Twitter Thread by Katrina Leonoudakis ■ Miyamura Appreciation Club



Katrina Leonoudakis ■ Miyamura Appreciation Club
@katrinaltrnsl8r



Machine translation can be a wonderful translation tool, but its uses are widely misunderstood.

Let's talk about Google Translate, its current state in the professional translation industry, and why robots are terrible at interpreting culture and context.

Straight to the point: machine translation (MT) is an incredibly helpful tool for translation! But just like any tool, there are specific times and places for it.

You wouldn't use a jackhammer to nail a painting to the wall.

Two factors are at play when determining how useful MT is: language pair and context.

Certain language pairs are better suited for MT. Typically, the more similar the grammar structure, the better the MT will be. Think Spanish <> Portuguese vs. Spanish <> Japanese.

No two MT engines are the same, though! Check out how human professionals ranked their choice of MT engine in a Phrase survey:
https://t.co/yiVPmHnjKv

When it comes to context, the first thing to look at is the type of text you want to translate. Typically, the more technical and straightforward the text, the better a machine will be at working on it.

There are plenty of industries that leverage machine translation semi-successfully: technical documentation, spec documents, patents and legal documents in certain language pairs, etc.

Of course, machine translation isn't perfect. Language is complicated, and there are tons of little obscure grammar and vocab rules that are difficult to program in.

Think: how English has assigned a different word for every "flock "of animals (murder of crows) or Japanese's system of counters (did you know that octopi and squid are counted with ■?)

Machine translation post-editing (MTPE) has become a huge part of the translation industry in the last decade, largely in the more technical fields.

In her presentation Standardized Competencies for the Professional Practice of Localization Project Management at ATA 2019, MIIS Professor Alaina Brandt stated that MT is widely undervalued in the industry, despite placing in the top 4 top translator skills in a survey of PMs.

Unfortunately, many language service providers (LSPs) have begun to leverage MTPE as cost-cutting measures for translation. They can charge their clients the same amount, but pay their translators less to "post-edit" a machine translation.

As we've already discussed, depending on what's being translated and the languages involved, the human translator either gets something fairly OK or a dripping bag of alphabet soup.

Poor applications of MTPE make human translators miserable--and likely, your clients, too.

For these reasons, there's been a lot of pushback from human translators against broad use of MTPE in the industry. (Check out the tag #humantranslator, an initiative created by EMMY award-winning DE-EN translator Sandy Jones @sj_translations!)

So by now, you've probably already assumed that MT is terrible for anything inherently creative. Let's break down WHY that is with some of my favorite concepts from translation theory!

(You thought you were going to get out of it this time? Who do you think I am?)

As always, we're bringing it back to polysystem theory, the concept that every linguistic creation (reports, books, articles, games, musicals, poems, speeches), or "text", is inherently tied to the social and cultural context in which it was created.

I'm a living example of this every day, when I (Californian) outside and it's 60 degrees and I go "It's cold!" to my (Utah snowbird) fiance, who laughs, "This isn't cold. It's nice!"

Same word. Different sociocultural context.

I like to describe it as a fish living in its own cultural fishtank. The salt levels, PH, temperature, all need to be perfect for the fish to thrive and grow.

But what happens when you take that fish out of the tank and plop it into a completely different one?

If texts were fish, we'd modify the water first: bring up the PH, adjust the saline level, heat the water a bit, move the tank away from the window.

Unfortunately, culture is hard to change, so we make these changes to the fish itself for it to thrive in its new environment.

This process--of adapting the text to function equivalently in the target culture--is commonly referred to as localization.

It's not just for translation, too--moving a fish from an British tank to an American tank requires localization, too. ("What the hell is a car 'bonnet'?")

Complicating this is yet another facet of the fishy text: context. The phrase "to the moon" hits different in a historical documentary about 1969 than in current Twitter discourse.

Often, there are multiple onion-like layers to context: the larger culture (Western), the smaller culture (r/WSB), the structure of the text platform (Twitter/Reddit), etc.

Each one of these layers forms a 'system' of contexts and cultures that a translator needs to understand to fully grasp the meaning and purpose of a word, phrase, or document.

Many contextual systems--a polysystem. Polysystem theory!

Herein lies the issue with machine translation: in its current state, we absolutely cannot teach a machine the nuances of culture and context. Hell, not even human translators are expected to know every single culture and context off the bat. (See: zoomer memes).

Human translators, however, can do research, consult with humans familiar with those contexts, read similar texts to learn about those contexts, etc.

Machine translation engines cannot. They know what they know, and they'll try their damndest, but in the end, they guess at what "sick" might mean in the sentence "You were totally sick just there, dude!"

Here's what baffles me (and likely you too, dear reader!): for decades now, Western culture has ridiculed MT. From mocking Engrish translations of food menus to singing songs translated through 50 languages on Google Translate, we are all hyperaware of how crappy MT is.

And yet.

Western culture ALSO has the idea that MT is a catch-all. From 'babelfish' devices that go viral every month to my mom asking "when my job will be made obsolete", we assume that a MT will be "good enough".

Despite memeing on MT all the time.

Sometimes, it's this ignorance that leads companies to utilize machine translation. It'll be good enough! At least they'll be able to read it!

Sometimes, sadly, it's capitalism. Sure, it's not good, but if it'll get a few more people to buy it, who cares?

Game Localization, and good loc at that, is expensive.

Not only are you paying for the cost of translation, you're also paying designers for graphics (see: P5R!), additional QA to ensure the translations display correctly, and additional marketing reps in other languages.

Sadly, cutting costs by using MT, or MTPE, always does more harm than help when you're translating a game. Even the words "Yes" and "No" can have dramatically different translations depending on where they appear in the game (a dialogue selection vs. "Would you like to save?").

(Side note: did you know the terminology for game systems vary within languages, too? What is a "joystick" on Switch may just be a "stick" on Playstation, and messing up this terminology can lead to the first party not allowing you to release your game on their platform!)

This is getting long, so let's cut to the chase. How do we prevent this from happening?

First and foremost, we've got to inform developers and producers in the industry of the value of good localization--and why human translators are the best way to ensure your loc is good.

As we've seen time and time again, good translations help your game and in some cases, can sell it to other native speakers. Bad translations hurt your art and tell your international base you don't care about them.

One last note: professional translators, don't be afraid of (or feel ashamed about) using MT as a tool for your own translations! There's nothing wrong with using Google to parse a difficult sentence or give you some ideas.

Just... don't copy-paste it into your game translation.