

Accept no imitations! How to get the most out of AI language tools

It's like a dream - artificial intelligence can produce a text or translation within seconds. But such a text is produced by statistics, not semantics. And as in dreams, there is no guarantee that it is true or even makes sense.

Amanda Crain has been working as a translator and editor at the University of Tübingen's Public Relations Department for 15 years. She reports on her practical experience with AI language tools.

The computer scientist and cryptanalyst Alan Turing tackled the question "Can machines think?" in 1950. Instead of redefining "thinking" – a concept that didn't apply to machines the way people imagined – he reframed the problem as what he called "the imitation game," asking: Can a machine perform a task normally associated with humans such that a human cannot tell the difference? Turing predicted that there would one day be machines that would be able to pass what became known as the Turing test.

Turing's expectations have been fulfilled in many ways: today we have systems that use complex algorithms to analyze vast amounts of data, to recognize patterns within it, and to reproduce and vary those patterns. In some situations, AI chatbots for example can pass for human, thus passing the Turing test. Yet the popular expectation that AI could learn to match humans in all things is far from being met. As Turing recognized, the technology does not "think" or reason or ponder or judge the way humans do. It operates on mathematical probabilities. Yet these are of only limited use when applied to the improbability and seeming illogic of human language. Therefore, any AI-generated text can only be a starting-point for a piece of writing intended to communicate a truthful and reliable message.

When AI programs are given relevant data and applied to suitable tasks, they are both fast and immensely useful. Yet the subtleties of human language are hard to quantify. If we are to use AI for text production, we must understand not only what AI systems can do, but also what they cannot do.

When it comes to academic writing, AI is alluring. Instead of staring at a blank screen and having to come up with a structure and ideas, users are now able to start the writing process by using prompts to generate a text on the desired subject. It saves a lot of typing. But users need to be aware that AI language processing works by trawling through internet texts to calculate the most likely next word or phrase for a given prompt; as yet, its horizon is no broader than that. It plagiarizes, remixes, and regurgitates. It cannot produce anything original or insightful, except by accident. And what it produces may or may not be true. An AI text program can make any thing up if, by its mathematical parameters, that thing is probable.

AI solves problems by making countless repetitive rules-driven decisions – something humans are not good at. For instance, AI can convert your Harvard-style bibliography into Oxford style in seconds – although you would be wise to check it.

But AI has major limitations in processing human language, precisely because much of what we say is not probable or predictable. AI language processing systems face the further drawback that the training data available to them is distorted and not representative.

The example of AI translation

Programs like ChatGPT draw on such enormous data pools that it can be impossible to establish from where and on what criteria they assemble text. But in more controlled AI language processing, such

as translation, the picture becomes a little clearer. AI translation is in effect intensely prompted text generation; this turns every sentence into a case study.

The benefit of hindsight

I have been using translation systems almost daily for more than a decade. The data sets used in translation memory systems of a dozen years ago were tiny compared to the later machine translation and today's "large language" and "large reasoning" models. But their manageable size did allow the user to gain insights into how automated systems go about the mathematical construction of human language.

Even in the current data deluge, it is possible to identify language situations in which AI cannot yield anything but a translation that is misleading or simply nonsense. That is the very latest point at which a human needs to step in, with human understanding and a big red pen.

Million-to-one chances crop up nine times out of ten.

The patterns of human expression are so subtle and varied that any single, mildly creative or original sentence is, in itself, an extremely improbable collection of words. Think of the predictive text on your phone: the most likely answer is not always right, and the right answer is often not very likely.

Depending on the quality and the kind of text it is given, an AI-based programs like DeepL and Google Translate may be able to turn out a passable translation. They choose the statistically most likely words, lining them up like beads on a string. This reductive approach enables them to deal capably with texts constructed in a typical pattern: "Johannes Kepler was born in 1571 in Weil der Stadt"; "Insert tab A into slot B"; "Once upon a time there was a king who had three daughters."

Yet even the stable patterns of biographies, operating instructions and fairytales are more varied and structurally complex than we realize. Once we depart from very basic patterns – which we do all the time – translation, like all writing, depends on meaning, not math.

Without the human capacity for context, translations can go badly astray. Here are a few of many examples. Because of the ever-changing AI data pools, they are no longer reproduceable. But this type of error will always occur because human language references outside contexts which AI cannot include in its calculations.

- In the first, a keen new professor picks up a screwdriver and works on some equipment:

Professor S. hat sich gleich am ersten Tag im Labor die Handschuhe angezogen und mit mir am Strömungskanal herumgeschraubt.

