

[초청강연]

학 술 발 표

Introduction of NLP and MT

- EverTran 강명주 이사 -

Introduction of NLP and MT

강명주 이사

mjkang@evertran.com



www.evertran.com

Agenda

- I. NLP Technology Overview**
- II. Product Overview**
- III. Introduction of Visual TranMate**
- IV. Demonstration**



Domain of NLP Technology

Natural Language Processing Technology

Information Retrieval

- Information Search
- Data/Web Mining

Voice Processing

- Voice Recognition
- Text to Speech
- Voice Composition

Language Processing

- Machine Translation
- Optical Character Recognition
- Document Summarizing & Clustering

- 3 -

EverTran Co., Ltd.

Enjoy Translation with Tools!

Domain of Language Processing Tech

Machine Translation

Definition by Hutchins and Somers:

Computerised systems responsible for the production of translations from one natural language to another, with or without human assistance...

The central core of Machine Translation itself is the automation of the full translation process.

Vgl. Austermühl, Frank. 2001. *Electronic Tools for Translators*. Manchester: St. Jerome Publishing. S.157

Definition...

Machine Translation

What to do?

In its broadest sense MT can include computer applications which convert a file or Web page in one language into a file or Web page in another language.

- 4 -

EverTran Co., Ltd.

Enjoy Translation with Tools!

Domain of Machine Translation

Human vs. Machine



Computer

| | | |
|------|--|--|
| FAMT | | <i>Fully Automated Machine Translation</i> |
| HAMT | | <i>Human Aided Machine Translation</i> |
| MAHT | | <i>Machine Aided Human Translation</i> |

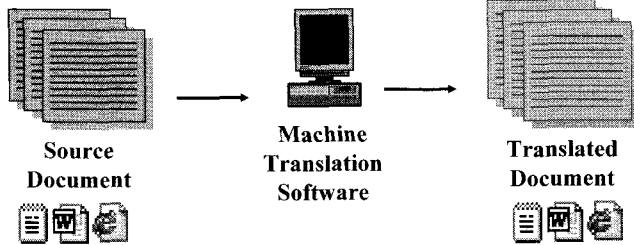


Translator

Domain of Machine Translation - FAMT

Fully Automative Machine Translation :

The source language text is fed into the computer as a file or Web page and the computer produces a translation automatically without any human intervention.



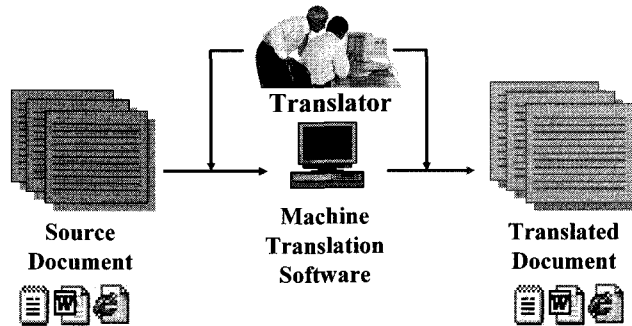
Domain of Machine Translation - HAMT



Human Aided Machine Translation :

The translation is carried out by a computer but the source language text is modified by a human translator either before, during or after the translation process.

Key word: human intervention in the translation process.



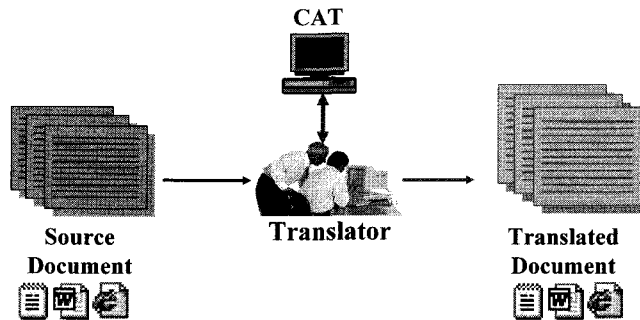
Domain of Machine Translation - MAHT



Machine Aided Human Translation :

