

**Technical Seminar on Copyright, Intellectual Property and Translation Tools  
Barcelone, 13 October 2007**

---

**Intellectual Property and CAT: the PwC Experience**

Ross Smith, Head of Translation Services, PriceWaterhouseCoopers, Madrid

The purpose of this paper is to study intellectual property issues in relation to CAT tools at PricewaterhouseCoopers (PwC), the world's leading professional services firm. We will focus mainly on policies and procedures at PwC Madrid, but reference will be made to other offices in Europe where possible, particularly PwC Paris, where extensive use is also made of TM technology.

**TM Organisation at PwC Madrid**

In order to analyse the situation with respect to the ownership of translation memories, it is advisable to start by explaining how translation memory files are organised at PwC Madrid. All translation memories (TMs) are generated using SDL-Trados software, specifically SDL-Trados 2006 Professional, and can be separated into three main levels. On the first level is a gigantic and somewhat untidy Main Memory, which has been in use since we started utilising Trados in the 1990s. This memory is used for accounting documents in general anything which does not fall into a specific category.

On the next level are four subject-specific memories, used for the main thematic areas which lend themselves to being translated using computerised support tools. These are:

- Annual accounts (also commonly called financial statements, i.e. the balance sheet, income statement, cash flow statement, notes to the accounts, etc.) of business corporations which are drawn up in accordance with International Financial Reporting Standards (IFRS)
- Annual accounts and other documentation specifically pertaining to financial institutions: banks, savings banks, ratings agencies, etc.
- Legal documentation such as contracts, minutes, rulings, appeals, etc.
- Audit opinions issued by PwC.

On the third level we find specific individual memories which are maintained for translations of the annual accounts of the major PwC audit clients, or for certain large-scale projects. Annual accounts make up around 75% of the total volume of work regularly carried out using our TM tools and under the new IFRS requirements, these documents can be very large. It is advisable to keep separate memories for them to assure that client-specific terminology is applied and so that they can be transferred easily from one translator to another if necessary, since not all members of the translation service are on the PwC network with direct access to the translation memories.

SDL Trados is used by both in-house and freelance translators for suitable documents. The most commonly used components are the Workbench and TagEditor, along with the MultiTerm terminology bank and WinAlign alignment tool. It is also the main software used at the Paris office.

### **Different scenarios for document ownership**

The circumstances concerning the origin of documents used to initiate or build TMs at PwC, and the ownership of the translations once these are produced, vary considerably and therefore making decisions as to who holds the relevant intellectual property (IP) rights to them is not always easy. In general terms the following scenarios can be identified:

1. The first scenario is perhaps the most typical: the source document is written within a PwC department by an employee of the Firm (most probably in the Audit division) and is translated by another employee in the Translation Department using SDL-Trados tools acquired under licence by PwC. The PwC line of service concerned (legal, audit or advisory) is then invoiced for the translation by the translation service (which belongs to the Legal division). In this case, therefore, the document is authored, translated and purchased by parties within the firm using CAT technology licensed to PwC.
2. It is also common, however, for the source document to be written by a client of the firm, particularly in the case of annual accounts. In accordance with current accounting regulations, auditors can assist clients in the drafting of their accounts but under no circumstances may they actually write the accounts for the client. In such cases, therefore, the documentation belongs to the client, which gives PwC a copy of the document, without any transfer of ownership, purely for the purpose of having it translated, usually from Spanish to English. The documentation arrives at the translation service directly from the client or, more commonly, via a member of the audit department. The translation is performed by a PwC employee and the translated document is sold to the client either as part of the overall audit service or as an additional service. Accordingly, the source and target documents both belong to the client, as author in the first case and as buyer in the second, while the translation memory itself is created by the PwC in-house translator who translates the document (as well as any segments already created in previous translations for the same client). To complicate things further, there are instances in which the documents provided by the client are not actually drafted by the client: an administrative judgement or ruling, for instance, or documents submitted to a court in an appeal process. Here, the documents are in the client's possession, but they are not "owned" by the client in the sense of its having either created or purchased them.
3. It can also happen that a client of our Audit department wishes to have a translation of certain financial documentation for a given period, and in order to help in the translation process, and also to assure consistency, they provide PwC with the previous year's documents and a translation thereof made by another translator or perhaps by the client itself. Alternatively, such material can often be freely obtained from clients' websites and can be utilised by translators to resolve doubts and guarantee consistency. To take advantage of this supporting documentation, the previous year's source document and translation are aligned (in our case using SDL Trados WinAlign) and the resulting file is exported to create a specific translation memory. In this case, therefore, the source and target documents belong to the client as author and buyer, and the entire content of the TM created by the alignment tool derives from client-owned documents as well, at least at the moment when the translation commences.

