



Local Language Computing Policy Initiatives

Country Paper: Islamic Republic of Pakistan

Presented By: Mansoor Jafri

Thimpu, Bhutan

January 22 – 24, 2007

Background: Languages in Pakistan

- ▶ “URDU” is the designated official and national language of Pakistan
- ▶ Urdu is spoken and understood by more than 75% of the population of ~ 140 million
- ▶ Urdu was officially planned to replace English as the official language by 1988 – today, this is still the official plan
- ▶ 92 % of the population does not speak English
- ▶ Out of 69 Listed Languages, 58 are listed Major Languages

Background: Literacy Rates in Local Languages

- ▶ Overall literacy rate in Pakistan is $\sim 44\%$

Gender	Total	Urban	Rural
Both	44	63	34
Male	55	70	46
Female	32	55	20

- Provincial Literacy:
 - Punjab 47%, Sindh 45%, NWFP 35%, Balochistan 25%

- Newspapers & Periodicals

Language	Total No.	Circulation
All	1997	6,527,000
Urdu	1533	5,236,000
Punjabi	28	34,000
Pushto	35	31,000
Balochi	17	12,000
Sindhi	80	569,000
English	290	637,000

Background: IT Infrastructure

- ▶ Internet Service Providers ~ 150
- ▶ No. of Internet Users ~ over 2.2 million
- ▶ Tele-density ~ 11%
- ▶ Length of Fiber Optic ~ 5500 km
- ▶ Internet at dialup @ ~ Rs.15/hour

Local Language Computing Policy

▶ 1973 Constitution:

- Urdu is the National Language. By 1988, Government processing to be in Urdu
- Provinces may adopt other Provincial languages in addition to the National Language

▶ National IT Policy

- Standardized Urdu code plate will be launched and a concerted plan to encourage the development of open source and licensable Urdu software would be undertaken. This will enable plug-ins for popular office and e-mail packages to be made available. This initiative is expected to drive the development of other Urdu and Regional software packages for word processing and data base applications.
- The government will encourage the use of open source operating systems and low cost or free English language Office software for normal operations.
- The intent of this initiative is to encourage people to develop skills in working and writing core software for applications and developing tools, which will go beyond the development of the local languages. The application programs for translation, speech to text conversion, databases, ASPs for popular packages will need to be written in currently and newly evolving software

Local Language Computing Policy

(cont)

► National Language Authority (est. 1979)

- To consider ways and means for the promotion of Urdu as the National Language of Pakistan and to make all necessary arrangements in this regard to submit recommendations to the Prime Minister for expediting the use of the National Language.
- To facilitate introduction of Urdu as the Official language throughout the country by developing dictionaries and other reading materials for in-service training of the personnel working in Government/ Semi-Government Offices, Courts and other Institutions.
- To coordinate the work of all Urdu.
- To facilitate the adoption of Urdu as the language to be used in all Competitive Examinations, in cooperation with the Federal and Provincial Public Service Commissions Development Boards.
- To carry out such other responsibilities as the Prime Minister may entrust to the Authority in regard to the promotion of the National Language.

Local Language Computing Policy

(cont)

- ▶ NLA Center of Excellence for Urdu Informatics (est. 1998)
 - Construction and Development of Urdu Informatics Center
 - To help academic institutes and universities establish Urdu Informatics Departments and provide support from Urdu Informatics Center. (e.g. Federal Urdu University to establish the first pilot Urdu Informatics Department)
 - To provide the GOP with option in adoption of Urdu as official, judicial and instructional language of choice.
 - To conduct research and development in all matters relating to Urdu standardizations and have declared as Center of Excellence and ultimate authority for all matters relating to Urdu Language

Relevant Government Ministries and Departments

- ▶ Ministry of Education
- ▶ Ministry of Information & Broadcasting
- ▶ Ministry of Information Technology
 - E-Government Directorate
 - Pakistan Software Export Board
 - Pakistan Computer Bureau
- ▶ Ministry of Science and Technology
- ▶ National Language Authority
- ▶ Ministry of Social Welfare and Special Education
- ▶ Other Related Ministries
 - Ministry of Interior
 - Ministry of Labor, Manpower & Overseas Pakistanis
 - Ministry of Law, Justice & Human Rights, Islamabad
 - Ministry of Local Government and Rural Development
 - Ministry of Population Welfare
 - Ministry of Tourism, Islamabad
 - Ministry of Women Development

Other Relevant Organizations

- ▶ National Database and Registration Authority (NADRA) Islamabad.
- ▶ Jang Group, Mashriq, Nawa-I-Waqat
- ▶ All Urdu Publishing Companies of Pakistan
- ▶ All Universities of Pakistan
- ▶ All Islamic Madrasas of Pakistan

Local Language Computing Policy Initiatives

- ▶ **Project Title:** Nafees Nasta'leeq font project (funded by IDRC, APDIP UNDP and APNIC) and Urdu Machine Translation, Text-to-Speech system and Lexicon systems through a three year 30 person project funded by E-Government Directorate of Government of Pakistan.
- ▶ **Company Name:** E-Government Project Ministry of IT, UNDP, APNIC, APDIP.
- ▶ **Web Address:** <http://www.pan10n.net>
- ▶ **Description:** PAN localization provided technical assistance for the development of Nafees Nasta'leeq font project (funded by IDRC, APDIP UNDP and APNIC). Also they are currently working on the development of Urdu Machine Translation, Text-to-Speech system and Lexicon systems through a three year 30 person project funded by E-Government Directorate of Government of Pakistan. Currently they are working on the Urdu Spell-Checker project and advising on Urdu Terminology Translation project for Microsoft, USA.

