**Globally Speaking**

**Podcast 047**

**Ethics in Machine Translation**

M I'm Michael Stevens.

R I'm Renato Beninatto.

M And this week on Globally Speaking, Renato, we're talking about a very timely subject. And that is the ethics around the machines taking over.

R Well, ethics is always a hard topic, because ethics ... It's something that changes over time, right? You have some basic foundation of things that are considered ethical, but as society evolves and as things change, ethical problems from the past cease to exist, and new ethical problems appear.

M Absolutely. And they often surprise people. You look at what companies like Facebook go through when it comes to governments being involved in promoting news, and there's a big ethical issue there, they don't know how to do it. And there's this one ethical issue related to automation. I know you have less of a concern. You know, there's a lot of calls out there for people, that politicians need to get involved today with self-driving cars, who's going to put all of these people out of work.

In the past, when we talked about am I going to lose my job on that podcast, you seemed to have very little concern about losing your job, or the ethics around that. And that's not really what we're talking about today, though it is one aspect of this.

R Absolutely. I think that the key aspect for us is the element of ethics applied to artificial intelligence. Let's take the element of job out of the discussion, because in the conversation that we had in previous podcasts, my opinion is essentially that we don't need to defend jobs that exist today, because there are going to be new jobs that are going to appear as a consequence of the changes that are happening.

So, we're not here defending the job of a horse carriage ...

M Buggy whip manufacturer.

R Exactly. Because this does not exist anymore, and the reality has changed. Then we talked about threats of the self-driving car replacing all the truck drivers and bus drivers, and all of those things. But ...

M We're also in a market where more words are translated every day through machine translations than in the entire corpus of human translation for all of time before.

R One of the funny things, is that back in 2003, when I was with Common Sense Advisory, I spoke at a conference in Prague. It was the UNICO conference. And I, for fun, I wrote something that was called the Automated Translation Manifesto. And I went back to read it. At the time when we wrote it, it was supposed to be something that was to be public and distributed widely. But I think that there was a copyright notice so it is protected, so it defeats the purpose of a manifesto. But that's another story.

But what I discussed at the time was very interesting, going back and reading the material, was that I was proposing ... Because people were afraid that the machines were going to take over. As people are today, and this is 15 years ago. And what I was proposing was an open source, I think I called it the limits of machine translation, that this should be an environment of open development and that everybody who used machine translation, whether in written or spoken form, would know that it was done by a machine or it was done ... or you didn't need to say that it was done by a human, but you would differentiate and alert people that it was going to be done by a machine.

And I wanted it to be, I said that automated translation needed to be available from all languages to all languages; it's to promote the communication of the world, and looking back at that, it's amazing how much progress has happened in the last 15 years.

M It really has. And that conversation had been going on, I think, for a long time before that, sort of the earliest accounts I've read about are as early as the 17th century, where people were talking about what's going to happen when translators are using machines to help them save time. Then by the early 20th century, like 1930s, you had patents being issued for translating machines, and by the 1950s, machine translation was in production in an academic sense, though way too expensive and the quality not that good.

So, we've been in this discussion for a while and now it's happening, and so the ethical question comes to the forefront.

R This is coming, it's not specific to the translation industry, right? If you look at the programs of most of the artificial intelligence and machine learning, advanced computing, quantum computing conferences that are going on around the world, academic and commercial, you will always see a session on ethics.

So, this being a new phenomenon, people feel the need to fill that space with rules and ethics. Well, one of the problems with ethics, though, and this, I'm being a little sarcastic here, you're the reverend, you're the one who has studied this stuff!

M All ethics were settled hundreds of thousands of years ago, so we're all good.

R Exactly. The problem of ethics is that ethics only works for the guys who break it because all the people that follow the ethics, very often feel like suckers, because they're being fair; they're playing by the rules, and there's always a jackass that comes through the window and breaks all the rules and then creates another set of ethics.

This is something that happens and comes with innovation. Right?

M Yeah. I heard a quote, from Phil Knight that is him quoting a Chinese philosopher that says, "You're only remembered by the rules you break."

R There you go.

M Yeah.

R So, some people feel ... you know, that I've been always involved with the translation industry associations and organizations and nonprofits, and every translation association creates a code of ethics, and sometimes I read some of them and it feels like that 1984 book.

Oh, I'm sorry. I'm mixing up. Animal Farm. I'm talking about Animal Farm, where the pigs become the leaders and they create the rules, and the number one rule is all animals are equal and then after a while, the rules becomes 'all animals are equal, but some animals are more equal than others,' because as time evolves, you create a bureaucracy, you create a certain elite in an organization, and those ethical and romantic, in a way, values that were designed in the beginning seem a little childish after 20, 30, 40, 50 years, because they don't follow the development of the real world.

