Integration into Workflow - what makes MT work?

Chris Pyne, Sietec Systemtechnik GmbH & Co. OHG

Abstract

The notion that MT systems and tools can be treated in the same way as any domestic appliance; "just plug in 'n play" is a misconception which has led to frustration on both sides of the fence, both on the part of the customer and the MT producer.
There is a growing awareness that certain factors are essential for the successful use of any language tool in a specific environment. By examining environments where language tools are (or not) objectively successful, it should be possible to crystallise out the essence of "what makes an MT system work" or "how can an MT system be successfully integrated"

Introduction

My subject is a (somewhat self-critical) appraisal of machine translation in 1994 and directions it may take in the next 5 years. I must say now that my comments and evaluations are based 99% on my experiences in Germany in the past decade.

Sietec is a subsidiary of the Siemens Nixdorf company. Sietec has various key business areas, archiving technology, system integration and linguistic systems.

My responsibility is consultancy and sales for our linguistic systems. Probably the best known of our products is the machine translation system **METAL.**

I do not intend to turn this into an advertising show for our products, but rather use this as a platform to raise issues which I believe occupy a lot of people involved with machine translation tools - either as users or manufacturers.

The term machine translation here is used loosely to cover any system which claims to make translators more productive, make their lives easier... etc.

I didn't know we had a problem!

Many firms assume that purchasing an MT system will lead to an immediate solving of problems such as overdue translations, supposedly high external translation costs etc. The "Band-Aid" mentality is understandable but is not a firm foundation for setting up any MT system. The nebulous title of this paper describes a necessity for success - integration into workflow.

We are often approached from companies who seem to have a vague feeling that their documentation and translations are not as good or efficient or streamlined as they should be. They are looking for a quick and simple solution. The first thing to be said to them is that there are a lot of solutions available, hardly any are simple. You cannot get something for nothing.

A good way to start in such a firm is to discover just how information flows at present, bringing the human protagonists together - data processing manager, translators, technical writers. What ensues is nearly always interesting. The pattern tends however to be the same.

- Communication within such a firm between technical writers and translators is poor - both at the human and digital level

- At the human level translators tend not readily to be accepted by the technical engineers, often with a scientific or engineering background. The fact that translators often have to cover a much wider range of technical areas seems to go unnoticed.

This communication problem is an area that must be tackled no matter what MT tools are being contemplated.

- At the digital level there is often no complete reuse of the documents produced by technical writers or other departments because the translators are not working with the same editing tools as the technical writers.

These anomalies must be investigated and corrected. Complete reuse is vital for integrating tools successfully. Acceptance is vital if translators are going to be able to influence the production stages before the translation.
The texts are often not written with translation in mind. This means they need to be totally reworked from the point of view of grammar, ambiguity, length of sentence and amount of ideas in one sentence.
Good document formatting is vital. Generally the expensive technical authoring tools available are abused as electronic typewriters.
Inconsistent use of terms.
Poor milestone control between departments.

What is the problem?

Graphic user interfaces have reached the stage where virtually any product can be made to look good on the surface. Demo files supposedly show what the system or tool can produce. These are factors which can raise initial interest, but we need something more than this for successful products. A potential customer needs to know how to integrate the tool, make it work in his environment with his texts. This goes far beyond the training for just which hot key has to be pressed and in what sequence. This we could call thinkware/orgware - it relates rather to the backend of the system; it may not be pretty but must be practical and address the legitimate needs of users. A translator from the United Nations recently said that many of the modern PC-based translation memory product manufacturers present at the event had talked a lot about the GUI's but not about the backend; the organisation, the structures and the benchmarks.

Some of the background areas worthy of mention would be:

- Sound bilingual alignment procedures - whereby any of the procedures must be integrated into the organisational concept
  Resulting from this recommendations concerning which texts to align.

- Sound administration procedures where workflow considerations must relate to interactive and batch sequences, hierarchical access and read/write structures
  Resulting from this recommendations concerning the kind of person who may have to fill this role

- Strategies for importing terminology into the system
  Resulting from this recommendations as to what can be default imported, preparatory data structures which may save time afterwards.

- Benchmarks relating to relationship between retrieval times and archive sizes.
  Resulting from this, recommendations concerning workflow management, efficient implementation of lengthy batch processes etc.

One of the greatest problems concerning the integration of MT tools is the effect it has on the translators' working environment. This change can be drastic and must be something that the translators actively support. Looking back at failed translation system installations in Germany, patterns again begin to emerge:

- The translators were not prepared for the system.

- The translators were not consulted about the system and felt that they had been left out of the decision process.