DeepL rendered this as:

Professor S. put on her gloves on the very first day in the laboratory and screwed around with me on the flow channel.

The program reproduced a predictable pattern but could not recognize that to "screw around" has quite another meaning from its innocent literal translation in German.

- Sometimes the right context requires extra definition. As here, where DeepL gets lost in the woods:

Jeder Forst ist zugleich auch ein Wald, aber nicht jeder Wald, und wäre er auch noch so groß, ein Forst.

DeepL: *Every forest is at the same time also a forest, but not every forest, no matter how large it is, is a forest.*

- Even *sein* is not always simply *to be*, as a look at Article 1 of the German constitution shows:

Die Würde des Menschen ist unantastbar.

DeepL and Google Translate yield variations on:

Human dignity is inviolable.

In English, this is a declarative and clearly untrue statement, because human dignity is violated every day. Article 1 is a declaration of principle:

Human dignity shall be inviolable.

The calculation game

If a word in the source language has four possible translations – one occurring 40 percent of the time and three each occurring with twenty percent frequency – AI translation will choose the statistically most frequent option. Put simply, it will choose wrongly 60 percent of the time when that word occurs.

Star or starling? A bird becomes a celestial body:

So wie ein großer Schwarm von Vögeln ganz andere Bewegungen am Himmel macht, im Vergleich zu der Bahn eines einzelnen Stars.

DeepL: *Just as a large flock of birds makes completely different movements in the sky compared to the path of a single star.*

Bacteria take the waters:

Bakterien in heißen Quellen der Tiefsee

DeepL: *bacteria in hot springs in the deep ocean*

Edited: *bacteria in hydrothermal vents on the ocean floor*

And there is a special opportunity for university members to clean up when a lecture-free *universitärer Verfügungstag* becomes *university disposal day*.

When the feature is also the bug

AI can make “informed guesses.” This may solve one problem while creating an utterly different one which no human would expect.

University of Tübingen researchers reconstructed an ancestral natural antibiotic. With the German language’s beautiful capacity for compound words, this became *Urantibiotikum*. DeepL did not recognize this word, so corrected it to two words it did recognize – *Uran* and *Antibiotikum* – and translated them instead: *uranium antibiotic*.

A game of telephone

When translating between several languages, DeepL and Google Translate use English as a “pivot language” through which all other translations run. That means, going from German to French, your German text will first be processed into English, then that result will be processed into French. This can introduce or amplify errors.

The recycled data economy

AI-generated texts are posted on the internet and then are gobbled up as new training data. This self-affirming cycle leads to the boosting of trends as AI suggestions are taken up and amplified. (*De/ve* is currently a popular word.)

In translation, common mistakes become institutionalized. (*Tage* can mean *days*, but is frequently used in the sense of *Tagung*, a meeting or conference. A translation into *days* leaves non-German speakers wondering what students do for the rest of the year if there are only five *Study Days*.)

First the quick fix, then the hard yards

Just as we walk past the bike to the car when we are tired, late, or it looks like rain – we are going to use AI for text generation simply because it is quicker. In making this Faustian bargain, we switch from the creative process of formulating and expressing our own ideas, to the critical process of working out what is wrong and fixing it. We must examine every sentence carefully, because there is no guarantee that what an AI-generated text says is true, or even what we think it says.

Alan Turing didn't speak in terms of "intelligence;" instead, he envisaged machines with the calculating power to mimic human behavior so well that humans could be fooled. In 2026, we take the Turing test every time we log on to the internet. The question to be asking now is not "How smart are the machines?" but "How smart are we?"

Amanda Crain, February 2026

Tips for using AI translation programs

When using AI translation programs, keep the following in mind:

- Consider data privacy: Whatever you upload to a free translation website will be in the public domain, instantly.
- Beware of rubber-stamping a translation of a source text in your native language. **You** may understand every word of the translation, but that may be a false positive. After all – you know what it is **meant** to say. An international audience may be left utterly confused. Translating back into the source language can help identify and fix errors.
- Pre-edit the source text: Ideally, avoid sentences with more than one relative clause; make clear Subject-Verb-Object sentences; scan for words that could easily mean something else and replace them or rephrase to avoid ambiguities.
- Input full sentences or short paragraphs so that you can check your translation piece by piece.
- Remember that every text is written for a particular audience, in a particular social and cultural context, and that the act of translation will place it before a different audience, possibly unfamiliar with that context. The translator must always ask: will the target audience get this?
- Be aware that at any moment, AI translation can introduce a mistake you've never dreamed of. There are no guarantees.