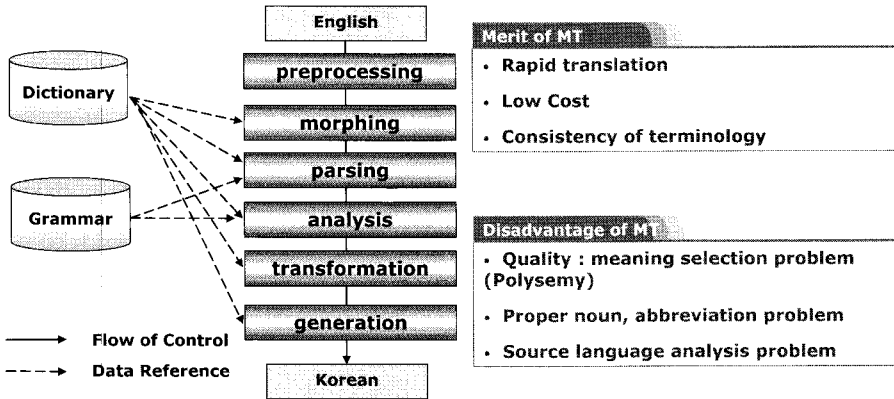
The translation is performed by a human translator, but translator uses the computer as a tool to improve or speed up the translation process using Translation Memory(TM) and Terminology.

This process is usually also called Computer-Aided Translation (CAT).



Key Technology of FAMT

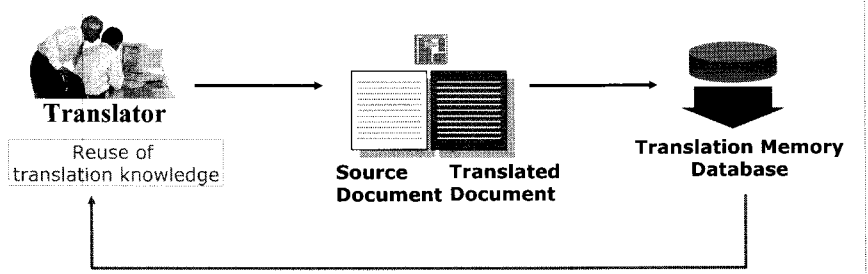
Morphing/Parsing/Analysis

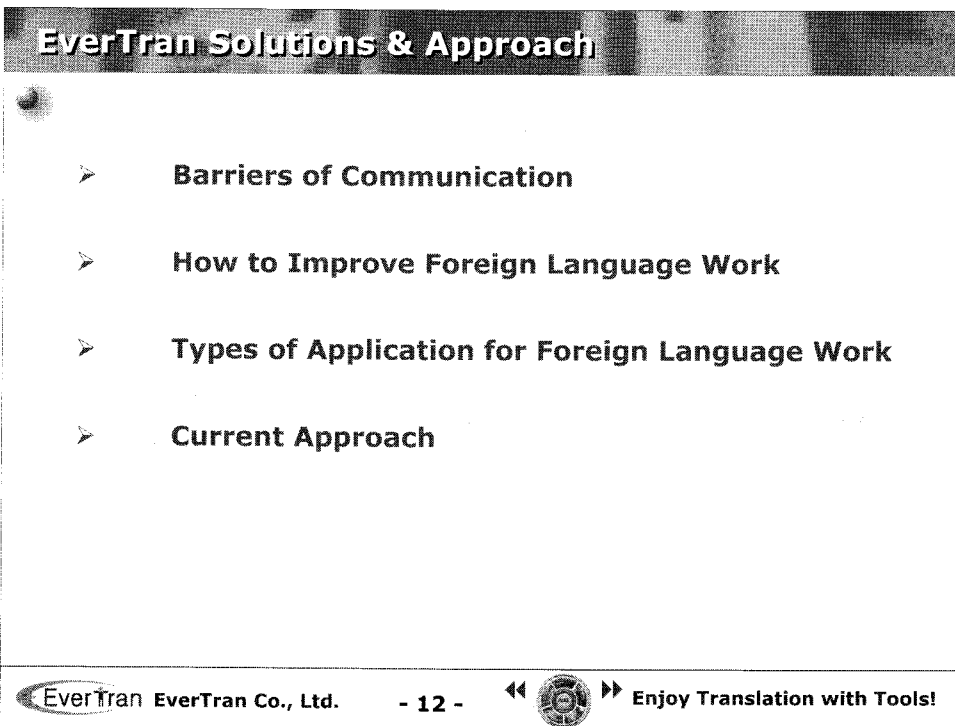
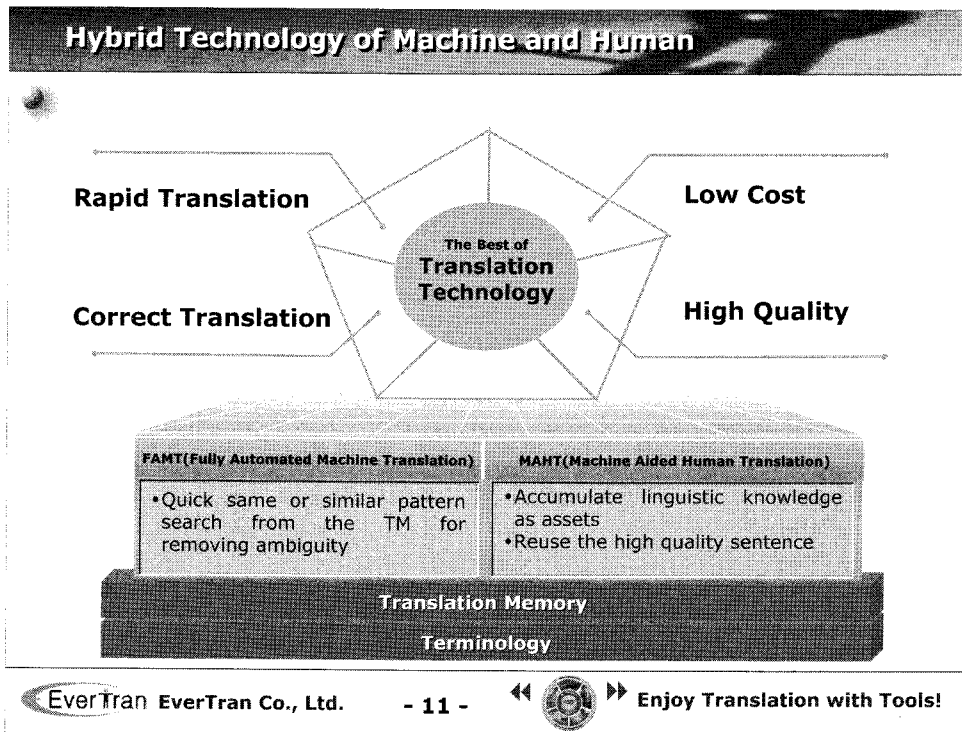


