Future Development for Thai Language Processing

Dr. Virach Sornlertlamvanich
ETL Japan. NECTEC Thailand.
ORCHID Project

Open linguistic Resources CHannelled toward InterDisciplinary research

– To overcome our language barriers and to preserve our cultural heritages.
Universal Networking Language

- To replace the HTML.
- Started in 1996
- 15 nations, 16 languages
- UNL, dictionaries, grammar rules, UNL servers
Emacs for Thai

**MULE**: \( \text{MUL} \)-tilingual enhancement to GNU Emacs

- Enhancement Thai editing: insert, backspace, word wrap, word move, etc...
- Dictionaries lookup: RI, LEXiTRON
- etc.
ORCHID Thai POS
Tagged Corpus

- Language resources
- http://www.links.nectec.or.th/orchid/
Plain text with structural mark up

http://come.to/virach/