



Research Component on Technology Concluding Thoughts

Sarmad Hussain

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

Technology Output

- Most countries are making very good progress
- Afghanistan
 - completed deliverables
- Bangladesh
 - issues with OCR, translation accuracy for parallel corpus
- Bhutan
 - TTS work delayed, but now progressing with support of NECTEC
 - OCR work on hold, but hope will re-start
- Cambodia
 - Work progressing well on TTS, OCR and language applications
 - Work on translation of parallel corpus on hold
- China
 - Limited status reported; work delayed for non-technical reasons

Technology Output

- Indonesia
 - Good progress on POS, Corpus translation and tool development
- Laos
 - Work on TTS and OCR progressing with the support of NECTEC
 - Work on language processing slow progress
 - Work on plug-in proceeding with support from RS
- Mongolia
 - Good progress on localization, language processing and speech recognition projects
- Nepal
 - Good progress on grammar development, reasonable progress on OCR

Technology Output

- Pakistan
 - Good progress on localization of SeaMonkey, OO and other tools
- Sri Lanka
 - Progress on mobile h/w recognition system
 - Reasonable progress on e-learning system
 - Slow progress on TM tool development

Overall Picture

- Excellent progress in technology development
 - Except in few cases
- Great collaborations turning work into successful research and development
 - Within partners
 - Across other organizations
- Developing a rich repository of resources for Asian languages, which will be very helpful even beyond the project

Issues

- HR availability, development and retention
- Linguistic Training
- Simultaneous English and local language competence
- Need for further synthesis of work
- Need for formal training at universities

Future

- Collate and develop a linguistic resource repository
- Write experiences across partners for various works
 - IDN
 - Lexica Resources, POS and Tagging
 - Text Processing, Computational Phonology and Speech Processing
 - Technology Policy
- Continue the collaborative work
- Online parallel (?) terminology repository
- Figure out linguistic resource distribution mechanism for longer term
- Experiment with commercialization for sustainability
- Phase 3?

Thank You