- This in turn produces a negative chain reaction. Translators are then trying to prove what the system cannot do rather than what it can do. It is very easy to make nearly any translation system look stupid initially - it is more difficult to make it really productive.

- Day-to-day working processes were not analysed or adapted to suit the tools. Management thought the system could produce excellent translations from day 1. This is a point I shall be returning to when we examine our own situation vis-à-vis METAL etc.

- Management was not prepared for initial system learn phases where productivity gains are not apparent.

Conversely, looking at successful translation system installations in Germany and Switzerland, we see that charismatic personalities, normally the translation department heads, play a decisive role. It is necessary for the translator to break out of the traditional passive role allotted to him/her in the firm and become active. Translations should take place in dialogue with the departments producing the material to be translated. What I am suggesting is that the role of the translator

especially within a firm or organisation could change dramatically in the light of technical developments. The proposed roles could be some or more of the following:

- Translation is a service industry, this is nothing new. The question is what services can the translator provide for his/her company.

- Of course, the first service is translation of documents directly either manually or using tools.
- A second service could be helping others to translate straightforward undemanding texts themselves.

Here the translator is also acting as know-how provider. He/she is constantly inputting his/her own know-how into a knowledge-based MT system.
This could be used and integrated at many levels.

- Most obviously for other translators but also for other employees needing a certain amount of translation support.
- This in turn may be information level translation, where at present the poor translator is told afterwards "I don't know why it took you that long - all I wanted was the gist!"
- These are clients accessing a knowledge server. Of course this structure must be reproduced in the system architecture (client-server).
- Finally, if this is not being done already, company-wide terminology control.

Translators are still very sceptical about these scenarios and tend to rather want to horde their knowledge and experience, which they feel will then improve their job security. It will not. This tends rather to become a vicious circle. The amount of mundane foreign correspondence, e-mails, manuals, repeats of previously translated manuals and repeats of repeats of previously translated manuals will increase. When departments in big companies discover internal bottlenecks, the tendency is to start looking for outsourcing options. The end result is an accelerated stagnation of the internal translation services.

Implementation

As I said earlier, management expectations regarding initial quality tend to be too high. The initial investment in any system is only the start of a commitment to make it work. The user cannot be left "standing in the rain" at this point - realistic hand-holding and sympathetic hotlines are very important.
Vendors must be realistic about the system's ability to learn in relation to speed and quality of the learn process.

According to some of our own users, a lexicon build-up time of 6 to 9 months has been usual up to now. This is far too high and can be reduced by using some of the following strategies:

- Using the system's own experience and intelligence to default new terms in a batch process.
- Providing off-line editing sub-tools which make this work effectively.
- Giving the idea new users a head start by efficiently implementing existing terminology out of an in house term base.
- Another possibility is to provide standard terms from specific technical areas beyond the range of the standard system lexicon supplied.

Success with MT does not come overnight. The most successful implementation strategy has been one of introduction in small steps, "homeopathic doses" as one our users put it. Only then is it sensible to move on. It is of course a big step to provide non-linguists with access to the system - this can only be done for specific areas after much preparation.

On the client side, interactive pre-editors using linguistic analysis modules for the source languages will help to drastically improve the translation quality of texts submitted by and returned to non-linguists.

What sort of system does a company need?

It is not much use providing non-linguists who would like to have translation support for simple texts such as desktop messages or gist translations with a system that only provides translation services requiring linguistic decisions and manipulations and leaves large amounts of source text untranslated. A system which makes such requirements on the user automatically defines and restricts the user profile. Ideally the user should be able to access and define the range of translation services required from his/her client connection himself.

These could fall into the following areas:

- Source text analysis and control

- Retrieval of defined corporate (global) bilingual memory modules

- Retrieval of personal (local) bilingual memory modules

- On-line translation of formatted documents

- Terminology lookup

Nearly all of these options are either available now or will be available in first versions within the next few months.

This options range will open up new market areas beyond those traditionally contemplated. The need for integration will become even more important at the same time.

Sietec

We have also realised that consulting and not pure hard selling is the only way to boost active use of MT. Companies with too many skeletons (users who feel incorrectly advised) in the closet will enjoy a brief period of success before being buried again by their own strategy.

We are interested in developing standards that reflect the way active users see and foresee the system. It is vital that an MT company listens to its users. A users association is a very necessary sounding board for new ideas. As I said earlier, many of our users have a deep seated interest in seeing not just their installation but the MT world come to life. The biggest mistake would be to take this interest for granted.

Training

We are actively pursuing a policy of integrating our MT system into the course content of degree translation courses in German further education establishments. Within the next two years, translators with MT experience will be coming onto the job market in Germany, saving companies initial training costs and ensuring that these new company employees can be integrated and assimilated into companies' workflow patterns effectively.