Local Language Computing Policy Initiatives

- ▶ **Project Title:** Computerized National Identity Card and Passport
- ▶ **Dated from and To:** March 10, 2000
- ▶ **Company Name:** National Database and Registration Authority
- ▶ **Web Address:** <http://www.nadra.gov.pk>
- ▶ **Description:** The National Database Organization (NDO) was established under the Ministry of Interior as a sister organization to undertake the task of handling the data being collected through National Data Forms during the Population Census in 1998. The National Data Forms (NDFs) were designed by a committee of experts encompassing maximum social and fiscal indicators as implements for future planning, documentation of the economy, and for creation of a comprehensive Citizens' Database. To cater to a organized and joined approach in both registration and social fields, NADRA was formed on 10th March 2000 by merging the National Database Organization (NDO) and the Directorate General of Registration (DGR). It is an independent corporate body with requisite autonomy.

Local Language Computing Policy Initiatives

- ▶ National Language Authority Government of Pakistan has introduced a Microsoft Office 2003 Urdu Interface on 23rd Dec. 2005, (Office 2003 Edition اردو مواجہ پیک) and also Urdu Key board, Pak Nastaleekh font, and alphabets.
- ▶ One of the major localization initiatives has been the National ID cards project promoted by the Ministry of Interior in the premises of the HQ NADRA Islamabad, which is working successfully on national as well as international level.
- ▶ PAN achievements included on Nafees Nasta'leeq font project funded by IDRC, APDIP UNDP and APNIC.
- ▶ E-Government Directorate (Ministry of IT) has completed Urdu Machine Translation, Text-to-Speech system and Lexicon systems through a three year 30 person project with the collaboration of PAN Localization Agency.
- ▶ Sarmad H and Afzal M (2001) have defined the procedure of Urdu Computing Standards "Urdu Zabata Takhti" in Pakistan.
- ▶ Introducing of different English to Urdu Translators like AGHAZ, POS (Parts of Speech), and MT (Machine Translation).
- ▶ Establishment of different softwares like Shahkar, Surkhab, Nastaliq, Nizami and In Page etc.

NLA Achievements

- ▶ Urdu as a medium of expression in the examinations held by Federal Public Service Commission and Provincial Public Service Commissions. Translated approximately 55,000 pages of official documents, including Government Rules and Regulations, and various Reports to facilitate the introduction of Urdu in offices, courts and educational and other institutions, the following material has been prepared and published;
- ▶ No of publications 550 titles including:

- Dictionaries (General & Technical) 80
 - Text books 80
 - Qaumi English-Urdu Dictionary (Based on Webster's Dictionary)
 - Dictionary for Children
 - Dictionary of Official Terms and Phrases
 - Dictionary of Scientific and Technological Terms
 - Dictionary of Legal Terms(Based on Black's Law Dictionary)
 - Dictionary of Educational Terms
 - Dictionary of Urdu Pronunciation

- Dictionary on Economic and Financial terms
 - Dictionary on Law
 - Dictionary on Political science
 - Dictionary on Library Science
 - Dictionary on Criticism
 - Dictionary on Psychology
 - Dictionary on Economics
 - Dictionary on Philosophy
 - Dictionary on Scientific & technical terms

NLA Achievements

- ▶ Reviewed and revised official dictionary of “Majlis-e-Zaban-e-Daftari”, comprising 35,000 terms. Prepared and published Urdu Thesaurus, comprising 1,000 subjects/groups of words
- ▶ **Training Material**
 - Books on Official Urdu
 - Books on Official Correspondence in Urdu
 - Books on Urdu Typing and Shorthand
- ▶ **Translation Project of Great Academic Books**
- ▶ Training of Federal Officers in Urdu Correspondence
 - Over 6000 Government officers completed official Urdu correspondence courses through the Allama Iqbal Open University

Fundamental Questions

- ▶ How many languages to localize in?
 - Preserving language/culture VS Freezing progression/evolution for
- ▶ Urbanization VS Rural Development
 - Urbanization also leads to “mono culture”
- ▶ Demographics – declining growth rates per official population policy
 - Urban population more dens
- ▶ Technology and Knowledge Independence VS Interdependence VS Dependency
- ▶ Economic Development via Access to Foreign Markets
 - Know your customer (know English and other cultures)
- ▶ Intellectual Property Rights Issues
 - Pro-developing countries IPR

Challenges

- ▶ Lack of realization and Will for localization
- ▶ Lack of organizational and individual “critical mass” for rapid progress on localization
- ▶ Lack of free open source platform for local language projects
- ▶ Lack of electronic content in local language
- ▶ Lack of education
- ▶ Lack of pervasive ICT infrastructure

Future Plans

- ▶ Identify localization efforts, organizations, stakeholders and create a national platform for knowledge and experience sharing
- ▶ Advance national policy on localization based answers to fundamental questions and challenges
- ▶ Target “Urdu” language for localization as:
 - Urdu provides highest readiness for localization (mandate, expertise and content)
 - High impact (greatest population, greatest area coverage, largest domain coverage)
 - Good momentum to launch off from
 - Identify a few high-impact target domains/areas (e.g. agriculture)

Future Plans

- ▶ Refine and develop existing technology and knowledge for localization
 - Dictionaries, lexicon, machine translation, text-to-speech
- ▶ Promote open source more actively and build local expertise
 - Promote Linux and open applications' localization
- ▶ Fund research for Urdu content over mobile devices (voice and text)
- ▶ Continue ICT Infrastructure Development with special focus on ICT4D
- ▶ Develop Special Programs for other Group region-level languages (e.g., Arabic, Chinese, etc)
- ▶ Build framework for dialogue on localized IPR policies

Thank You

Questions/Information: mjafri@moitt.gov.pk