

ESTeam BTR

G. Magnusdottir, L. Cranias, Y. Aglamissis, C. Roniotes,
A. Samiotou, G. Dukas, Z. Cranias et al.

ESTeam Inc.
Markou Botsari 15
145 61 Athens
Greece

email: zilos@hol.gr
phone and fax: +30 (0)1 8085 704

The ESTeam BTR system is a toolbox for building customized solutions for multilingual translation processes. It contains large language resources and language independent processing components. ESTeam BTR features a combined practical approach of the basic MT and TM technical solutions and is designed for multilingual information browsing. It is intended to work where it is usually impossible to apply human translation due to several factors, the two most important being time and quantity. The system is intended to perform on large-scale translations, specifically with on-line resources and on problematic language quality. A multi-user version of the system has been implemented as an independent, fully-automatic translation installation requiring no human translation intervention, at any stage.

Large scale, multi-user, fully automatic and **multilingual** are the keywords that have prompted the development of ESTeam BTR.

Multilingual, resulted in the pivoting architecture for both translation memory and machine translation components, where all development is into one language (English) and links to the other languages are automatically created through that language.

Large scale, Multi-user, instigated the development of an industrial size translation memory. This inevitably led to the introduction of on-line automatic ‘post-editing’ of machine translation results, which does not affect the information content, and to ensuring robust translations even where a word is misspelt. The system architecture itself combined with specially designed algorithms enhance on-line, multiple access and maintenance of the translations.

Fully automatic, prompted the elimination of human interaction throughout the translation process, including the ‘post-editing’ module and even during development, by introducing data verification across languages.

The MT component of ESTeam BTR is lexicon and data driven. Domain assignment at the lexical level, and restricted, shallow morphology are applied but no syntax. Translation units which cannot be treated by our ‘post-editing’ filter will be mostly word or phrase translations. In applications for which the system is not customized, the MT component is the only means of translation. MT-lexicons are translated into English and linked across languages (PIVOTING).

The translation memory component comprises an index of several million units. Its design is completely different to all stand-alone solutions currently on the market, since its task is to link multilingual, non-parallel original language texts through their translation correspondence in English (PIVOTING).

ESTeam BTR incorporates new languages with record speed. Three to six months are required for including a new language with its links to all the previously implemented languages. Thus far, the system includes five languages: English, French, German, Spanish and Italian (22 language pairs) and is currently being expanded with Danish and Dutch (release date end '97 - in all 42 language pairs).

Customization ensures translation quality and coverage above 90%, for tested applications, in terms of content accuracy and comprehensibility.

ESTeam BTR is implemented in C, running on an Oracle UNIX server with a minimum storage space of 3GB and 128 MB RAM.