

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Language Computing Policy in Pakistan



Sarmad Hussain

Center for Research in Urdu Language Processing
National University of Computer and Emerging Sciences
Pakistan

sarmad.hussain@nu.edu.pk

Language and Policy

- 60+ languages spoken in Pakistan
 - Indo-Aryan
 - Indo-Iranian
 - Tibeto-Burman
 - Dravidian
- All except one (Gujrati) use Arabic Writing System
- Urdu is the national language
 - Lower courts, police FIR, land records, in Urdu
- English is the official language
 - Higher courts, govt. notifications, govt. communication, in

English

Language and Policy

Punjabi	60,647,207
Sindhi	18,500,000
Seraiki	13,843,106
Urdu	10,719,000
Pashto	18,863,716
Balochi	5,681,000
Brahui	2,000,000
Hindko	2,500,000
Farsi	1,000,000

Language and Policy

- Urdu not the most spoken language but second language for most people, as medium of instruction in schools is Urdu
 - Punjab – Punjabi, Seraiki, Hinko
 - Off: English
 - School: Urdu
 - Sindh – Sindhi
 - Off: English
 - School: Sindhi, Urdu
 - Sarhad – Pashto
 - Off: English
 - School: Urdu
 - Balochistan – Balochi
 - Off: English
 - School: Urdu

Language and Policy

- Urdu was to be adopted by Govt. by 1988 (Constitution of 1973)
- National Language Authority (est. 1979)
 - To promote Urdu as the National Language and prepare Urdu to be adopted as an official language by developing terminology dictionaries and training material
- Urdu Science Board
 - To translate scientific material into Urdu
- Other Authorities and Academies
 - Sindhi Language Authority
 - Seraiki Academy
 - Punjab Language and Culture Center
 - Pashto Academy
 - Balochi Academy

Language, IT and Policy

- Language not specifically addressed in IT Policy in 2000 but realization to promote Urdu computing – many relevant organizations and projects
 - National Database and Registration Authority (NADRA) launched in 2000 to develop Urdu ID Cards for all Pakistani Citizens
 - Center of Excellence in Urdu Informatics at National Language Authority in 2000+
 - Urdu Lexicon Project- E-Govt. to develop Urdu MT, Urdu TTS and Urdu Lexicon
 - Land Record System in Punjab commissioned
- Current Focus on Language Content
 - Expert group to devise content development projects through MoIT initiative called NIDU

Language, IT and Policy

- Much greater role in next IT policy, currently being drafted:
 - Review current situation and quality of I.T. Education in Pakistan; identify future areas that need to be focused, in terms of human resources skills and specializations; and outline a roadmap for spreading I.T. education in Pakistan to reduce the digital divide in line with future trend
 - Review current situation of using Urdu language in I.T.-centric processes and functions; identify impediments in common availability of Urdu-based software applications and automated services; and recommendations for promotion of Urdu language in ICT

Some Related Initiatives

- Higher Education Commission Scholarship Program over past 3 years
 - 3000 International Scholarships
 - 5000 Indigenous Scholarship
- Punjab ICT Literacy Program
 - 15 Computer Labs in 4,286 High Schools of Punjab
- Land Record System
- Center for Excellence in Computational Linguistics
- Language Processing Society of Pakistan
- Regular biennial conference on Language and Technology

Some Continuing Challenges

- Too much focus on Urdu
 - Other Languages of Pakistan?
- Not clear policy on which organization owns the standardization process
 - NLA
 - MoIT
 - PSQCA
- Need more R&D groups in Pakistan in language technology
- Need stronger IPR regime to protect translation of research into development
- Need to translated content oriented policy to action. Not clear how that is possible in a participatory way

Thank you

Questions