

## **TRANSLI™ trusted automated translation software at the service of justice**

**Atefeh Farzindar**

NLP Technologies Inc.

1255 University Street, Suite 1212

Montreal (QC), Canada, H3B 3W9

[farzindar@nlptechnologies.ca](mailto:farzindar@nlptechnologies.ca)

### **Abstract**

This paper presents the comparison between two versions of TRANSLI™ - the automatic translation management system developed for machine translation of legal documents from English to French and from French to English by NLP Technologies Inc.

NLP Technologies is an enterprise devoted to the adaptation of advanced information technologies for processing large volumes of documents, such as the needs reflected in the legal industry. NLP's solution addresses our clientele's volume and time management issues. TRANSLI's statistical algorithm has been trained on a large corpus of legal documents as to significantly reduce production time and costs. Moreover, we've adapted and trained our own SMT engine with volumes of legal documentation, namely court decisions, in order to provide a language solution to the Canadian judiciary.

Canadian jurisprudence is predominantly produced in English; 75% of decisions are written in English and 25% in French. According to legislation, the Federal Courts are required to provide a translated version of decisions in order to grant access to information in both official languages. Quebec counts more than 30 administrative tribunals, the majority of which invest considerable re-sources in order to appease their translation needs. The legal industry is marked by continuous publishing and translation cycles, large volumes of digital content and growing demand to distribute multilingual information. The capacity to administer high volumes of translation requests without delay thereby becomes imperative.

Our goal is to study the ability to train translation systems on a specific domain or subject area like the legal field so as to radically increase translation accuracy. Our extensive research efforts have enabled us to build Translation Memories in an array of fields ranging from legal to medical vocabularies thus enabling us to provide terminological consistency and quick production turnaround regardless of the domain. Additionally, our technology has the capacity to be trained in novel domains ensuring that our services can cater to any of our clients' needs. An interesting aspect of our findings in the legal domain was that the review and the post-editing processes of judicial translations are an important part of an SMT-integrated workflow. Reviewers with subject knowledge need to have direct access to the translation process in order to provide a feedback loop to the SMT training process.

Evaluation of first and second version of TRANSLI is described in the paper. First version of TRANSLI was only based on trained data however second version is using post edited feedbacks in translation cycle again. The post editing data is inserted in translation memory for future use. We have compared the two version of TRANSLI in this paper. The comparison metrics are accuracy (percentage of the machine translated data that have not been touched after post editing), WER (word error rate), TER (term error rate), improvements in true casing (restoring case information to badly-cased or no cased text) and finally by the number of 'real' insert and deletes in post editing. Two systems are evaluated based on the mentioned factors then statistically proved that the second version is more efficient.

## Biography

Dr. Atefeh Farzindar is the founder of NLP Technologies Inc., a company specializing in machine translation and translation management systems. Dr. Farzindar received her Ph.D. in Computer Science from the Université de Montréal and her doctorate of languages, computer and cognition from Paris-Sorbonne University. She gathered academic and industrial experience in various positions in industry and academia.

She is Adjunct professor at the Department of Computer Science at the Université de Montréal and Honorary Research Fellowship at Research Group in Computational Linguistics at the University of Wolverhampton, UK. Dr. Farzindar is a member of the Board of Directors of The Language Technologies Research Centre (LTRC) of Canada and the chair of the language technologies sector of the Canadian Language Industry Association (ALIA).