

# Lao Localization Project under PAN Localization Project

## Report of Phase 1

Presented by:

Phonpasit PHISSAMAY, Project Manager  
Valaxay DALALOY, Project Coordinator

# Plan and Output

## ■ Research Plan

6/2004-6/2006

- Collation Sequence
- Keyboard Driver
- Encoding Utility
- Sorting Utility
- Font
- Lexical
- Spell Check
- **Grammar Check**

6/2006-3/2007

- Laopad
- OCR
- Corpus
- Plug in OO

## ■ Research Output

6/2004-6/2006

Collation Sequence (100%)

- Keyboard Driver (100%)
- Encoding Utility (100%)
- Sorting Utility (100%)
- Font (100%)
- Lexical (80%)
- Spell Check (80%)
- Find/replace utility (100%)
- Line Breaking utility (100%)
- Lao-English-Lao Dictionary (80%)

6/2006-3/2007

- Laopad (Localize Notepad) (100%)
- Lao OCR (70%)
- Text Corpus (50%)
- Plug in OO (10%)

# Implementation Team

- Established Natural Language Processing Group, which consisting of 3 teams
  - Architecture team consists of 5 persons, which 2 from STEA, 1 MIC, 1 NUOL and 1 Newspaper agency
  - Linguistic team consists of 3 persons, which 1 Language Research Institute and 1 from Vientianetimes newspaper and 1 from STEA.
  - Programmer team consists of 7 persons, which all are from R&D division of STEA.

STEA is responsible for all the management and coordination facilitation

# Human Resource Development

- For the Project team
  - 3 senior staffs has attended the 1<sup>st</sup> training in Lahore (2004)
  - 3 technical staffs has attended the summer training in Lahore (2006)
  - Training on C++ and Font creation in Vientiane; 18 staffs (12 from project staffs and 6 from University) has attended. The instructor from Pakistan (Mr. Amar)
  - Training on VB.Net in Vientiane; 10 project staffs has attended. The instructor from Pakistan (Mr. Nadia).
  - OCR Training in Vientiane; 10 project staffs has attended. The instructor from Sri Lanka (Mr. Dulip).
- For the other Local researchers and private company
  - Being the LAB for the students from various university in Laos. (New fonts creation, Keyboard and other research activities)
  - Joint development of some software application (Database that using project output)
- For the users
  - 1 National seminar on Lao Localization in Vientiane.
  - 2 Joint organized Localization workshops (define standard) in Vientiane
  - 2 Seminars on Project output in Luang Prabang province and Champasak province

# Publication

- [www.laol10n.info.la](http://www.laol10n.info.la) (Project website)
- 350 CDs of the project outputs.
- Research papers in both hard and soft copies for both Lao and English language of each reach development.
- Regularly advertise the research output in the Local newspapers to encouraged the public to widely use the research products
- Under development of Lao Language Computing book (Lao language, Localization technology and user guide for the research product) which expected to be presented in the National Localization workshop in 13-15 March 2007 ( IDRC, NUCES and Country Project leader are invited to attend this workshop)

# Project benefit

## ■ Policy and standard

- Unicode Fonts: The research output on Lao Unicode Fonts and ASCII to Unicode converter has encouraged the widely use of Unicode
- Keyboard layout: The research output on Lao keyboard layout and driver, which Microsoft is already include in the Windows Vista will encourage national standard
- Sorting utility: Define the Lao collation standard

# Project benefit (2)

## ■ Content:

- 2 Press offices are using Line breaking and spell check utility for writing the newspaper.
- Sorting utility and Unicode converter are being used for creating the database and accounting system in some offices.
- Some webpage in Lao language are using project product such as:
  - [www.pasaxon.org.la](http://www.pasaxon.org.la): This is daily public newspaper in Lao language
  - [www.lnr.org.la](http://www.lnr.org.la): This is the daily radio in both Lao and English

# Issue

- The works related to the design of the system and algorithm are developed by the foreign experts while the project staff just doing the program coding and input data. The staff are still not able to update and modify program by themselves. Even these people has got a lot of training of those subject.
- The research output are not widely used in the public, because each utility are not insert into the software application (MS office), therefore the user have to take “work” from the software application to operate in utility and copy the result back to software application, which really not convenient and they are refuse to use the utility.
- The awareness of the Lao language computing are still very low in the country, which many ICT project seem to excluded the development of the Lao language in those applications and technology, especially related to the online service and network communication.
- There are lack of Lao language computing policy. Many organization take its own initiative, which many activities were duplicated and conflicted among each other.

# Conclusion

- The efforts of Lao localization project has been recognized and appreciated in many academy institutes and some government organization and ICT private companies.
- The project output is enhancing the utilization of Lao language in the Information technology, especially Lao language content are being develop in many local websites.
- The institutional capacity to carry the research and development of Lao language information has been established and strengthening.

The Natural language processing group and the people of Lao PDR are express its sincerely and highly appreciation to the PAN Localization Project, which not only benefit the Lao people but also to many people in this region who not speak English

Thank You

For your kind attention