WORDSONLINE

DOMO ARIGATO

AN ODE TO MODERN TRANSLATION

TECHNOLOGY

Did you ever wonder how the universal translator in *Star Trek* would actually work? How could a machine take every language and translate it in perfect idiomatic American English?

According to the *Star Trek: The Next Generation Technical Manual*, the universal translator in the series is an"extremely sophisticated computer program" which works by "analyzing the patterns" of a previously unknown foreign language, beginning with a conversational speech sample. As it becomes more extensive, the "translation matrix" becomes more accurate and reliable, which enables instantaneous translation.

This sort of translation is no longer entirely science fiction. You have probably used machine translation on your social media feeds or to understand a foreign website and you may have even used Microsoft or Google's voice translation. You may have even noticed how much better machine translation has gotten in recent years. Have you ever wondered what's been happening and what's behind the development? How did machine translation develop and how has it gone from comically clunky to pretty good in the space of a couple years?

It turns out that good machine translation has been a long time coming and that the ideas behind it developed long before the technology. The first strand dates back to at least the 9th century and Al-Kindi, an Arabic cryptographer. He developed systemic language translation techniques, including cryptanalysis, frequency analysis, and probability and statistics, which are used in modern machine translation. Al-Kindi was a polymath, knowledgeable in fields including philosophy, astronomy, chemistry, meteorology, mathematics, optics, medicine and music. It is thought that in his careful study of the Qur'an he became aware of the characteristic letter frequency of Arabic and that this could be applied to cryptography and translation.

Many years later, in 1629, another polymath, René Descartes, proposed a sort of machine translation. In theorizing about a universal translator, he proposed an ideal universal language, with equivalent ideas in different languages sharing one symbol, and simple and regular grammar. Each language could then be translated into the universal language and then to any other. But even in the letter in which he proposed it, he realized that

> ***"A universal language that would work at the level of the imagination, describing the actual 'things' of the external world, could only produce uniform results in the perfection of Eden or the ideal of fiction."***

In the 20th century, these ideas started to be put into practice in various forms. By the 1930s, there was a patent application for a translation machine, though it was not developed. Also in that decade, a Russian proposed a translation system similar to Descartes's idea. His system, based on Esperanto, had three steps: first, a native speaker of the source language organized the text into a logical form, then a machine translated that logical form into the target language, and finally, a native speaker of the target language normalized the output. But his proposal remained unknown until the late 1950s, by which time development was going in a different direction and wouldn't come back that way for several decades.

That 1950s saw the real start of machine translation as we know it, using the rapidly developing computers. It kicked off in 1949 when Warren Weaver, a researcher at the Rockefeller Foundation, presented a proposal for computer-based machine translation, based on information theory, successful code breaking in WWII and theories about universal principles of language. By 1954, a group of scientists in Georgetown successfully did a fully automatic translation of more than sixty Russian sentences into English. Based on their success, they thought the problem would be solved within three to five years.

Similar research was being done in the Soviet Union. The success and competition led to significant funding of machine translation research in the US, but progress was much slower than expected and in 1966, a committee found that expectations had not been met and funding in the US was dramatically reduced. Star Trek was more optimistic: the original series (and its universal translator) premiered in September 1966. Of course they could assume hundreds of years of technological leaps and maybe some help from advanced alien cultures.

Luckily, we probably won't have to wait either hundreds of years or for a friendly visit from another star because research did continue despite the cessation of US government funding. Companies like SysTran (founded in 1968) and Logos (founded in 1970) continued development in the US and research continued in places like Japan, France, Germany and Canada. Machine translation also started to be used in real life. In 1970, the US Air Force installed the Systran system and in the Commission of the European Communities followed in 1976. The Université de Montréal developed the METEO system which Canada installed in 1977 to translate weather forecasts from English to French. By 2001, it was translating close to 80,000 words per day or 30 million words per year.