M So, to sort of dig into the question a little deeper, when a translation buyer is working with a company or an individual translator and they are paying for human translation, And that's been the agreement, what is different that they're paying for there than machine translation? Because they're still getting words back.

The output ultimately looks the same in some ways, but what is it they're actually under the impression they're getting?

R It depends. This is how you look at the world, right? So, my partner, my co-author in The General Theory of the Translation Company, Tucker Johnson, just wrote a blog entry recently talking about the ethics of machine translation, and I think that he created a very good separation of concepts that helps us decide how to behave when it comes to this decision, whether we should let the client know that machine translation's being used or not.

And this is how you look at it, right?

M Yeah, I think ... what Tucker provides is a helpful framework, but when someone's under the impression they're getting human translation, what they're really buying the translator's, the term is 'agency.'

R Yes.

M Machine doesn't have that capability. So one example I came across, and it was a quote from Franz Och, who was behind Google Translate, one of the minds there, and he was talking about when Google Translate was working on a project, and it translated Heath Ledger in the English into Spanish as Tom Cruise.

R Yeah. I've done that myself, I had a meeting in a hotel in Madrid, and the name of the hotel was, I don't know, Corazon de Maria or something like that, and the English translation it became Sheraton. I couldn't find the hotel.

M That's right. And the idea was the translator made this change because the target was Argentina, and at that time, no one knew who Heath Ledger was in Argentina, but they knew who Tom Cruise was, and so this sort of separated it out that it was a very good translation there, whereas the machine wouldn't have that agency to be able to make that…

R Adapt and transcreate and so on. What you need to see here is, and the point is that, what is the client buying? Is the client buying a product or is the client buying a service? And these are two different things.It's like ... the example, if you're buying a sausage, you need to know how the sausage was made.

If you're buying it as a product, as an end product, as something to be consumed, you don't need to know how the translation was made. You don't need to know what was the process, how many translators were involved, what TMS was involved. What project management system was used. What the client wants to know is that they get their product on time, with good quality, and that they can use it immediately for the purposes that they want.

So, if you think, and if the client is buying the translation as a product, he doesn't care how you do it, he's just finding you, the LSP or the individual translator, as the vehicle, as the expert, to provide them with a product that they want and that they need.

But if you change that picture and if the buyer buys it as a service, and they negotiate how the process is made, they negotiate conditions, they negotiate what is the profile of the translator, so they might require confidentiality issues, they might require the translators to be certified or to have a certain background or education and so on, the approach is a little different, because then, when you're negotiating the opportunity, the business, the deal, how you provide the service becomes part of the negotiation, and if there is a commitment to deliver a service to the client, and you committed not to use machine translation, you cannot use that ethically. I think that that's the basic discussion in the process.

M If the client has committed to a process and a pricing model that matches that process, and then without their knowing, you change the process without changing the pricing model with them, you have violated the ethics of that contract.

R Absolutely.

M Whereas, if they are saying "Hey, here's a fixed cost. Get as much done in however many languages you can, in this period of time, and deliver it back to me." At that point, you have an open up. I liken it a lot to software development contracts. When companies bid on those, they rarely say "Okay, I'm going to use developers in this part of the world, who are writing it in this kind of code or this." There's some basic agreements of what we need is an iPhone app at the end of the day. That’s the final deliverable.

Now, if it's a more complex thing, you may have clauses that say yes, and in some cases open source software will be used in a piece of developing this larger software. And the client agrees to that ahead of time, so the company developing the software can put that in without having ... Knowing that they're going to get cost savings by not building it from scratch there.

That's rarely done. That has not been the sort of agreement in most client/buyer side to language providers engagements in our industry.

R Yeah. And there is always this element, and I can hear the uproar of you, our listener—"What about confidentiality? What do you say to the problem of confidentiality?”

I think that essentially your job as a provider, as an individual translator or an LSP, is to make sure that the tools that you use guarantee the level of confidentiality that you agreed to with your client.

So…read the contract. Your job as a professional, as an expert in the language business, is to deliver an excellent product to your client, with all the guarantees that you can provide, to the best of your judgment.

So, if I remember correctly, if you use Google Translator as an API, confidentiality is protected, because the content doesn't go to the pool that is used to educate the engine.

But in Microsoft Translator, you didn't have that same kind of caveat and guarantee in the translation conditions user contract, however you want to call it.

