telugu, Tamil, Malayalam, Urdu, Sindhi & Kashmere. Indian scripts may look different in shapes, but they follow similar alphabetic order. Script grammar is also similar. Alphabet consists of vowels and consonants. They are ordered on the basis of phonetic utterances. *What you write what you speak.* Pronunciation of a word is the concatenated string of pronunciation at letter-level. Vowels and consonants have distinct shapes. Pure consonant is a virtual consonant without vowel sound. When vowel follows the (Pure) consonant its modified shape may attach on top, on side or on bottom around the consonant. This modified vowel-grapheme is called MATRA or vowel modifier. Consonants can combine themselves. Characteristics of Indian Languages may be summarised as below:

- What You Speak Is What You Write (WYSIWYW)
- Script grammar describes transformation rules
- Relatively word-order-free
- Common phonetic based alphabet
- Common concept terms (from Sanskrit)

4.1 Technology Development for Indian Languages

India was aware of the technological changes and the local constraints. Development of Language Technology in India may be categorized in three phases:

- 1976-1990 : **A-Technology Phase**
Focus was on **Adaptation Technologies**; abstraction of requisite technological designs and competence building in R&D institutions.
- 1991-2000 : **B-Technology Phase**
Focus was on developing **Basic Technologies**- generic information processing tools, interface technologies and cross-compatibility conversion utilities. TDIL(Technology Development for Indian Languages) programme was initiated.
- 2001-2010 : **C-Technology Phase**

Focus is on developing Creative Technologies in the context of convergence of computing, communication and content technologies. Collaborative technology development is being encouraged to realise.

Government spending during FY 1991- FY 2000 was about US\$ 3 Million. During 2001-2004 government spending for Indian language technology development projects is about US\$ 5 Million.

4.1.1 TDIL Mission Program

Under the TDIL Mission-mode program, 13 Resource Centers were set-up covering all 18 languages and 7 centers to spread IT penetration in economically backward States, and some projects for developing specialized technologies and products. Third party evaluation of language technologies thus developed is done by STQC.

Resource Centres for Indian Language Technology Solutions are IIT Kanpur, IIT Mumbai, IIT Guwahati, Anna Univ., Univ of Hyderabad, C-DAC Noida, C-DAC Thirunanthpuram, C-DCA Pune, MS Univ. Varodara, TIET Patiala, JNU Delhi., Utkal Univ./OCAC Bhuvneshwar and ISI Kolkatta.

The CoIL Net Centres for Content Development and IT Loccasation are in the economically backward Hindi speaking states. These include BHU (UP), IIT Roorkee (Utranchal), Bansasthali Vidyapeeth (Rajsthan), IIITM Gwalior (Madhya Pradesh), BIT Ranchi (Jharkhand), Govt, Polytechnic Patna (Bihar), Engineering Collage (Chhattisgarh) and the support centers at C-DAC Pune. IIT Kanpur & IGNCA Delhi.

TDIL Mission mode program was initiated during 1999-2000 with

Vision statement : *Digital unite and knowledge for all; and*

Mission statement : *Communicating without language barrier & moving up the knowledge chain.*

Seven focused areas are: *Knowledge Resources, Knowledge Tools, Translation Support Systems, Human Machine Interface System, Localization, Language Technology HRD, and Standardization*

Long Term Goals include: *Speech to Speech translation, and Human Inspiring systems*

4.1.2 ZOPP Workshop for Consensus Building

During the second Zopp workshop held in March 2002. Discussion points included:

- Expectations : networking, sharing, goals, collaboration with industry, institutional issues
- Killer Applications : Parallel Corpora (text & speech), OCR, MAT