4. Finally, there is the situation in which the Translation Department provides an external translator or language service provider with a copy of a main TM (probably the one used to translate corporations' or banks' financial statements) or a client-specific TM to help the translator and to assure consistency in style and terminology. In such cases, the external parties concerned are told that the TM is the property of PwC and they are reminded that they are being provided with the TM for their use only, for work commissioned by PwC, and that it is not transferable. We do not ask for any kind of payment, since it is in our interest that the translator concerned should use the TM for the reasons of style and consistency already mentioned. The situation here, therefore, is that the source document is authored either by a PwC staff member or a PwC client; the TM belongs to PwC but the translation is performed by a third party freelancer; and the final document, in the first instance, is purchased from the external translator by PwC's translation service, and therefore belongs to PwC, though it will usually be passed on and invoiced to the client concerned.

All these scenarios are possible and occur in practice. It should be noted, however, that in the history of the PwC translation departments in Madrid and Paris, no client has ever asked to be given the translation memory used to translate its documents as well as the translation itself.

### **Who owns what in these scenarios?**

Only in the first of the cases described above does the ownership issue seem clear. All of the elements in the translation process – the source document, the translated document and the translation memory itself – are created by members of PwC's staff using computer tools licensed to the firm. We may safely say here, therefore – to the extent that one can safely say anything when IP is involved – that PricewaterhouseCoopers Madrid holds the intellectual property rights to the TM and its content.

In the remaining three cases, things are not so easy. With regard to situations such as that described in scenario 2 (the client creates or furnishes the source document and purchases the translated document, while an in-house translator uses the TM), which are probably the most interesting for our present purposes, the opinions of experts concerning the ownership of intellectual property rights differ radically.

On the one hand, it may be argued (based on certain aspects of European IP legislation) that the TM belongs to the translator simply by virtue of the fact that the translator makes the memory through his own intellectual effort and expertise. This is the "I made it, so it's mine" approach which most translators regard as irrefutable<sup>1</sup>. There are further arguments, put forward by the Spanish lawyer Jorge Marcos<sup>2</sup> and also by the first company to set up a market for translation memories, TM Marketplace<sup>3</sup>, to the effect that since the source document is segmented and atomised into such small particles within the TM, there is nothing large enough in the content of the TM to merit IP protection. As Marcos puts it: "Except in extraordinary cases, it is unlikely that a sentence can be an object of intellectual property". In addition, this segmentation transforms the source document to such an extent that the party performing the transformation – the translator – becomes the intellectual owner of the newly created item.

This approach is broadly supported by PwC. In the opinion of Javier Ribas, a partner at Landwell (PricewaterhouseCoopers correspondent law firm) and one of Spain's leading specialists in the intellectual property field, the memory belongs to the translator, if an independent contractor, or to the translator's employer, by virtue of the fact that the segments it contains are mostly standard or general in nature (if not, they would not be of much use in other future translations) and because single sentences (except for advertising slogans, etc.) virtually never qualify for IP protection. Ribas considers

that the applicable type of IP protection is that pertaining to databases: what is being covered by the law is not the author's creative effort, but the investment in time and money made to compile and store the most frequently used sentences and phrases.

On the other hand, there may be grounds for claiming that the TM belongs to the party that actually commissions the translation. For instance, the well-known localisation specialist Jost Zetzshe, writing in the journal *Multilingual Computing and Technology*<sup>4</sup>, states that: "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." This could be called the "I'm paying for it, so it's mine" approach, which is regarded as logical by some language service users. Their argument, in essence, is that the segments used to create the translated work cannot be differentiated from the work itself. From a translator's perspective, this approach is rather like a brick manufacturer claiming that he owns a newly constructed building because it is made with his bricks, yet from the client's viewpoint it seems quite coherent. Another specialist lawyer consulted for this paper also pointed out that, in fact, IP legislation refers to "an original creation expressed in any medium", but says nothing about the size of the creation. In principle, therefore, a single sentence, or phrase, *could* be an object of intellectual property, though it is difficult to think of any examples outside advertising slogans and minimalist poetry.

In scenario 3, not only the source and target documents derive from the client, but also the content of the TM since is generated by aligning the client's documentation. Surely, in this case, the TM must also belong to the client? Not necessarily, according to the above-mentioned White Paper by TM Marketplace, due to the arguments concerning atomisation and transformation. In any event, even if at the start of the translation the TM content derives wholly from client documentation, as soon as the translator gets down to work new translation segments will be added and existing segments will be modified. It could be advocated that in such cases the best solution is the joint ownership of the TM though again, this could be contested by both the end purchaser based on the "I'm paying for the whole job" approach and by the translator (in addition to the segmentation issues), who may come back to the "sweat of one's own brow" argument, claiming that the alignment process is an intellectual task which in itself transforms the original work beyond anything that could be covered by IP rights.

