# KIPO's MT Activity & IP5 Mutual MT Project

August 30, 2009

K-PION

Young-Pyo Kim

**Technical Cooperation Division, KIPO** 





### **1. KIPO's MT Activity:** *What the KIPO is doing...*

- ✓ History
- ✓ Machine Translation Services
- ✓ Quality Improvement
- ✓ Future Plan

## **2. IP5 Mutual MT Project**

- ✓ Overview
- Technical Implementation
- ✓ Short-term Tasks



# **KIPO's MT Activity:** *What the KIPO is doing...*



#### **Historical Review of KIPO's MT Services**

- ➤ 2000: Introduction of a Japanese to Korean MT system
  - Rule-based engine
  - Technical terms dictionary
- ➤ 2005: Introduction of a Korean to English MT system
  - Pattern-based engine
  - 300,000 pattern DB, technical terms dictionary
- > 2008: Introduction of an English to Korean MT system
  - Pattern-based engine
  - 300,000 pattern DB, technical terms dictionary



#### **KIPO's MT Activities for Patent Documents**





#### **MT** Quality as Evaluated by Patent Attorneys

## > MT Quality Evaluation

- Evaluation period: 1~2 Times / Year
- Evaluators: 5 patent attorneys from the U.S. and Korea
- Evaluation method: randomly selected 200 sentences for evaluation
- The highest and the lowest scores were excluded in order to obtain an evaluation result that is not biased.



#### **MT Quality as Evaluated by Patent Attorneys**

## > Standard of Score: criteria established by KIPO

Score	Level
5(100%)	Entire sentence is correctly translated, but may not be perfect
4(80%)	One or two words are translated incorrectly, but the essential meaning of the sentence is preserved
3(60%)	Keywords are correctly translated, so the sentence is partially understandable, but the essential meaning of the sentence is not preserved
2(40%)	One or more words or phrases are correctly translated but keywords are not translated and so the essential meaning of the sentence is not preserved
1(20%)	Missing sentence or translation different from original text in content
0(0%)	Absolutely incomprehensible



#### **MT Quality as Evaluated by Patent Attorneys**

## > History of MT Quality Evaluation

Туре	2004	2005	2006	2007	2008	2009
K2E	54.25%	76.5%	79.3%	80.6%	80.4%	81.3%
E2K					81.3%	82.1%
J2K			97.1%	96.8%	96.9%	97.4%

2005: Specialized MT system customized for patent domain

2009: MT Dictionary Upgrade



- > MT system customized for patent domain ('04~'07)
  - MT engine was originally developed by ETRI and then made more advanced by Siriussoft (2004 - 2005)
  - Established 3.2 million technical terms (2004 2005)
  - Established unique sentence patterns in each of the fields of patent documents, such as Title, Abstract, Claim, and Description (2005 – 2007)
  - Optimize the dictionary, sentence pattern for the examination reports (2006-2007)

Target	Before (2004)	After (2007)	Improvement
Patent Documents	54.25%	80.60%	26.35%



## > MT Dictionary Upgrade ('08~'09)

- K2E: 300,000 terms incorporated into the MT dictionary
- E2K: 300,000 terms incorporated into the MT dictionary
- J2K: 300,000 terms incorporated into the MT dictionary

Service type	~2007	2008	2009	Total
K2E	3,200,000	300,000		3,500,000
E2K	3,000,000		300,000	3,300,000
J2K	1,200,000		300,000	1,500,000



## **1 MT Quality Improvement & Feedback**

 Improving translation quality by making good use of user feedbacks (unregistered word, typos, grammatical errors)





## **2** Automatic Extraction of New Terms

- Terms which do not exist in the dictionary are automatically retrieved from patent specifications.
- Such new terms are then translated by expert human translators.
- The pairs of new terms and their corresponding translations are then added to the dictionary.

## **③ Regular/Irregular Spell Check**

- New terms are spell-checked before they are added to the dictionary
- Existing terms are also irregularly spell-checked.



## **④** Quality Assessment

- Once or Twice a year

## 1, 2, 3, 4

## **Regular System Update**

- MT dictionary Monthly
- MT engine Quarterly



## **Future Plan**

- > Implementation of K-PION services for trademark examination history/reports and design examination history/reports (2009)
- > Implementation of a K-PION service for PCT gazettes published in Korean language (2009)
- > Implementation of a foreign language translation service for KIPO examiners (2010~)



# **IP5 Mutual MT Project**



#### **Overview of the IP5 Mutual MT Project**

#### Goal: To offer high quality machine translation services







#### **Technical Implementation**

## > Quality Improvement of Machine Translations

- Upgrade MT systems through feedbacks obtained from the review of machine-translated documents (Short-term Task)
- Control errors in original documents
- Exchange information and share know-how on MT
- Develop direct mutual MT systems (under agreements between the offices involved)

## > Harmonization of MT services

- Harmonize the service contents of the different MT services
- Implement a free text-translation service



#### **IP5 Short-term Tasks** – *what is next?*

## > Resources required for the review of sample sentences and the upgrading of MT systems

- quality standard and guideline for conducting reviews
- error description method
- number of sample sentences to pick
- sentence sampling method
- method and frequency of MT system updates



# Thank you !

## E-mail:zedero@kipo.go.kr