M Yeah. There's specific programs that Microsoft has to offer that provide that level of security, that are separate.

R And you can hear our listeners saying "But what about that story about Translate.com, or that website that the Norwegian government fed a bunch of information there, and the next thing you know, you could recover all the information from that site?"

This was a little scandal in our little industry that didn't make the headlines anywhere else. But that problem was specific, because in the terms and conditions of that specific platform, it was said that confidentiality was not guaranteed.

So your job, as a professional, as an expert, as a user of machine translation, is to do your due diligence, and to endeavor your best efforts to guarantee to your customer that the product that they're buying, or that service that they're buying, meets all the requirements that they have.

M Yeah. And I think, as a buyer, it's fair to have an agreement with that company, and if you want to explore ways to get better productivity for lower costs, decide where the areas you are willing to compromise as far as automation and use of machine translation. It is ... Very few people would say that it is acceptable to pay to have content translated and then deliver it back just having it run through Google Translate. And almost every client I work with has a way of making sure that that's not what's been done very efficiently.

R Absolutely. Absolutely.

M It's a bad practice from LSPs; it's a bad practice from translators to do that.

R But let me share with you something funny, Michael. A common thing, and you hear these discussions in translator forums and so on, is that when somebody sees a bad translation, the first reaction that they have is "Oh, this was done by Google Translate."

And then you start reading that bad translation, and you see that there are mistakes that a machine translation wouldn't make. Like spelling mistakes, or translate the same term in three or four different ways. There are some mistakes that, you look at them, and it's "This was done by a bad translator," or "This was done by someone using a dictionary."

But the immediate reaction that everybody has is "This was done by a machine." And that the machine has a very, very specific set of mistakes that it makes. And I'm going to refer again to that great series of podcasts that we did about neuro machine translation.

The problem with neuro machine translation is that the translation is very fluent; it's beautiful, and you need to be extra careful in doing the post-editing and the review, because the translation looks so natural that you don't see that they might have changed the polarity of the sentence, they might have changed the meaning of the phrase because of a stylistic adaptation that it does, because it looks very, very natural.

M Well, transparency in the contract for translations is really key. That's why contracts exist, is to help mitigate the ethics of the business deal that's going down. I think we've had a good look at what it means for a translator and a buyer to have that conversation.

There's a whole other group of people that maybe someday we'll explore, who are thinking about the ethics related to this. It's the people who are developing the MT engines. What kind of biases are in the translations, what kinds of biases are in any type of artificial intelligence and machine learning?

People who write source text are thinking about ethical implications of machine translation. Post editors, the newer role that has come out of this new world. People who are going in to get training as translators, people who are teaching. This impacts our entire industry, and I think this is a good first conversation about ethics related to it.

R Yeah. And what I would say, what I would keep in mind when getting into discussions of this nature, is to step back, and not to have an ultra-orthodox position on anything.

One of my favorite columns ever over the years was the Ethicist, The New York Times, they used to have, on the Sunday magazine, they used to have this column and I used to read it. And I think there is a podcast by this guy, I forgot his name. But it's called the Ethicist. And people would come up with very mundane issues related to ethics. How to behave in relationships, to relatives, to gifts, to bribery, to all kinds of things on a regular basis.

And one of the things that I remember, one of the comments that he made is that it's all about context, it's all about situation, it's all about communication. So, my advice, when you're getting to a point like this is talk about it. And don't take “There’s no way I'm going to do it." Or, "This is impossible; this is against human nature," or whatever position, because you might have somebody else that will take a different position and view different things.

Ethics is something that is very connected with religion. And in some countries, the religion behavior, the religion beliefs, and some of the behaviors are affected differently than from our Northern Atlantic view of the world.

M Individualist. My rights over your rights, rather than…yeah.

R And these are discussions that happen in medicine, these are discussions that happen in economics, these are discussions that happen in all aspects and all walks of our life.

And the point is, just talk about it. Discuss it…the parties that are negotiating are the parties that are have to talk. And we, being as a consultant here, my recommendation usually is that when you're talking to a client, act as an expert. Because you're the expert in your business. You're the expert in the language services.

So as an expert, explain to the client whether you're selling it as a product, so you deliver in the best way that you can do it. Or if you're selling it as a service, and then you can negotiate how the project is done and so on. And, charge accordingly.

M Yep. Yep. It's great. The ethical conversation moves us beyond if, and into the how are we going to do this—how are we going to apply it?

R Exactly. Yes. Machine translation is not in the "if" realm anymore. It's in the "how." And it's not even in the "when," because the when is now.

**END OF CONVERSATION**