Key Technology of MAHT

Translation Memory/Terminology

- Translation Memory(TM) : Database made from pairs of source and translated sentences
- Searches the best match in the existing translation memory for a sentence (translation unit) to be translated.
- Reuses the translated sentences saved in the past





Barriers of Communication

Globalization
Localization

Language Barriers
Cultural Barriers

Multinational Company

EverTran EverTran Co., Ltd. - 13 - Enjoy Translation with Tools!

Innovations to Improve Foreign Language Work - 1

High Quality Low Cost

Human Translation Service
Publishing-level Translation Quality (Contracts, patent applications, manuals and proposals)

Machine Translation System
Human-understandable translation quality with rapid speed (Searching and acquiring information, Review)

X

Quality

Price, Time (Cost)

EverTran EverTran Co., Ltd. - 14 - Enjoy Translation with Tools!

Innovations to Improve Foreign Language Work - 2

| Category | Innovation | Method | Outcome |
|---------------|--|--------------------------------|--|
| Innovation I | <ul style="list-style-type: none"> • Manage process, quality and cost with the help of Computer Aided Translation | CAT (Visual TranMate) | Improved quality & productivity in translation |
| Innovation II | <ul style="list-style-type: none"> • Construct terminology DB (Special Dictionaries) • Construct TM • Post-editing (User Dictionary, etc) | Hybrid of MT and TM (AutoTran) | Improved quality in MT |

