

Andy Benzo: Ethics in an AI World

Hello. Welcome to the Artificial Intelligence in Translation Summit session that I will talk about ethics. My name is Andy Benzo. I'm a lawyer and a translator from my native Argentina. I live in San Diego now and belong to the ATA, the American Translators Association. I will talk about ethics in an AI world.

So, as a lawyer, first thing we do is let's define the terms. So what is an AI world? An AI world is a world in which artificial intelligence can assist, augment, in some cases, replace human capabilities. The use of big data and artificial intelligence raises several ethical questions. But before we go into the ethical questions, let's define what is ethical and what is legal. Ethical and legal are two distinct concepts, very different, but they both guide human behavior and help you make decisions in different ways. Ethical principles are a set of moral values that dictate the society, what you live in, and they consider things right or wrong.

Legal principles, on the other hand, are defined in the legal system of a country or a jurisdiction. The ethical principles are grounded in personal or societal beliefs, that tells you what is morally acceptable. The legal ones, we find them in codes, in law, in regulations, or in judicial opinions. Depends on the system.

And the main difference between ethical and legal is we conform to ethical opinions, we comply with legal principles. Why? Because it's a choice. We choose to comply with ethical principles. And the violation of those could result in a societal consequence or a reputational consequence, but it's not a legal thing that we have to comply with. Legal principles are written.

We know what they are. We know we have to obey because the authority is going to make us obey. If crossing the street with your car in a red light, that's illegal. And if you do it in yellow, it's a, well, it's a choice and I'm going to get caught or not. But if you, if they get you in the red light, that's illegal.

So we, they can lead to, the violation can be fines, imprisonment, any other legal consequence. So when I was asked to do something about ethics, I analyzed eight different code of ethics. Just to find what are the most common principles and I came up with four.

Of course, this is not by way any extensive list. This is just what I thought was the principal one. So I studied the Code of Ethics of the American Translators Association, Asociación Argentina de Traductores e Intérpretes, Asociación Española de Traductores, Correctores e Intérpretes, ASETRAD, el Colegio de Traductores Públicos de la Ciudad de Buenos Aires. Colegio de



Traductores Públicos del Uruguay, Institute of Translation and Interpreting, ITI in the United Kingdom, International Federation of Translators, which is in France, and the Israel Translators Associations. Out of all these, and out of all the research I read, I came up with the four main basic principles, the fundamental ones, and by no way this list is exhaustive.

First, confidentiality. Translators must uphold strict confidentiality in any document that we are translating. The client relies on that, and we rely on that, and we give our word. The second is integrity. Translators must represent and work within their qualifications, competencies, and capabilities. What does that mean?

My degree is a literary and technical translation, but I haven't done it in 25 years, so when you are not an expert in that, when I became a lawyer I started doing only legal translation for obvious reasons and I stayed doing that but because that's my expertise and I believe that this time and age with the advance of artificial intelligence we all have to be experts, experts in one, whatever subject matter you like.

We have to be honed in our skills. The third one is accuracy and fidelity. And this must be accurate to the, faithful to the source text content, context, and nuances. And the fourth one is professional competence. And this is, um, when we... We have to strive to maintain a very high level of competency, very high, understand every day more in their study, our language pairs and subject matters.

And that is why we come to conferences, besides seeing our friends and, you know, catching up with their lives, but also we want to keep up with the latest advance in technology, to keep learning what is the latest thing that's going on. Because our profession is very isolated. We, you know, every translator is somewhere in the world and we all contribute to the same.

But, we, although we speak freelier now, more freely with Zoom and we catch up with FaceTime, we are each one in our own homes or wherever you work. And it's very important to know what happens in the industry. Although we are not an industry, we are professionals. We are a profession who provide a service. Definitely, artificial intelligence is not replacing us. It's not perfect. It can help us, but it's not replacing us.

So moving on, how are these principles affected when we use AI in translation? First one, confidentiality, which is the first that we saw before. So when we use artificial intelligence systems, we run the risk of that the data could be accessed or intercepted. So whenever we do that, whenever we use one of these systems, we have to make sure that everything is anonymized or de-identified so we prevent any leak of personal data.



The second one is misinformation and this goes with integrity and with fact-checking and source verification and I have an anecdote which is not mine, it was in newspapers all over the United States, in June last year. No, this year, I'm sorry, this year, there was a case of a lawyer in Manhattan who prepared a brief, and you know the briefs are, you know, 50 pages, but they prepared a brief, and he cited all the cases, and he went to court, and gave it to the federal court judge, and the federal court judge started asking, okay, what is this case and what is this case?

And the man said, well, actually, I got it from ChatGPT.

The judge was a little bit taken aback and said, okay, did you verify any of this? And the lawyer said, "uh, no". ChatGPT has given him six fictitious case citations. And you know how the American system goes, that you cite the cases to prove your point. But no one in the office chose to go and fact-check or sort of verify any of these citations.

So the judge was very, very upset to say the least and he was fined a very huge amount of money. So we call it hallucinations whenever it becomes something that has nothing to do with it. We have to be very aware when we're working with AI systems. The next one is bias. By no way artificial intelligence can be neutral.

Because it's based on text and content generated by humans. And we all carry some bias. It's the law. I mean, it's the law of nature. The next one is culture sensitivity. And in the same token, We need to recognize that AI does not understand cultural and social sensitivities, and if we are not very conscious of what we do, we may provide offensive or insensitive translations.

The last one, or the fourth one that I talked before is consistency, and this has to be with professional competence, the output of AI, I wrote sometimes, but generally never consistent. So we have to make sure as professionals that the terminology, the style, the tone is uniform. Okay, so we provide a very good product to our client.

