

TRANSLATION • INTERNATIONALIZATION • LOCALIZATION • LANGUAGE TECHNOLOGY

104  
3  
X  
2  
1  
4

# MultiLingual

#72 Volume 16 Issue 4

COMPUTING & TECHNOLOGY

## DOING BUSINESS IN CHINA: THE RULES ARE CHANGING

Page 25

## THE GEOPOLITICS OF CONTENT DEVELOPMENT

Page 46

### LOCALIZATION

## LOCALIZATION TESTING FOR EUROPEAN LANGUAGES IN CHINA

Page 37

TM MARKETPLACE

# TRANSLATION MEMORIES: THE DISCOVERY OF ASSETS

*Recognizing opportunities and  
overcoming obstacles to TM sharing*

**JUST ZETZSCHE**

**N**ot this again! I can imagine many readers having just such a response to this headline. After all, the topic of sharing translation memories (TMs) has been discussed again and again over the last five to ten years, not least in the pages of this magazine.

### PROBLEMS

In her excellent article “Sharing Translation Database Information” five years ago (*MultiLingual Computing & Technology* #33 Volume 11 Issue 5), Suzanne Topping outlined the problems that are usually associated with sharing TMs. Although she described the necessity (and the reality) of database exchange, she concluded that the legal ramifications and quality considerations associated with the exchange of TMs were so complex that it would take a long time before a framework was developed that would allow for a legally and qualitatively acceptable exchange of TMs.

Note that her article was partly motivated by a very lively discussion on the TRADOS user list at [http://groups.yahoo.com/group/TW\\_users](http://groups.yahoo.com/group/TW_users). Most of the participants in that discussion were freelance translators or single-language vendors, and the discussion was ignited by one list participant opening his FTP site for an actual TM exchange.

Now fast-forward a few years and change the focus group a little. The LISA 2002 Translation Memory Survey (published in 2003 and available at <http://www.lisa.org/products/survey/2003/tmsurvey.html>) drew its participants from the whole range of representatives of the translation industry with a strong focus on translation vendors and end clients. It

concluded, “At present the notion of a TM asset marketplace is problematic and those surveyed recognize this, with many citing intellectual property and copyright concerns as reasons why they would not exchange TM. Others cited concerns about quality in memories from external sources or said that they would only accept memories developed from materials specific to the same client as their present project.”

This conclusion is virtually identical to Topping’s, and the problems posed for translation vendors are described in very similar terms as those of freelancers: “Given that most contracts specify that clients own the memories developed for their projects it would be impossible for vendors to sell memories without client permission, something most clients would be unlikely to grant given the competitive market value of translations in the world market.”

The final statement of the LISA survey does shed one ray of hope: “Clearly the establishment of a viable marketplace for translation memories will require the solution of present problems, yet some of those surveyed indicated that they felt that the exchange of TM assets will eventually be a huge market; interest on the part of some is quite keen.”

In short, some kind of framework for the exchange of TMs is highly desirable. But legal issues, especially intellectual property (IP) rights, are perceived to be the largest obstacle. Quality issues are considered to be the second biggest problem.

### THE FIRST PROBLEM: LEGALITIES

**W**hat exactly are the rights that are generally associated with IP in regard to TMs?

I am not a legal expert, but it seems to be common sense that the IP of the TM that a freelancer and/or a language vendor assemble while translating for an end client remains with the end client. Even in the unlikely case that the freelancer or the language vendor would successfully argue that they hold a certain IP ownership of the translation that they performed, the IP to the source text part of any TM remains most certainly with the end client. As the LISA report points out, this means that unless the end client specifically gives the freelancer or the language vendor the IP rights to the source text, TM exchange seems impossible from the freelancer’s or language vendor’s perspective.

Interestingly, the central party in all these considerations — the end client — has been almost completely ignored as an active entity in the discussions so far. Part of the reason is that it has taken most end clients several years to catch up with the technological developments in the translation industry. From the end client’s perspective, price breaks due to TM matching have been a welcome and common occurrence in the relationship among freelancers, language vendors and end client; however, the TMs typically remained under the care of the freelancer and/or language vendor. Maintaining TMs under the end client’s control was often considered unnecessary or expensive because it meant the training or hiring of personnel and the acquisition of TM systems that would have to be used for the proprietary database formats of the TMs.

Very few end clients recognized the true value of TM assets early on and controlled those assets, even if the translation work took place externally.