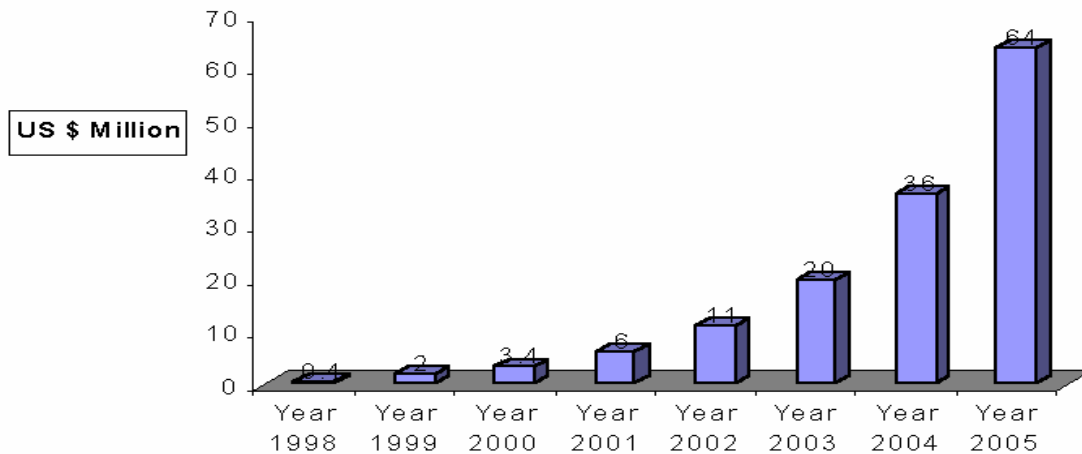
- Tools : dictionary, Spell checker, Morph Analyser, Fonts & Conversion Utilities
- Future Directions
 - Generic applications with open architecture, interoperability, flexibility & scalability (all Resource Centres to develop)
 - Cross Lingual Information Retrieval (top 5 world languages e.g. Chinese, Spanish, Russian, Japanese, English, and Indian languages)
 - Concept based Indian Networking Language (INL)
 - Speech_to_Speech Translation

4.1.3 Industry Interface : *COIL-Tech*

TDIL Programme is closely working with *COIL-Tech* (Consortium for Innovation & Language Technology) which interfaces during development Process, jointly works on technology incubation and commercialization.

Findings of the survey by Forst & Sullivan 2003 on Local Language Software Market revenue.

Local Language Software Revenue



Language Technology Business Meet is organized to showcase prototype technologies and facilitate dialogue between LT developers in academia and the industry for possible transfer of technology or collaborative development aiming at productization. 43 Technology Handshakes were signed during the LTBM, November 2001.

4.1.4 Innovation Management of Distributed Multilingual Projects

Technology Innovation Audit of the sponsored projects is essential in order to promote standardization and sharing of technologies. Audit steps may include:

- Concept, Design and Implementation audit

- Alpha Testing with Peer Developers
- Beta Testing with a small number of potential users
- Certification of IL Software (IS:14639-1998 standard for software evaluation)

Peer review of the projects and enforcing Beta testing of products or services yield satisfactory results; Culture of collaborative technology development is also strengthened.

Technology management focuses on *Consolidation, and Integrating Innovations into Products/ Services*

Public Domain/General Public License (GPL) approach is encouraged for faster development and rapid spread. IT localization clinics promote wider dissemination and organize internship training.

Bilateral/International cooperation in Language Technology and Applications will be encouraged to share knowledge and enable the neighbouring nations to catch to the state of technology absorption and participation.

Academic institutions are good in research and technology development. They are often averse to productising the technology. They prefer publishing papers. They are often reluctant to share their codes. Mechanism for IPRising their ideas and products is also not very conducive. University rules vary. Researchers in academia want to go for that last 2% of performance, but they need to be reminded that it is better to get *sufficient* solution out fast and then continue to enhance it. Test and evaluation of their technologies is difficult due to lack of documentation and non-adherence to industry practices. Hence we, in the TDIL mission HQ, had to devise strategies for innovation management, consensus building for collaborative development, third party test & evaluation, time targets for integrating technologies, launch of products, comprehensive information dissemination, standardisation, peer-competition, industry-academia interaction, and supporting virtual R&D units under the industry consortium. Open source software approach is carefully pursued. Software, placed in public domain, does not have much meaning if there is no measure of performance mentioned along with. There is need to properly coordinate various efforts relating to evolving standards, building up standards database, peer-review, benchmarking constantly upgrading and integrating into larger complex systems.

STQC (Standardisation, Testing and Quality Control) under Ministry of Communications & Information Technology have been entrusted with the task of third party evaluation of language technologies/products. This will ensure claims with performance measure and certification of language technologies in the market. This will enforce discipline on the part of technology developers, facilitate dialogue with industry for commercialisation.

