Enterprise Application of MT: Progress and Challenges

Craig Plesco
Nestor Rychtyckyj
Ford Motor Company
Dearborn, MI 48121

cplesco@ford.com
nrychtyc@ford.com

Abstract

At the AMTA conference in 2012 we first reported on the deployment of a global enterprise machine translation service at Ford Motor Company. As a multi-national company doing business around the world Ford has many requirements for translation. These translation requirements come in many different varieties and often lead to different solutions. As we have deployed machine translation for use throughout Ford, we are usually the first point of contact for new requests. In this presentation we will discuss the progress and challenges that we have faced over the last several years.

Our internal translation service uses the translation software from Systran Software Inc. (through the Systran Enterprise Server (SES) and Apptek Inc. We have developed over 90 Ford and industry-specific dictionaries that can be used to customize a translation request. The translation requests can be processed through a text box or with various supported file formats. In addition, we have developed a system interface for translations that allows applications to programmatically access the translation systems with background or real-time requests. These translation servers are hosted internally and provide a secure environment for our users. In order to address translation requests from new users, we have developed a process and template for users to describe their needs and expectations for translation.

We currently support 19 languages and process more than 200,000 translation requests daily. Batch translation requests are scheduled on a staggered timeline to enable reliable throughput. Our users include manufacturing, quality systems, plant floor systems, dealer concerns, warranty claims and other applications. As word of our service has spread throughout the company we are frequently approached to support other users and applications. These requests fall into a number of different categories.

In some cases we determine that MT is not the right solution. These include translations of legal or corporate documentation that would need significant customization and post-editing to deliver usable results. If we determine that MT is a viable solution we then perform text analytics on the source text to extract the terminology that will need translation. This extraction is done using natural language processing and ontologies with the goal of identifying “not found” terms and their frequency of usage. These “not found terms” are then translated manually and added into the translation glossaries.

Another common request is need for new languages which are not commonly available. In these cases we often work with the users and the vendor to apply statistical approaches to develop a language model more quickly than formal language modeling and with acceptable business accuracy. This was the approach that we used with Systran to build an English-Thai translation system. We have also developed and deployed a translation system with Systran for Romanian where the source text is converted to a parse tree prior to translation which results in much more accurate results.

We have found that Machine Translation has many significant advantages and uses in a corporate setting. However, it is critical to understand the user requirements and manage their expectations in regards to translation accuracy.