Install Updates for “Translating Tibetan”? 
A Presentation for the Technology Panel at Tsadra’s “Translation & Transmission” Conference, 2017

This presentation is about “installing updates” for translating Tibetan—how do we translate using technology skilfully?

First, let’s simply note that there is a “suffering of translating”—translating is not an easy task, and we all have experienced the difficulty of translating Tibetan!

I imagine you all can come up with many personal examples of “suffering” while translating...

The next step in IT (information technology) is, of course, troubleshooting or diagnosing the issues. In other words, what’s causing our suffering? What issues lie at the heart of the problem? What’s giving us trouble?
Here, I’d like to point out a tendency we have to cling to our “legacy systems”—in IT parlance, a “legacy system” is a system that, to put it nicely, is nostalgic or quaint—to put it less kindly, it is a system that is outdated.

And I’d like to divide these “legacy systems” into two, the first of which is the “technology” side. Technologically, very few Tibetan translators use anything more sophisticated than a word processor, like Microsoft Word.

That means we only have access to the plain text of the document; we only have access to it on our local machine; and any tool we use, we have to use it manually (by physically typing a word into the search bar, for example).

Word processors are, of course, very useful tools—but these have existed since the 1960s. Your smartphone has hundreds of times more processing power than that, and your computer thousands of times more!

If all you’re using to translate is a word processor, then you’re using an outdated legacy system. You aren’t making use of the power new processors have to analyze the much larger data sets your drive can now store...
The second kind of **legacy system** we’re clinging to is not technological, but human. The human legacy system responsible for quite a lot of suffering in Tibetan translation is the **Grammar-Translation** pedagogy we’re using to teach the Tibetan language.

This pedagogy, or teaching method, rose to prominence in 17th century Prussia. It has been well overturned by modern second language teaching techniques for hundreds of years, yet has managed to survive in isolated pockets of second language learning (like Tibetan)... Modern teaching methods would go a long ways toward giving future translators the **real language skills** they need to translate!

Finally, our reliance on English as the predominant target language, as well as the language in which we discuss points of **Tibetan** language, makes little sense. Instead, using **Tibetan** as a communicative **lingua franca** for translating **Tibetan texts** has the potential to open translators to a whole world of native-scholar expertise...

Now that we’ve diagnosed the issues, it’s time to suggest some practical, easy-to-implement solutions for the Tibetan translator.
“Have you tried turning it off and then back on again?” — the first suggestion of anyone experienced IT, and I think it’s applicable here.

All we need to do to solve the problem (and perhaps it’s easier said than done) is delete our legacy systems, install updates, and re-boot—to turn everything off and then turn it back on again!

Let’s take a look at the updates we need to “install” for technology first:

For one, we need to start using etext formats with markup to take full advantage of modern resources that are common in other languages like English—for us, that means our Tibetan IT specialists need to perfect word spacing, POS (part of speech) tags, and lemmatization for Tibetan.

Next, we need to use shared, global formats and platforms for our source texts, as well as our translations. Using a modern CAT tool (like smartCAT) goes a long way in this regard, and I’ll speak more to that in just a minute.

Finally, we need to avail ourselves to the tools of modern digital humanities. That means familiarizing ourselves with corpus search tools for reference; using a simple dictionary in this day and age is a bit like being a biologist who does all their
research at the museum on jars of preserved specimens. Corpus tools give us rich content and contexts—the word in its natural habitat, rather than as preserved by a collector for posterity!

(I’ll be uploading tutorials on how to use Voyant and AntConc for that in the future, so subscribe and stay tuned)...

Second, we need to install updates for the human side of our work, for the human part of language and translation.

That means using modern teaching methods, and modern translator training; we need programs that recognize what natural languages are, and how human beings learn them, and what it means to “translate” in this day and age.

Finally, we need to start using Tibetan as a tool for communication. We need to start communicating with the communities who hold the texts we wish to translate, not in our language on our terms, but in their language, and on their terms...

If we want to call ourselves “Tibetan translators”, in other words, we should learn how to speak Tibetan...

While I’ve gone into the education side of things in past presentations (most notably IATS 2016), since this is a technology panel, let’s focus on the how-to for that side of things today.

How do we begin using modern technology in order to translate Tibetan? What’s our “path” to installing technology updates?
Let’s start by introducing the basic software update—instead of word processing, we can use CAT tools! CAT stands for “Computer Assisted Translation.” Briefly put, CAT tools draw on TM—Translation Memories—to make automatic suggestions during the translation process. These suggestions are based on your past work!

In addition to this phrase-level, automated help, any glossaries you upload or terms you enter are consulted automatically. If a glossary term shows up in a segment, the CAT tool displays it. No need to search!

Finally, these tools are interactive and live; real-time collaboration is possible from anywhere.

In detail, it works like this: All your old translations are stored, segment by segment, in the TM (Translation Memory). If a new segment you are translating is unique, you will translate it—and the CAT will automatically store it in your TM for reference during future translations!

However, if you’ve already translated a similar segment, you can insert it from your TM—and make the necessary edits to improve it. The new segment is again stored for the future.

Finally, there will be times you’ve translated exactly the same segment before; simply insert the TM, and continue to the next segment.
In the long-run, especially over large, multi-year projects, and especially for full-time translators, using CAT tools can save you 40-50% of your time.

Language comes in patterns, and CAT is very adept at recognizing and re-presenting those patterns for you!

All we are doing, really, is leveraging what machines are good at to our advantage. Machines are good at processing lots of data, very quickly; they’re good at repetitious, boring work; they’re good at matching patterns, crunching numbers, and being consistent...

All things, by the way, that human beings tend to be bad at. Let’s let the machines do the arduous and boring tasks that cause us so much suffering in the translation process!

We can also ask ourselves: What are human beings good at?

Human beings are good at having human experiences; we’re good at empathizing with other people’s points of view, and understanding “where they’re coming from”; we’re good at decoding socio-cultural cues and subtactual messages; finally, we’re also good at getting a feel for how things sound—and how they sound “good” or “right”, as well as how things sound “wrong” or “bad.”

Machines, at least ‘til now, still aren’t good at these things. Let’s do the human things that we, as humans, are naturally good at! This will also allay our suffering...
If we want to maximize our capacity for translation, we should leave computers the tasks they are good at, and take on those skills we can develop, that machines can’t!

As human beings who want to translate Tibetan, we should be speaking Tibetan with other humans, and especially with native speakers in native cultural contexts.

And we should be letting machines memorize terms for consistency, and allow them to find and match patterns for us in our translation workflow.

**If we let machines do machine things and humans do human things, we will suffer a lot less to translate Tibetan!**

Let’s take a closer look at the tools and features of **smartCAT** for Tibetan.

For the YouTube demo, please follow this link: [https://youtu.be/vfELrgS1ms0](https://youtu.be/vfELrgS1ms0)