In more recent years, though, things have changed a bit, and many mid-size to larger end clients are increasingly controlling TMs. There are a number of reasons for this. They tend to work with more than one vendor and therefore have to be in control of the TMs if they plan to gain full leverage. They are increasingly aware of the danger of being locked in with a particular vendor and/or tool set and require the delivery of the TMs in a common or exchange format. Tool vendors have been increasingly aggressive in marketing their CAT solutions to end clients. TM exchange formats such as TMX are widely supported and used. And the ongoing discussion about TM ownership has dramatically raised the awareness among many clients of their powerful new asset.

To summarize, a great number of end clients either already hold assets in the form of TMs or could hold them if they made sure that these were passed on to them as part of the deliverables. Many currently use these assets to minimize translation costs and to maintain a certain independence from translation vendors. However, the end clients could astronomically multiply the value of these assets as the only party with the legal standing to share these TM assets with other parties.

## THE SECOND PROBLEM: QUALITY

First of all, what is contained in TMs? In the vast majority of cases and across most industries, TMs contain multilingual content of already-published information. In fact, in a great number of cases the published content is not the actual product itself but descriptive information, such as instructions or lists of ingredients. Though it's unlikely, all of this already-published information could theoretically be collected and turned into TMs through processes such as alignment. The notion that TMs contain highly confidential information which, if shared, might jeopardize business operations is not true in almost every case. While it is possible that truly confidential information could make its way into the main TM that an end client maintains (such as blueprints, memos, or other data), it seems highly doubtful. After all, in general the TMs are maintained to be used by outside vendors.

Furthermore, if the TMs do only contain information that is already published and if the end client and/or translation vendor has followed widely accepted standards of translation, editing and proofreading, the level of quality should be sufficient — certainly sufficient for the original end client to have enough confidence to have this published. Furthermore, most TMs either contain internal documentation about degrees of change

management, subject area and user name or are accompanied by external documents that describe the state and history of the database. While none of the above will serve as a final proof of the quality of the TM content, this in combination with the reputation of the original end client should give a good picture about the level of quality of the databases.

SDL's David Pooley said in Topping's article, "I find it difficult to see how translation databases can be shared. . . . Each client has and will require in their translation a particular style when authoring documents. It is unlikely that there will be much leverage to be gained for clients from different vertical markets due to the disparate subject matter. Even within the same market this is unlikely to change. I don't think that the client would be content to have their documentation and Help in a similar style and terminology to that of their competitors."

In essence, these arguments remind me very much of the pre-TM era, when very similar arguments were made about the creativity of translation. Some people maintained — and still do — that because TM forces translators to write consistently in repetitive situations, the quality of translation would radically diminish and translators would become mere computer experts rather than the linguists whom they were trained to be. Most people today respond to this statement with bewilderment, but my feeling is that a few years from now, we will look at Pooley's statement with similar bewilderment.

Pooley is right to argue that different industries have little overlap, but that's not the point. TMs are much too valuable to be shared aimlessly. There's no reason for a pharmaceutical company to use the TM of a car manufacturer or software vendor. But why would the software vendor, the car manufacturer or the pharmaceutical company not be interested in using an industry-specific TM to reduce translation cost and at the same time provide a terminology and style that is understood across the industry?

Translation of functional texts should be just that — functional. That means it should be as understandable for the target group as it can possibly be (which, by the way, would include a pleasant style of writing). No reasonable translator tries to introduce a new style or terminology just because it is different from a competitor unless the source text makes him or her do that. But in those cases there won't be many matches to start with.

## MICROSOFT GLOSSARIES

Microsoft has been the great visionary in this respect. No other company in any other industry

that I am aware of has published TMs as extensively and as widely used. Throughout the last 12 years, Microsoft has invested significantly into publishing a large portion of its TMs into 44 languages. Its motive for doing this and for regularly updating these databases is to promote consistent terminology in an industry where Microsoft is certainly one of the leaders in terms of innovation and sheer size. The Microsoft glossaries (actually a misnomer because they are in fact TMs) are easily the most widely used reference materials for any Windows-based software localization project today. And while it is difficult to estimate the success of the Microsoft glossaries in measurable unification of software terminology, it is no coincidence that "File," "Edit," "View," and "Help" are generally translated the same across languages for Windows-based applications.