Scenario 4 is becoming increasingly common, as translation departments and companies supply TMs to external professionals with a view to assuring consistency and quality. Unless there is a specific contract on IP rights to the TM (see below), the most logical solution seems to be that ownership is shared, the content of the initially supplied TM belonging to the supplier, and any segments added thereafter belonging to the freelancer.

It must be pointed out, of course, that most of these issues can be resolved from the outset by arranging a contract between translator/language service provider and the client which clearly specifies the ownership of the TM concerned, or which allows the work being translated to be freely transformed, reproduced or distributed. This matter has been stressed by all the legal specialists consulted. However, as we all know, such contracts are not usually signed, for a variety of reasons. In many cases, translators do not want clients to know they use TM tools, regarding this area as literally none of the client's business. In others, the translator might be wary of frightening a client away by appearing to require undesirably complicated contractual conditions for the job (most clients will probably not know what a translation memory is and might suspect that somehow or other they are being asked to sign away their rights to the subsequent exploitation of the translation they are paying for), and would therefore rather not bring the subject up. For whatever reason, contractual cover is not usually available and therefore the ambiguities described above persist.

## **Selling or licensing TMs**

In principle, the TMs created by PwC Madrid for the Spanish-English translation of banking and IFRS financial documents could be of considerable value to freelance translators and language service providers, and might therefore be worth selling or licensing. To date, however, this issue has never been considered as a real option because our CAT software is regarded as a tool for assisting translators, not a means of creating a saleable commodity. This view could change in the future, however, if all potential intellectual property issues, and confidentiality concerns, are clearly resolved. The market is already moving in this direction, thanks particularly to the creation of the TM Marketplace website where in theory our translation memories could be exploited for commercial purposes.

We must acknowledge that at PwC there is one very large barrier to such a move, this being our image vis-à-vis our clients and the auditing/consultancy market in general. Clients' names and financial data will figure, unavoidably, in the sector-specific memories mentioned above and even though the source documents used to create these TMs are publicly available, we cannot risk a major client causing a fuss because they perceive that we are somehow disclosing their internal information without express permission, or using their documentation to generate cash without giving anything in return. For the reasons explained earlier this perception would almost certainly be false but the harm would have been done in any case merely through a suspicion of dishonest practice or potential disclosure of sensitive data, since image is a vital factor in the professional services sector.

If trading in and licensing TMs becomes a common and accepted practice in the language services industry then PwC could become involved on an active basis, but until then we will probably have to restrict ourselves to watching how the market develops from a prudent distance.

## **Payment for translations using TM tools**

The policy at PwC's Madrid and Paris offices in general is not to adjust prices solely on the basis of the technology used to translate documents. We consider that the purpose of CAT tools is to help translators reduce the volume of repetitive, non-specialised work so that they can concentrate their energy on the actual task of translating, and to foment quality and consistency. Their purpose is not to force translators to cut their rates. We are against pricing systems which involve different rates based purely on the list of percentages generated by the Trados Workbench (e.g. 100% matches no charge, 75-95% half rate, below 75%, full rate), since this fails to take into account a range of relevant issues, particularly the question of the varying reliability of the TMs themselves. Having said that, we have always considered it reasonable for a client to request a discount if there is a notable volume of repeated text in two or more documents, and we maintain this approach when using TM tools.

Therefore, in the event that a certain volume of text can be recycled (for instance, in a lengthy appeal process going through the courts with a large amount of text being reused in the procedural documentation in each judicial instance, or in the case of a company's accounts which are largely unchanged from one year to the next), we are prepared to take this issue into account, but without making things too complicated. For instance, we may charge 85% - 100% matches at our review rate (just under half the translation rate) and the rest at our normal translation rate. This is because, in our opinion, 100% matches can rarely be trusted completely and should always be reviewed, particularly in TMs in which various translators have been users. If we apply this policy with our own client and the translation is subcontracted, we ask the external translator concerned to invoice us on the same basis.

---

<sup>1</sup> As reflected in the survey edited by Antonio Valderrábanos titled *Who owns what? Some Insights on TM Ownership* ([www.transref.org](http://www.transref.org))

<sup>2</sup> Marcos, Jorge: *Un enfocament jurídic de les Memòries de Traducció* ([Hwww.fti.uab.es/tradumatica/revista/num0/sumari/sumari.htm](http://www.fti.uab.es/tradumatica/revista/num0/sumari/sumari.htm))

<sup>3</sup> TM Marketplace: "*White Paper*" on sharing translation memory data ([www.tmmarketplace.com/whitepapers/align](http://www.tmmarketplace.com/whitepapers/align))

<sup>4</sup> Zetsche, Jost: *Translation Memories: the Discovery of Assets in Multilingual Computing and Technology* (Volume 16, issue 4) ([Hwww.multilingual.com](http://www.multilingual.com))