4.1.5 Media Lab Asia

The goal of this collaboration is to bring the benefits of innovation and technology to the common man. Towards this end, a network of research labs has been

established on the campuses of the IITs at Mumbai, Delhi, Chennai, Kanpur & Kharagpur under 4 broad categories include :

World Computer (Lowcost PC)

Rural Operating Systems; Speech Interfaces For Local Dialects; Visual Language; Interfaces for All; Interlingua Web; Multi-Literate Interface; Literacy Learning Through Pictures

Bits for All (Universal Connectivity)

Rural WiFi, DakNet, Digital Gangetic Plain, Off-Line Internet Access, Rural VoIP

Tomorrow's Tools (Language Interfaces)

Mapping For the Masses, Community Access to Sustainable Health (Ca:sh), Building Robots Creating Science (BRICS), Digital Craft Revival, Digital Human Body, Digital Music, InfoSculpture, Suchik, Polysensors, Complex RF Impedance Analyzers, UV-VIS Spectrometer, Power Sensors, Think Cycle

Digital Village (Consolidation in delivering value to the masses)

Sustainable Access in Rural India, Community Connection, Digital Mandi, InfoThela

4.2 Major Achievements in ILT

▪ Translation Support Systems (TSS)

- English to Hindi (Angla-Bharati) <http://anglahindi.iitk.ac.in>. (very satisfactory above 85% consistently okay)
- Indian Languages to Hindi (In the process of development)
- Hindi to English (In the process of development)

▪ Human Machine interface Systems

- **Optical Character Recognition (OCR)**
(accuracy 7ILs viz. Hindi Marathi, Bangla, Tamil, Telugu, Gurumukhi, Malayalam, above 97% OCRs in other IL s are in the process of development)
- **Text to Speech system (TTS):** (Hindi, Bangla,)
- **Continuous Speech Recognition: CSR** (Hindi)
- **Translingual Reading Machine for blinds** (OCR+MT+TTS : E - H)

▪ Knowledge Resources

- **Bilingual dictionaries:** Between English and Indian languages tri-lingual E-H-IL dictionaries are under development.
- **Parallel Corpora** – One Million page Parallel Corpora is under [600 Thousand pages ready; texts in 11 Indian Languages]

- **Localization of LINUX systems**

INDIX, Localized LINUX operating system, has been developed to support major Indian languages

- **Knowledge Tools**

Morph Analyzer, Syntactic Analyzer, Spell checker, Messaging system ,
Authoring Systems, Word processors, code conversion utilities have been developed.

- **Standardization**

DIT is the voting member of the Unicode Consortium. Proposed changes in the existing Unicode Standards have been finalized in consultation with respective State Government and Indian IT Industry and presented in the **UNICODE Technical committees**.

Indian Scripts Font Code (INSFOC) Standards have been developed.

Indian Script to Romanization Tables (INSROT)are ready.

- **Information Dissemination:**

TDIL Web-site <http://tdil.mit.gov.in> This Web Site contains information for various TDIL activities, achievements and provides access to a variety of content and free downloads in Indian Language. **Free Downloads include:** Indian Language keyboard driver & fonts and other tools, corpus, content, conversion utilities, dictionaries Machine aided Translation systems, language learning software etc.

4.3 Industry Picks up Indian Languages

Web Sites Supporting Indian Languages

- Web Dunia: www.epatra.com supports 11 languages
- Mithi.com : www.mailjol.com supports 12 languages
- Langoo: www.langoo.com supports 12 languages
- C-DAC: www.cdacindia.com with multilingual support

MNCs Support their products with Indian Language Supports, For Instance,

- Microsoft supports indic scripts on Windows 2000 and office XP, Windows 2003
- Oracle 8i & 9i RDBMS and Lotus support major Indian languages
- Star Office 7.0 also supports major Indian languages on Sun OS, Windows and Linux environment
- Google search engine provides search in Hindi for Unicode-compliant web-enabled content.

