

Language Table Sub-Committee Meeting on IDNs for پاکستان .ccTLD

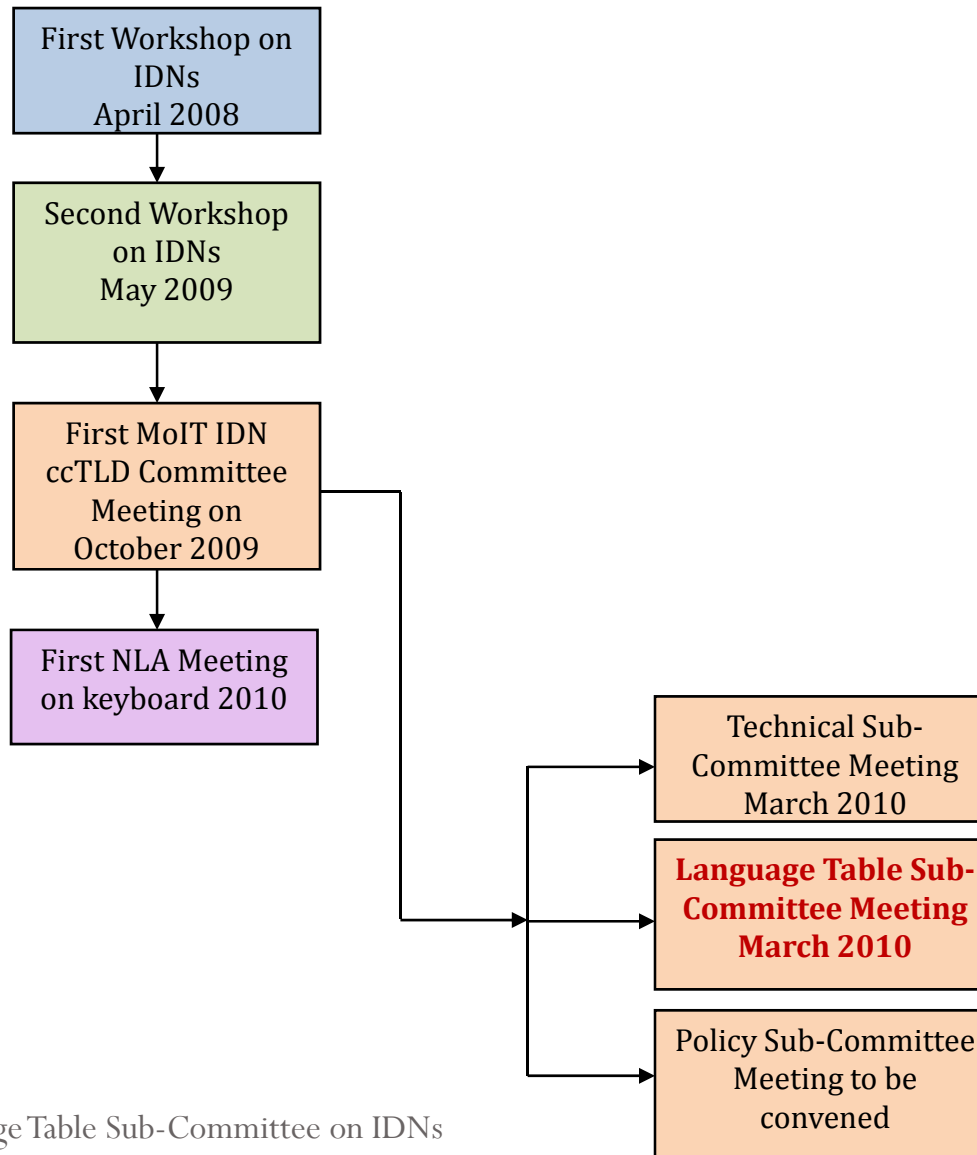
Sarmad Hussain and Rabia Sirhindi

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

Outline

- Status of IDNs in Pakistan – Work done so far
 - First Workshop (2008)
 - Second Workshop (2009)
 - پاکستان. Technical Committee Meeting (2009)
 - NLA Meeting on Keyboard (2009)
 - Technical Sub-Committee Meeting (2010)
 - Language Table Sub-Committee Meeting (2010)
- Work plan for this meeting

Sequence of Events



First Workshop on IDNs 2008

- The first workshop on IDNs was held at NUCES, Lahore on 20th April, 2008
- Attended by MoIT, language authorities and academies, language, linguistics and technology experts from academia and other relevant organizations
- Need for IDNs was explained
- A number of language and technical experts collaborated to develop individual language tables

Outcomes

- First draft of language tables created:
 - Balochi
 - Pashto
 - Punjabi
 - Saraiki
 - Sindhi
 - Torwali
 - Urdu
- Can be accessed at
<http://www.crulp.org/idn/download/LanguageTables.pdf>.