"The publication of our glossaries has been a success story for Microsoft from beginning to end," says Ursula Schwalbach, team lead of one of Microsoft's terminology management groups, based in Redmond, Washington. "We've achieved a remarkable degree of standardization for Windows-based terminology that has resulted in greater clarity for users. And in the process we feel that we've made a contribution toward greater understanding and clarity in the development and translation community. In fact, this success has led us to begin developing an improved and more user-friendly distribution model."

How are the Microsoft glossaries actually used? The market has provided some of those answers. At least four tools currently available are specifically geared toward working with the large CSV files that the Microsoft memories are published in. One of the tools converts the Microsoft memories into a TRADOS-specific text format and another into TMX for easy use as TMs; the other two are index or search tools that are primarily geared toward terminology searches. This means that while the Microsoft TMs are used as TMs (and it is important to note that their CSV formats make them prime candidates to be imported into a great number of TM systems), they are also used as invaluable terminology repositories.

The good news is that there are already a number of terminology mining tools for non-Microsoft TMs as well. Since the advent of TM applications, CAT developers have looked for ways to extend the possibilities of their existing product offerings. TRADOS' ExtraTerm, MultiCorpora's MultiTrans and SDL PhraseFinder offer the processing of TMX memory files to specifically create multilingual glossaries, setting the stage for quick harvesting of terminology out of "foreign" TMs (and other vendors are sure to follow).

TREATING ASSETS AS ASSETS

These examples should have made it sufficiently clear that the time has come to look at your own TMs not just as an important means to streamline translation work, but as assets that can generate a direct return on investment and as being infinitely expandable through competitors' TMs.

Ironically, when I recently talked to an executive of one of the CAT vendors about the idea of offering a marketplace for TMs, he was euphoric about the opportunity to get his hands on his competitors' TMs. However, he was still leery of sharing his own TM because of all the "confidential" information in it.

I suspect that the true nature of his confidentiality concern was more likely his frustration or embarrassment at how much money he had invested into building up his large TM. What other kind of confidential information would there be in the help files, manuals, websites, marketing materials and the user interface of his application? And why would he not be willing to financially profit from

licensing his TM to his competitors (and in the process help the consumer by working toward a more consistent terminology)?

Everyone tends to have knee-jerk reactions when it comes to certain topics. The sharing of TMs tends to cause this knee-jerking for many. Knees should calm considerably, though, when one considers the price of translation, editing, proofreading and multiple reuse of every single one of those thousands or millions of translation units in a TM — and the possibility of now turning this into a marketable asset.

Earlier this year, MultiLingual Computing, Inc., and International Writers' Group joined together to create a company, TM Marketplace, that offers an exchange forum for TMs. TM Marketplace is what its name implies: a marketplace for end clients to license their TMs to competitors or other interested parties such as translation vendors. More information can be found at <http://www.tmmarketplace.com>

New concepts are exciting. This concept, however, may be more than "just exciting." Some of the immediate or intermediary results that this could have for our industry could be a further development of TM as well as linguistic standards, the acknowledgment of similarities and common interests between end-clients, the re-definition of client-vendor relations and new frontiers for tool makers.

Most of all — and this is beyond speculation and wishful thinking — TMs have finally become true assets. 🌐

.....

*Jost Zetzsche is a localization consultant, translator, cofounder of International Writers' Group and the new TM Marketplace, LLC. He can be reached at [jost@tmmarketplace.com](mailto:jost@tmmarketplace.com)*



TRANSLATION MEMORY'S NEW ERA OF SHARING

TM MARKETPLACE

A CONCEPT WHOSE TIME HAS ARRIVED

As translation memory (TM) brokers, TM Marketplace connects corporate owners of translation assets with parties who want to license and benefit from those linguistic resources.

TM Owners — You now have an opportunity to receive an immediate return on your investment through licensing

TM Buyers — You now have the opportunity to access proven, industry-specific translations

For details on our services and how we ensure success for both parties, visit [www.tmmarketplace.com](http://www.tmmarketplace.com) or contact us at [info@tmmarketplace.com](mailto:info@tmmarketplace.com)



TM MARKETPLACE, LLC  
208 265 9465  
US TOLL FREE 888 533 7886  
[INFO@TMMARKETPLACE.COM](mailto:INFO@TMMARKETPLACE.COM)

[www.tmmarketplace.com](http://www.tmmarketplace.com)