MT AT SAP - A SUCCESSFUL INTEGRATION

Dirk Lueke,
SAP AG,
Neurottstr. 16,
69190 Walldorf,
Germany
Tel: +49-6227-34-3256
Fax: +49-6227-34-2798

ABSTRACT

The MT Service Group at SAP currently consists of 8 people permanently employed plus a total of 5 praticants, freelancers, and temporary personnel.

We see ourselves mainly as service providers for our specialized translators, but we also supply preliminary versions and end versions in fields such as user manuals, training material, release notes, error messages, and online texts such as F1 help and hypertext.

Today, machine translation at SAP is fully integrated into the translation process, totally accepted by the internal human translators, has an excellent profitability rate, and is yet looking into new fields of application.

Facts & Figures

SAP was founded in 1972 and today is the world leader in integrated business software solutions with the two products R/2 and R/3. R/2 and R/3 stand for realtime systems of the second and third generation.

SAP is regarded as one of the five biggest software companies in the world with over 4,000 installations, more than 6,500 employees, and 26 subsidiaries worldwide. Over 20 % of the revenue is invested in R&D.

Documentation at SAP covers a variety of documentation types, such as manuals, training material, release notes, marketing brochures, and so on. This documentation is currently translated into 18 different languages in which case the translation depth varies from language to language.

The Translation Department currently holds a staff of 65 English translators (+12 freelancers) and 30 translators (+ 35 freelancers and translation bureaus) for the various other languages and with over 50 technical writers being involved writing in the documentation.

Before the new product R/3 was launched in 1991, it was clear that there would be an increasing demand for large-scale translations in the shortest possible time for the existing R/2 System and for the new R/3 System and with the release periods of upgrades being reduced drastically. For that reason, the MT Service Group, or METAL Group as it was called then, was established in May 1990 for machine translation from German to English using the METAL System by Sietec/Munich.

MT Service Group

The decision to use METAL was largely influenced by the close cooperation between both companies. A previous test in 1982 with Logos was unsuccessful because neither Logos nor SAP were ready for productive machine translation. For Logos, it was the very first installation in Europe and SAP had just hired the first three English translators.

The METAL System and its service group was established with two people working full-time to build up the lexicon, that is, coding the SAP internal terminology, to do extensive tests with the

documentation material of which type of text would be suitable for machine translation, to integrate machine translation into the translation workflow, and to establish the necessary contacts with the specialized departments. Following a 6-month testing phase the management decided to continue the METAL project. The group was later enlarged by another member and together they spent roughly 9 months to set up the machine translation project.

In October 1991, METAL went productive, starting with the translation of the first manuals released for R/3. This was followed by large-scale extractions of texts from the R/3 System, such as F1 help.

By that time not all of the specialized translators were very fond of machine translation, because most of them regarded machine translation as a job killer. It took much time and effort to convince them that the objective of the MT Group was totally different. The group aimed at providing support to the specialized translators for large translation orders, such as manuals, and to supply them with roughly post-edited machine translations if time was running short.

For that matter, the group provided information on the machine translation project right from the start. Small presentations were organized to explain the everyday work and what machine translation is all about. Since at SAP the group still largely depends on the expertise of the specialized translators in the various applications concerning the definition of terminology it was also made sure that the translators were integrated into the machine translation process from the very beginning.

MT Workflow

When machine translation started at SAP, the translation orders to the MT Group were assigned from the coordinator of the English translation. But during the course of time this procedure slightly changed. Today the specialized translators contact the group directly if they need translation help.

The source text is transferred via the network to an exchange directory. From there, the text is picked up and copied into a local directory for further processing. In a first step the text is converted from its original format (either WinWord 6.0 or the SAP-internal SAPscript format) into a METAL interchange format (MDIF). Then the text is deformatted, that is, split up into translation units, and send off to preanalysis. During this step, the text is checked for terminology that is not contained in the METAL dictionary. This terminology is then written into lists that are handed on to the specialized translators for definition. These lists may also contain words that are spelled incorrectly, as they are unknown to METAL, which allows us a feedback to the author, for example, on spelling mistakes.

