

# The new slate in 'translate'

## Emma Goldsmith talks to Tom Hoar, the man behind Slate Desktop, about this new technological development for translators



Emma Goldsmith MITI is a freelance Spanish to English medical translator who made a career switch from nursing some 25 years ago. She writes a blog called *Signs & Symptoms of Translation* about medical translation and SDL Trados Studio.

Slate Desktop is a Windows application that was due for release around the time this issue of *ITI Bulletin* went to press. It has been designed to integrate with CAT tools to provide machine translation suggestions from an engine created on your local machine, based entirely on your own translation memories.

Tom Hoar is the founder and chief executive of Precision Translation Tools, the company behind Slate Desktop. I asked Tom if he would explain to *ITI Bulletin* readers in more detail how Slate Desktop will work, and also tell us more about the novel way it was marketed and pre-sold in the months leading up to its release.

**EG:** First of all, Tom, can you tell us what makes Slate Desktop different from cloud-based machine translation tools such as Google Translate?

**TH:** Let's start with what's the same. Slate Desktop and most cloud-based machine translation service providers (MTSPs), such as Google Translate, rely on statistical machine translation technologies. From there, Slate Desktop and MTSPs diverge significantly in both intention and design. Cloud-based machine translation engines serve millions of users across a broad range of subjects. They're designed to serve anyone with ease and convenience. Translators, for example, can simply upload jobs and receive translated suggestions through their CAT tools. However, the quality is debatable and, more importantly, confidentiality agreements often prohibit any use whatsoever of cloud-based machine translation. Some cloud-based engines can be customised, leading to improved quality for specific subjects, but customisation requires users to share their translation memories with the MTSP. Translation memories are not only confidential, but also strategic assets. Therefore, many translators and agencies abstain from customisation to protect their translation memories. In contrast, a Slate Desktop engine serves one translator working in one or a few subject areas. Translators can convert their translation memories into customised translation engines. The engines then generate suggestions that accurately reflect the focus, quality and consistency of the translation memories. Confidentiality

agreements and strategic translation memories are safe because all the work stays on the translator's computer.

**EG:** Will Slate Desktop work for all languages? And all language variants?

**TH:** On release, Slate Desktop will support 29 languages. That means translators can create engines across 812 language pairs. We will automatically add support for new languages with free updates as we develop the necessary 'tokeniser' utilities. There are several approaches for managing language variants. The simplest approach relies on the content of your translation memories. For example, if you want suggestions in British English, simply create an engine from translation memories with British English only as the target language.

**EG:** Using our own translation memories sounds very exciting for translators who have years of resources under their belts. But how useful will Slate Desktop be for people who are just starting out?

**TH:** Slate Desktop might not be very useful for new translators because it takes time for translators to build their own inventories. We estimate that roughly 130,000 segments are required to create reliable statistical machine translation engines. However, some customers report good experiences with as few as 50,000 segments, particularly if they work in a very specialised

field. Translators without their own translation memories can use publicly available ones, such as the DGT corpus, but then Slate Desktop mimics that writing style.

**EG:** Many CAT tools already have smart engines that repair or patch fuzzy matches from translation memories in real time, drawing on a variety of resources to do so. How will Slate Desktop do this better?

**TH:** I'm aware of these new features, but we haven't tested any of them. I think they also go by names such as 'subsegment matching' and 'auto-assembly'. So my answer here is mostly speculation, and I welcome readers' feedback on this point, but here goes: these features use language-specific rules to break apart translation memory segments and identify phrases. Then, more language-specific rules are applied to splice new suggestions. I perceive these features as a type of rules-based machine translation. That's great if you're working with languages that support these rules. But statistical machine translation is the ultimate subsegment match and auto-assembly tool. It uses machine-learning techniques to identify rules, instead of language-specific rules crafted by hand. This makes it accessible across many more languages. It will be interesting to see how the two compare.

**EG:** Will Slate Desktop be able to give priority to certain terms or glossaries?

**'Once configured, the CAT tool will present machine translation suggestions from your Slate Desktop engine like any other machine translation provider'**



**TH:** You can set up Slate Desktop to force specific translations – just create a table with side-by-side source and target language terms. Slate Desktop will use the table entries to override the engine's choice.