So the question is, should I use AI? Absolutely, it's a tool. Why wouldn't we use AI? We don't go to a doctor and he treats us like in 1980. So why would we, in our profession, be working like in 1980? But AI, artificial intelligence, has to help us to get better and faster. I use AI. Actually, there are many programs that use AI, not necessarily ChatGPT, but, you know, Grammarly, many others that we use.

So, how do we get that? How do we think of AI in our lives? Well, I call it my added improvement. That's my AI. It doesn't replace me. It will never replace us. Because our creativity, because of our knowledge, because it generates content, but it's within the content it's being fed. But we have no limit. So, it will not replace us.



But we have to use it as a tool. And we have to incorporate this added improvement in our toolbox. There's a lot of research, a lot of studies now that they place the human in the center. Absolutely. We were just talking seconds ago with the other slide. It's that we have to be in the center. We are the ones driving.

And one of the examples I say all the time is, even with all the technology. With all this sophistication in technology, when you get in a plane, like many of us take, you know, when we go to conferences, not this one, but others, you want two pilots in the cockpit. Not the plane driving itself, not the plane flying itself.

You want the pilot. Why? Because the pilot is the one who's going to make the decisions and what we want that in translation. We are the ones who have to make the decisions. That's fine to help and with the amount of data translated, it's fine that we do faster, but we are still making the decisions. So I was investigating all this and I came across the investigation of the paper, 2019 European Commission that published the Ethic Guidelines for Trustworthy AI.

That's not immediately applicable to translators, but I thought that the requirements that they're asking for developers and for everybody else in the industry are very interesting. So there are only seven requirements that they name. The report is very, very big. It takes a lot of time to read it.

You have in the other slide, you have the address, you can get it, you can download it, you can read it. It's very interesting. I think we should be in this. We should know what's going on, we should understand what's going on, and we should advocate for what we want. So the first one in the seven pillars is human oversight, and in this requirement there are four.

One is self learning or autonomous. The second is overseen by human in the loop, then is overseen by human on the loop, and then is overseen by human in command. So, of course, I like the last one, in command, but they had different series. So the first one, human in the loop is when you make any decision in the designing cycle, human on the loop refers to the capability of the human interaction with the AI system, monitoring the system operation, the output, and the human in command is the broader one, the economic, the societal, the legal, and the ethical impact of artificial intelligence. And also the human in command is the one like, should we use artificial intelligence in this project or not? So every one of these seven pillars have hundreds of questions in it.

I just got one, just to, like a sample of what we are seeing. So in this one I got, have the humans been given specific training on how to do the oversight? The second pillar is technical robustness.



So, is it compliant with the safety security standards? Do we have what it needs for this system to be safe?

The third one is a big one is privacy and data protection. So there were, I think it was the bigger part of the report. This is the system aligning with the current standards of data protection. And one example they gave is the ISO on artificial intelligence. Are we following all this? And this is how we want these artificial intelligence systems to operate?

So now we have number four, which is transparency. And the question will be, does the system have measures to continuously assess the quality of the output? This helps ensure all stakeholders, they can clearly understand the workings of an AI system, including how it makes decisions and process data.

The fifth one is diversity and non discrimination. Is there a strategy to avoid creating or enforcing bias in the data we are providing? The sixth one is societal and environmental well-being. Huge, huge about environmental well-being. How many, how much energy do we use? How many trees are we, you know, not helping with this amount of huge energy. Now, all of us want to be cognizant of how we are harming society or environment. The seventh one and the last one is accountability. Is there an AI ethics review board to discuss accountability and ethical practices? And this is not about, liability only, this is about risk, managing risk, and how are we going to do that?

We need to talk, this belongs to us. So ethical considerations and responsible development are crucial to shaping this AI world in a way that aligns with human values and principles. So, who should be involved in creating ethics principles for AI? All of us. This concerns all of us. We need to know.

We may not be in the nitty-gritty of the technology. I assure you, I will be not. But, in everything else, I can help with the legal, with the ethical. Let's all collaborate. This is a collaboration issue. So, developers, data scientists, linguists, professors, academia, research in academia, legal compliance officers, intergovernmental entities, , governmental agencies, all of us have to collaborate.

This is new and this is affecting the whole world. So now I want to leave you food for thought. So when you leave this session, you go and talk to your peers, your colleagues, your teachers, your professors, your whoever, and discuss these next questions I'm gonna give you, there are four questions about the future of AI.



And we all need to participate in it. So these are the questions. So the first one is how can we make AI systems robust? So the commands are executed without malfunctioning, without hallucinations, or being hacked. The next one is how can we increase our prosperity through automation while preserving the resources and goals of the people?

We don't want to displace a lot of workers because we use artificial intelligence. So how can we, that help us and create new jobs, create new sources for people to work. So we all have to think about that. Next one, how can we modernize our legal systems? That's, that's my heart, to make them more efficient and just keep pace with AI and to manage risks associated with artificial intelligence. We have to think about it. Who's going to be liable? What are the risks? What are, we should talk about it all the time. Not 24/7. No. I know you're not going to leave this session and you're going to go talk to your friends and partners 24/7 about ethics.

No. But we have to start. This conversation needs to be happening in society. And the last one is, to what set of values should artificial intelligence be aligned and what legal and ethical standing should it have? So before I finish, which is, I promise I'm finishing with this, is my favorite quote of Dr. Seuss, "Unless someone like you cares a whole awful lot, nothing is going to get better".

It's not. Thank you very much for being with me in my ethics in an IA world. And, like I said, my name is Andy Benzo. I'm very happy and very proud to be part of this Artificial Intelligence in Translation Summit. Thank you.

