

Jost Zetzsche

Language as a commodity



I'm not sure that the language industry is particularly homogeneous, but there are certain concepts that most of us agree upon so unanimously that we don't even think twice about them. The concept of "language as a commodity" is a great example.

A few minutes of Google searching turn up representatives of independent industry analysts, tool vendors, CEOs of large multi-language vendors and freelance translators united in saying that "commodity" and "translation" should have nothing in common and, in fact, that "the [present] image of translation is not positive, it is viewed as a commodity" (see Fernand Boucau of the Belgian Quality Translation Association in www.euatc.org/conferences/pdfs/boucau.pdf). Numerous well-known language vendors announce prominently on their websites that they offer "services" rather than "commodities" (implying, of course, that their competition does the exact opposite), and they are supported by a Forrester Report of 2000 (*The Multilingual Site Blueprint* by Eric Schmitt) that urges translation buyers to "rethink translation service requests for proposals" because "translation services are not a commodity."

So, the concept of translation as a commodity is so strongly despised because a) translation is not a ready-made product but a service and b) there is a great variation in price and quality of translation throughout the language supply chain. The emphasis is therefore on the quality and price of the service provider or the service itself and the quality and price of the finished product.

Brazilian translator Danilo Nogueira defines this more clearly: "Translation is a service business, not an industry or commerce. The basic difference between industry, commerce and services lies in *inventories*. Industrial establishments keep at least two kinds of inventory: raw materials and finished goods. Commercial establishments keep only finished goods inventories. Service establishments, however, keep no inventories" (see <http://accurapid.com/journal/06xlat1.htm>).

This article was published in 1998. Knowing that Nogueira is a passionate user of translation technology, I wonder whether he would still stand by this eight years later. The fact is that the "by-product" of databases of previously translated text (translation memories or TMs) is exactly what he describes: an inventory of raw materials and – in the ideal case – finished goods, that is, translated projects.

By using the now-so-common TM technology, we have almost accidentally created a whole new entity in our supply chains: data that could indeed become – I'm almost embarrassed to say it – a commodity!

You probably noted the conditional *could* in the previous sentence. Whether it is the fear of commoditization or a simple fear of change, we seem to be intent on keeping the TM data to ourselves.

And yet, there are plenty of scenarios for using our data.

- The most promising branch of machine translation (MT) technology today is statistical-based MT, a technology that allows the translation engine to "learn" languages by processing very, very large amounts of bilingual data (that is, TMs).

- It has long been recognized that a true breakthrough with TM technology can only be achieved if the authoring process itself is database-dependent as well, assuring not only greater consistency but, more important in our context, the tailoring of the source to a very high percentage of matches in the translation. The technology to achieve this is available (such as Sajan's Authoring Coach TMX or SDL's AuthorAssistant), and the data is already sitting in TMs, but it is unused beyond its original owner.

- Terminology research is one of the more time-intensive and cost-intensive aspects of the translation/localization process, and translated terminology is re-invented in all its possible variations over and over, despite the fact that the source terminology is often identical. Again, technology for "harvesting terminology" out of TMs is available, and the data could be as well.

- And clearly the most obvious way to use TMs is to use them as they were intended: repositories of data that can be leveraged in perfect and fuzzy matches in any number of ways.

Services should not be commoditized, and the language industry as a whole needs to focus on high-level quality services if it wants to gain its long-deserved recognition as an industry. But to be recognized as an equal partner by other industries, we must act on opportunities if we see them. Acknowledging that TM data should be a commodity is one opportunity that cannot be ignored. **M**

Jost Zetzsche is a translator, co-founder of International Writers' Group, author of The Translator's Toolbox, publisher of the Toolbox newsletter, and founder of TM Marketplace.

Takeaway is an opportunity to comment on language-industry issues. Contributions are welcome. Send them to editor@multilingual.com