After the unknown terminology has been defined by the specialized translator, it is coded by the MT Group into the METAL dictionary before the text is actually sent off to translation. The output then is a raw translation which is post-edited within the MT Group. Three different output standards, that is, time and effort spent on post-editing the text, are distinguished depending on the usage of the target text. Is it a preliminary version for information purposes only, a translation proofread by the specialized translator, or even a final translation version ?

The actual time that is required to translate, for example, a manual depends on a variety of factors: the capacity within the MT Group, the quality of the source text, the time the specialized translator needs to define the unknown terminology, and last but not least, technical problems that may occur.

As a rough estimate, you can refer to the following figures:

1992 3 people 2 million words were translated,

1994 4 people 3 million words were translated,

1995 7 people the monthly average is about 500,000 words.

Today, the group is very proud to say that machine translation at SAP is fully integrated into the translation process and totally accepted by our human translators. The G-E dictionary contains 60,000 coded transfer entries (in addition to 45,000 transfer entries that were contained in the dictionary when the system was bought). This allows for high-quality, high-speed translation in all SAP applications and to seek new fields in which machine translation could be profitable.

Future perspectives

There are a number of new areas that were considered in the early days as not being suitable for machine translation. This includes the translation of customer error messages (notes), R/Mail, communication, texts on various topics for information purposes throughout the company, as well as the introduction of new language pairs.

In order to be equally successful in these new areas, certain additional tools and improvements in the existing programs are required. For the notes translation at SAP we had a new program written both by Sietec (for the machine translation) and by SAP (for the export from and import into the R/ 3 System). This allows us to provide a 24-hour translation service for notes translation. We also triggered the development of various macros within the MT Group and together with externals, for example, to compare terminology contained in the METAL dictionaries with the existing SAP terminology and thus be more accurate in the translations, and to process METAL files in WinWord and thus allow the processing of METAL files on the PC rather than on different machines, as it used to be in the past.

Yet, there are still other requirements. For one thing, suitable tools are necessary to import large amounts of terminology or complete reference works into the METAL dictionaries to use machine translation more effectively in a wider range of areas. Here, a powerful defaulter would be helpful. Another aspect is the integration of a translation memory into the machine translation process to handle the change management or to use it for less frequented languages. And, because of the increasing demand, SAP would also like to introduce new language pairs, for example E-G, for machine translation irrespective of the producer.

The most important factor, however, is that all these tools should be operable under ONE interface. It should not matter to the user whether he is using G-E from METAL and E-F from Logos as long as he gets the desired result. The same applies to translation memories and terminology databases. There is also no point in entering terminology in all sorts of databases that serve different purposes. They should all be linked together so that terminology is entered just ONCE regardless of the purpose it is going to serve, that is, if it is used for machine translation (and here again regardless of the underlying engine) or for reference purposes within a company such as SAP.

There are too many spectators in the machine translation field and not enough actors. To make machine translation a successful product, even after Sietec decided to change its strategy concerning the further development of METAL, and to introduce it to a broader audience, we have to have more productive users to be able to considerably influence future developments. Producers should finally acknowledge the growing demands of the users and adapt their tools accordingly. Last but not least do they have to provide support with highly qualified and specially trained staff to guarantee for greatest customer satisfaction.

The future in machine translation lies in the combination of several engines under one common interface. Hardware and software investments went over the top in the past because the different systems were not compatible by any means or in any ways. If machine translation and its adjacent tools are to be successful, they have to address a broader spectrum of users with a variety of tools at a reasonable price and with qualified support instead of just focusing on a handful of large-scale users willing to spend enough money, time and effort to make it a success.

No, machine translation is not magic. It works and if all parties get together it can be even more successful.