**EG:** What happens when Slate Desktop comes across new terminology in a source text? Will it look for the nearest match? Will it leave a blank?

**TH:** Words that are missing from your translation memories altogether are called out-of-vocabulary (OOV) words. During translation, the engine uses the words it knows, and OOV words are passed through to the target sentence as source language words. You can add newly translated OOV words to your forced translation table to reduce repetitive corrections.

**EG:** How easy will it be to set up Slate Desktop? Will non-tech-savvy translators be able to get it to work?

**TH:** To create an engine, you first export your translation memories, glossaries and other bilingual resources to TMX or tab-delimited files and import them into Slate Desktop. Then, you label imported segments by subject or other identifier. Finally, you define which labels to use to create each subject-customised engine. To run the engine, you add its machine translation connector to your CAT tool. Once configured, the CAT tool will present machine translation suggestions from your Slate Desktop engine like any other machine translation provider.

**EG:** As you know, I was one of the backers of your Indiegogo crowdfunding campaign for Slate Desktop. Can you explain to *ITI Bulletin* readers what Indiegogo is and why you decided to market your product through crowdfunding?

**TH:** Thank you for your backing! Slate Desktop is the first Windows application of its kind, so we didn't know what to expect from potential customers. A novel application deserved a novel marketing strategy to quickly test the market and learn from the experience. I considered Survey Monkey, expert interviews and other research techniques, but nothing tests a market like customers spending real money. We could have set up an online store on our own website, but why would prospective customers have trusted us? I learned that some companies use Indiegogo as a test market platform. Instead of trying to raise a large sum, we focused on learning how to actively engage translators with our new product. We didn't know what to expect in terms of sales. We set the \$6,500 goal to test a reasonable sales volume of about 20 units in 60 days.

**EG:** In addition to the \$6,500 target, I saw that you had some 'stretch goals' that later disappeared from the campaign page. Does that mean the response wasn't as good as you'd initially hoped?



**'You can set up Slate Desktop to force specific translations – just create a table with side-by-side source and target language terms'**

**TH:** From the start, we designed two versions of the sales message. We ran one for the first 30 days with stretch goals to see if they would inspire more sales. We dropped the stretch goals for the second half of the campaign because the first half showed we needed to focus on simplifying our message. It was a great learning experience and we even overshot our money goal by 40 per cent at the very end. We have some great new customers and we learned what translators are looking for. Now that the market knows we're here, we have a lot to live up to!

**EG:** Campaign backers had to pledge for a product they hadn't tested. I'm sure some translators weren't prepared to take that risk. Will a trial version be available when the product is released?

**TH:** I agree. Most consumers want to test new software before buying it. Yes, in the long run, we will offer a trial version. Along the way, we'll try a variety of promotions.

**EG:** How do you see the tool being developed, once it's on the market?

**TH:** It's not only about how we develop the tool. Slate Desktop includes a programmable application programming interface (API). Translators can build extended features, inspired by their language skills and their own creativity. Our development will focus on language-independent features and improving the shared framework. We'll also create online resources for translators to share.

**EG:** How do you see the future of machine translation?

**TH:** I'd like to see 'machine translation' become a legacy term. It's an antiquated term that comes with 65 years of emotional baggage – some of which is justified – but we are living in a modern age where software has been abstracted from the machine. I believe quality estimation is the biggest remaining challenge in translation software, but the unknown is exciting to me. I think we will see some unpredictable advances in this field as translators start to experiment with this democratised technology on their desktops.

**EG:** Thank you very much for answering these questions. I look forward to putting Slate Desktop through its paces!

**TH:** Thank you. I've enjoyed our exchanges as we've got to know each other through the Indiegogo campaign. Several other ITI members are customers of our products and it's always fun working with them. I look forward to continued growth with the ITI membership. 

**Tom Hoar is a language technology veteran with 30 years of experience and founded Precision Translation Tools in 2010 with the vision of simplifying statistical machine translation technology for professional translators. Email: [tahoar@precisiontranslationtools.com](mailto:tahoar@precisiontranslationtools.com). For more details about Slate Desktop, visit [www.precisiontranslationtools.com/slate-desktop](http://www.precisiontranslationtools.com/slate-desktop).**