On Localization, Natural Language Processing Tools and Use Models of Technology for Increasing ICT Use in Rural Nepal

Bal Krishna Bal
Assistant Professor, Kathmandu University
Email: bal@ku.edu.np
Major Localization Projects and Initiatives in Nepal

- Font Standardization Project supported by UNESCO
  - Nepali Unicode Drivers, Unicode compliant fonts
- PAN Localization Project, supported by IDRC, Canada.
  - NepaLinux, OpenCD, Localization of hand held devices, Spell checker, Nepali Computational Grammar Analyzer, Internationalized Domain Names, Optical Character Recognition, Parallel Corpus, Parts of Speech Tagger, Parts-of-Speech Tagset, Stemmer and Morphological Analyzer, Language Packs for Linux Distributions, Font converters
- Bhasha Sanchar, supported by the European Union
  - Nepali Monolingual text corpus, Spoken corpus, Speech corpus, Parallel corpus, Fonts, Text-to-Speech, Corpus based Nepali dictionary, Computational & Corpus Linguistics course at Central Department of Linguistics, Tribhuvan University.
Major Localization Projects and Initiatives in Nepal...

- Microsoft Windows/Office Language Interface Pack, supported by Microsoft
- Dobhase, English to Nepali web-based Machine Translation System, supported by PAN ICT R&D.
- One Laptop per child program (OLPC) and the current OLE as a continuity to the program.
- Ubuntu Localization, Sahana Localization – Open Technology Resource Center (OTRC), FOSS-Nepal.
Major Localization Projects and Initiatives in Nepal...

• Nepali Wikipedia
• Midas EduKit
• Localization of Content Management Systems
  – Joomla, Drupal
• More recent Developments
  – Big software giants like Google and Microsoft focusing on the Nepali language.
  – Localized applications continue to be developed for different platforms on an individual basis
    • Transliteration applications, Keyboard software, Conversion systems, Game applications
Combining the strengths of Localization and Natural Language Processing

• Localization involves:
  – Developing/customizing any software application as per local and cultural needs
  – Translating the UI and other system terminologies of any application into the local language

• Natural Language Processing goes one step further, it enables the machine to generate and understand the local language.
Combining the strengths of Localization and Natural Language Processing

• Some typical examples:
  – SMS reader for mobile phones (localized interface together with text-to-speech application)
  – In-built spell checkers in word processors (spell checker with localized word processor application)
  – Digital Archiving Systems and OCR systems
  – Localized E-gov portal with regular services:
    • Citizenship card issuance
    • Passport issuance/renewal
    • Driving license issuance/renewal
Use Models of Technology

• How do we ensure the use of the developed technology?
• How do we tie technology’s use in our daily lives?
• Technology should be designed and developed keeping into consideration the motivations for use.
• The bottom line – “People use technology if they feel its use, its benefits, its impacts”
Use Models of Technology...

• Lets revisit the examples on slide 6
  – SMS reader (unintegrated)
    • Mobile phones with regular SMS service (standalone)
    • Text-to-Speech Application (standalone)
  – Word processors with spell checkers (unintegrated)
    • Word processors (standalone)
    • Spell checkers (standalone)
  – OCR systems and digital archives (unintegrated)
    • Digital archiving system (isolated)
    • OCR systems (isolated)
Some thoughts on combining the all three: Localization+NLP+Use Models

• Need to note – Localization and NLP are not hot cakes which would be readily consumed
• They have to be associated with some Use Models
• Lets take an example of a deployment X:
  – Some questions to consider:
    • Is there a direct connection between the deployment and daily use? For example, use of X for teaching learning and other activities of daily use.
    • Is the technical\financial support required for X locally manageable and available?
    • How closely is the community attached with X? Sense of ownership
    • More importantly, do the community people see the prospects of the sustained and continued use of X?
    • The last but not the least important point: “It is the community people who would need to take up the sustainability and support of X, not you or anybody else.”
Thank You

Questions