4.4 New Initiatives

4.4.1 **Speech_to_Speech Translation**

2005: Systems:

- Useful speech summarization systems in major Indian languages
- Useful text summarization systems (100:10:1)

- Useful multiple - speaker meeting transcription systems
- Medium-size vocabulary (3,000 W) conversational systems
- Speech driven personal assistant systems

2010: Basic Technologies expected:

- Unlimited-vocabulary spoken multilingual conversation
- Unlimited-vocabulary spoken translation systems
- Unlimited on-line understanding & generation of integrated natural speech, lips, facial expression and gesture communication
- Fully natural interactive communication

4.4.2 Intelligent Cognitive System

Knowledge Understanding & Acquisition of Languages, Interferencing and Interpretation (KUNDALINI)

Broad Objectives:

- To develop knowledge frameworks and access mechanisms based on Indian tradition.
- To develop methodologies and tools for knowledge representation, extraction, mining, gisting, inferencing and interpretation.
- To e-Content knowledge in Shastric Sanskrita and Indian heritage
- To develop Sanskrit based Networking Language as Machine Translation Interlingua.(Concept based Indian Networking Languages: *INL*)
- To develop cognitive models for **human inspiring systems** in the networked society.
- To promote partnership with agencies in other countries for all collaborative development.

4.4.3 Entrepreneurship Development

Gyanaudyog : ~~गुणवत्~~ (Knowledge-based Enterprise Entrepreneurship Development.) Program is envisaged to promote Small Office & Home Entrepreneurship in the area of Information Technology for catalyzing IT enabled services (ITES), especially for women empowerment.

Entrepreneurship development areas may include : Content Creation, PPT/Web-page design, e-Tutoring, Computer Aided Design, Software Localization, Remote Customer Interaction Services, Translation / Transcreation

The Coordinating center must ensure three techno- business supports:

- *Technology mentoring*
- *Financial Support Guidance*
- *Marketing Information.*

4.4.4 Affordable PC

SOHE – GANAK BHARATI program is envisaged to develop Affordable Computing System for Small Office, Home and Edutainment

- Ganak_Bharati will be a specially designed system having user interface, input & output completely in an Indian Language, standard monitor for display, and imprinted keyboard
- The operating system of Ganak_Bharati will be in public domain
- It would have interfaces available in all Indian scripts, integrated in the same version, and users can switch from one Indian script interface to another easily.

5. Summing up the Challenges Ahead

- **ML Open Source Software**
 - Shareable Software
 - Standards database and updating
 - Support service & Help line
 - Consortium approach
 - GPL with performance index
- **Benchmarking & Standards**
 - testing against international standards
 - active participation in evolving standards
- **Information Technology Culture**
 - **Awareness** : IT Clinic, Workshops, media
 - **BIPK** (Basic information Processing Kit) with user friendly, easy-to-use, affordable, scalable, interoperable and re-usable tools. BIPK may consist star office like processing facility, fonts, KB driver, spell checker, dictionary and conversion utility.
 - **Entrepreneurship** : Gyanaudyog workshops.
- **Localization**
 - Localization of software and content into local languages
 - Enlarging share in localization outsourcing (\$ 8 Bn By 2006:IDC)
- **International Collaboration in Language Informatics.**
 - Industry - academia cooperation in joint research & technology development projects.
 - Exchange of faculty and students
 - HRD programs in knowledge Engineering & Computational Linguistics

- **Rise, Raise & Race**
 - Possess basic language technologies
 - Promote Collectivistic Culture
 - Think globally & act locally
 - Collaborate for innovation
 - Speed up seamless universal access

References

1. Om Vikas, **Annals of Indian Language Computing**, international conference on Universal knowledge & Language-2002, at Goa.
2. **Development Dialogue**, 1999 : 1-2, Dag Hammarskjold centre
3. **World Culture Report** – Culture, Creativity and Markets, UNESCO, 1998.
4. G.Hofstede, **Culture and Organizations Software of the Mind**, McGraw Hill, New York, 1997.
5. **World Culture Diversity**, UNESCO 1995
6. **Vishwabharat @ tdil**, issue10, July 2003, published by Dept of Information Technology, Govt of India.
7. <http://tdil.mit.gov.in> contains updated information on language technologies in India.
8. Ashton Applewhite, “IT Takes a Village”, **IEEE Spectrum**, September 2003