Second Workshop on IDNs 2009

- A follow-up workshop was organized by the National IT Development and Promotional Unit (NIDU) Ministry of IT on May 15-16, 2009 at NUCES, Lahore
- Divided in two sessions
 - Open public session on first day to seek public opinion on all IDN issues (invitation through newspaper advertisement)
 - Closed meeting of a group of technical and language experts on second day to finalize issues based on public feedback

Outcomes

- Major outcomes of the second workshop are as follows
 - Single language table to be adopted for all Pakistani languages
 - پاکستان. to be IDN ccTLD string for Pakistan
 - ‘.’ to be used as label separator
 - ZWNJ and ‘-’ (hyphen-minus) to be used as character separators
 - Extended Arabic Indic and ASCII digits to be used in Pakistani IDNs where both cannot be mixed in a single label
 - Confusable or variant characters to be mapped onto each other during registration
- Complete minutes of the workshop can be accessed at <http://www.crup.org/idn/IDN2009/techdoc.htm>

IDN ccTLD Committee Meeting 2009

- The Ministry of IT formed a committee for the implementation of پاکستان .ccTLD registry
- Members include representatives from the following
 - Pakistan Software Export Board, Islamabad
 - National Language Authority, Islamabad
 - Computer Society of Pakistan, Islamabad
 - E-Government Directorate, Islamabad
 - Pakistan Computer Bureau, Islamabad
 - National University of computer and Emerging Sciences, Lahore
 - Pakistan Software House Association, Islamabad
 - Kohat University of Science and Technology, Peshawar
 - ISP Association of Pakistan, Islamabad
- First meeting convened on 22nd October, 2009 under the chairmanship of Joint Secretary (Development), MoIT

Outcomes

- Major outcomes of the meeting were the following
 - Minutes of the second workshop on IDNs held on May 15-16, 2009 at NUCES, Lahore were unanimously approved and adopted by the Technical Committee for implementation of پاکستان. ccTLD registry
 - Three separate sub-committees established
 1. Technical Sub-Committee (to resolve technical issues)
 2. Language Table Sub-Committee (to finalize language tables)
 3. Policy Sub-Committee (to develop policies)
 - NLA to work for including ‘.’ on Urdu keyboard

NLA Meeting on Keyboard 2010

- National Language Authority meeting held in 2010
- The purpose of the meeting was to review Urdu keyboard design in the context of IDNs
- The committee decided the following
 - Have multiple versions of KB based on major languages
 - Each version support all characters in language table
 - Additional characters for email, Internet, etc. to be included
 - To be done with least change in existing KB version for backward compatibility
- Final keyboard design after language table for Pakistani IDNs is finalized
- Complete detailed minutes to be released by NLA

Technical Sub-Committee Meeting 2010

- The first technical sub-committee meeting was held on 25th March, 2010 at Pakistan Software Export Board, Islamabad
- The following agenda items were discussed
 - Progress on setting up test-bed for پاکستان. IDN ccTLD
 - Requirements for setting up test-bed registry for پاکستان. IDN ccTLD
 - Discussion of future steps needed for the implementation of پاکستان. IDN ccTLD, components and completion time frame
- Outcomes
 - Technical issues relating to test-bed operation resolved
- Complete detailed minutes to be released by PSEB

Language Table Sub-Committee Meeting 2010

- This meeting builds on the earlier work done on IDNs
- The purpose of this meeting is to finalize the language table for Pakistani languages for پاکستان. IDN ccTLD
- Specific agenda items of the meeting are
 - Finalize a single language table for all Pakistani languages
 - Discuss normalization and extended normalization cases for Pakistani language table
 - Discuss variant table for Pakistani languages

Work Plan for the Meeting

- Characters included in the Language Table for پاکستان.
 - Combined Unicode character table for all Pakistani languages
 - Honorifics and diacritics table for Pakistani languages
 - ASCII character table for Pakistani IDNs
 - Join control characters for Pakistani IDNs
- Normalization table
 - Normalization provided by Unicode (NFC)
 - Normalization not provided by Unicode (Extended Normalization)
- Variant Table
 - Variant characters based on shape similarity – Exact shape variants and Similar shape variants
 - Variants based on orientation of dots
 - Variants based on confusability with Arabic characters