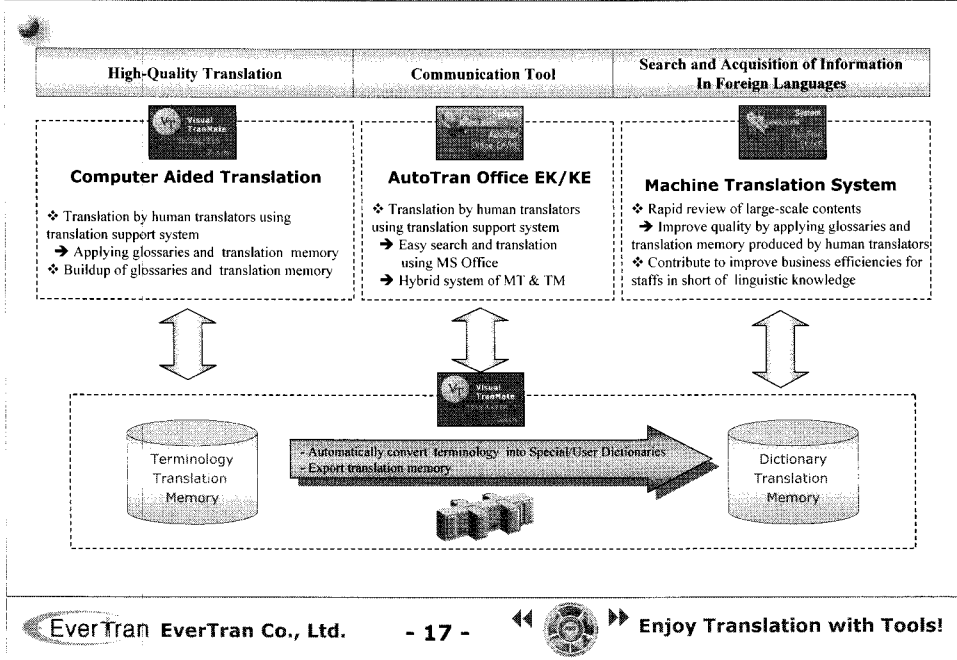


Current Approach

| Foreign Language related Work | Relevant Documents | How and Who? |
|--|---|---|
| High-quality translation | Manuals, Technical documents, Patent applications, Written contracts, Proposals, Marketing related contents(multilingual web pages and pamphlets), etc. | External translation companies (Outsourcing), Employment of translators |
| Searching and acquiring information in foreign languages | R&D, marketing, advanced research for patent application, contract reviews, etc. | Internal human resources or outsourcing if necessary |
| Communication | E-mail or business letters (communication with foreign companies or organizations) | Internal human resources |





EverTran Application Products

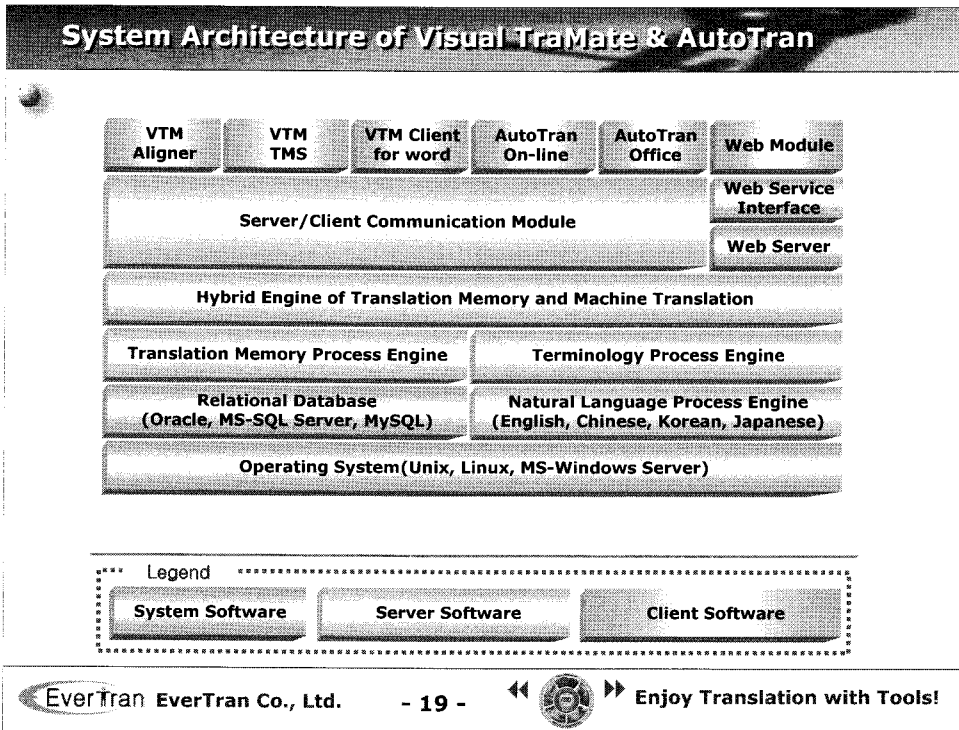


Benefit of Products

Benefits of introducing this system: Direct effects from translation processes and Indirect effects from assetizing linguistic knowledge by system management are expected.

| | | |
|-------------------------|---|--|
| Direct Effects | Reduce translation cost | Save cost for translation by reusing Linguistic Knowledge Save cost for review with consistent use of terms and sentences |
| | Improve translation quality | Reduce risk factors of communication error Generate consistent translation result by sharing Translation Knowledge |
| | Reduce man-hour for translation(Delivery) | Control manpower for increased translation jobs Advance delivery date with systematic schedule management |
| Indirect Effects | Share Linguistic Knowledge | Utilize TM and Term DB constructed by experts in each industry |
| | Standardize in Communication | Prevent misuse of terms by utilizing consistent terms and TM |
| | Assetize Linguistic Knowledge | Improve company's asset standing by assetizing accumulating linguistic knowledge |


EverTran Co., Ltd.
- 18 -

» Enjoy Translation with Tools!



Server Modules of Visual TranMate

| Components | | Functions | |
|--------------------------|--|---|--|
| VTM Server Module | Web Module | <ul style="list-style-type: none"> Management module for linguistic knowledge archived in the server Providing the administrator mode | |
| | VTM Server Module | Communication Module | <ul style="list-style-type: none"> Receiving the request Transmitting the outcome |
| | | Pre- & post-processing Module | <ul style="list-style-type: none"> Processing the outcomes Running modules corresponding to the request |
| | | AutoTran Office | <ul style="list-style-type: none"> MS-Office Add-In Module for Word, PowerPoint, Excel Search terminology and translation Memory DB Hybrid of MT and TM |
| | Natural Language Processing Engine (NLP Engine) | <ul style="list-style-type: none"> A module that provides searching output from terminology DB and translation memory DB applying morphological and syntactic analysis of user's queries | |
| AutoTran EK/KE | <ul style="list-style-type: none"> Client/Server-based FAMT | | |

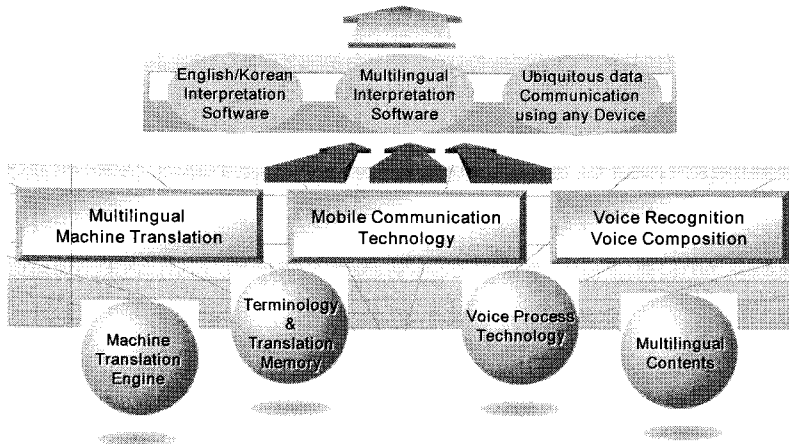
EverTran Co., Ltd. - 20 - Enjoy Translation with Tools!

Client Modules of Vsual TranMate

| Components | | Functions |
|---------------|--|---|
| Client Module | VTM Client | VTM For Word ❖ Human translation workspace integrating all functions required for human translation |
| | | VTM TMS ❖ A module that automatically extracts terminology candidates in the source language text and helps user's glossary buildup |
| | | VTM Aligner ❖ A module that analyzes source language text and target language text, and semi-automatically builds up TM |
| | AutoTran EK/KE ❖ Hybrid system combining MT technology and TM management and search technology in a single interface | |
| | AutoTran Office ❖ Module that added on MS-Office (MS-WORD, EXCEL, POWERPOINT, OUTLOOK) and IE | |

Technology Vision of EverTran

Natural Language Global Company
Any Language, Any Device, Any Where





Enjoy Translation with tools !!!

Thank you!

www.evertran.com

 **EverTran Co., Ltd.** - 23 -  **Enjoy Translation with Tools!**