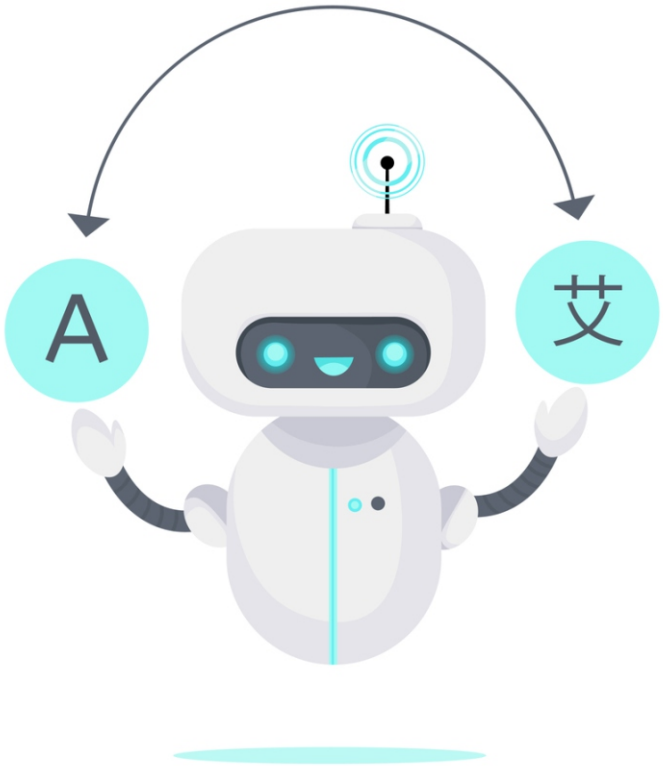
In the 1970s, increased globalization and the entry of players like Japan into the market plus the use of multiple languages by the EU and in countries like Canada (which had become officially bilingual in 1969) meant that there was demand, especially for inexpensive systems that could handle a range of technical and commercial documents.

Up until the 1980s, the focus had been on Rule-Based Machine Translation (RBMT), which meant that computers were trying to duplicate what they thought translators do. They used bilingual dictionaries and rulebooks for each language and fed the rules into a computer. There were three types of Rule-Based Translation. The first was Direct Machine Translation where you take each word, translate it and then form it into a correct sentence. It was pretty clunky.

Then there was Transfer-Based Machine Translation. This broke up sentences into bits and translated each bit before putting them back together. It was better with word order and such, but still did not work very well. The third form of RBMT was Interlingual Machine Translation. This tried to do what Descartes had proposed, which was to translate first to a symbolic language (Interlingua) and then from there into the target language. But, as Descartes had foreseen, it was simple in concept and very difficult to actually create.

RBMT, in short, did not work very well in practice, except perhaps in specific cases like the Canadian weather forecast translation. The method is predictable and results are replicable, but languages are just too irregular, with homonyms and usage patterns defying any machine translation based on rules.

In the 1980s, the Japanese started experimenting with a new technique called Example-Based Machine Translation. They were especially motivated because translating between languages like Japanese and English is especially difficult since the concepts and structures behind them are very different. Sentences need to be completely rearranged and a lot of information needs to be filled in from the context during translation. At the time, few Japanese spoke English but Japanese tech companies wanted to join the globalization game and promote their computing power, which meant reliable machine translation.



They were ready when in 1984, Makoto Nagao from Kyoto University proposed translating ready-made phrases instead of repeatedly translating the same words. Thus, if you knew that "Domo arigato Mister Roboto" from the 1983 Styx hit translated to "Thank you very much, Mr. Robot" you could use that information to translate the first part of the sentence "Domo arigato gijutsu" and then use a dictionary to look up the unknown words, which could be slotted into the same sentence structure, giving you a final translation of "Thank you technology". Scientists figured out that you could feed the machine existing translations instead of spending years specifying rules and exceptions. It was an important step.

The 1990s gave us another important step: Statistical Translation. Computing power had been increasing rapidly, which meant that computers could now be put to work analyzing vast quantities of translated texts to find patterns, without knowing any linguistic rules. Thanks to European Parliament and the United Nations Security Council and the increase in digitalized documents, these texts were also available online in all member country languages, giving the computers the millions of words that they needed.

Statistical Translation started by looking at the frequency with which words were translated into various equivalents and developing predictive algorithms. It quickly progressed and by 2006 everyone had moved from a word-based approach to a phrase-based approach, using many examples of phrases to predictively translate. Everyone, in 2006, included Google, Yandex, Omniscien, Systran and Bing. These were the first machine translators that the general public used online, and often provided laughable translations.

There things stayed for about ten years, with the statistical systems getting incrementally better. But in 2016,Google announced the revolution. They had moved their translation service to Neural Machine Translation (NMT). Other translation services quickly followed. NMT combined Statistical Translation with neural network programming, which permitted a great deal more complexity. Each word could be analyzed in relation to other words in the text, with vectors showing all the relationships. These relationships allowed the models to fit highly complex data, like languages.

With NMT's development, specifically an encoder-decoder architecture, the Cartesian Interlingual Machine Translation idea also makes a reappearance. NMT first encodes the source text and then decodes it into the target language, meaning that the same encoded text could be easily decoded into another language as well. Later developments allowed gave the model attention mechanisms, which lets the model learn where in the input to pay attention to while decoding the output.

NMT also uses Deep Learning to learn as it goes. Deep Learning is training the neural networks with more and more data so that the model becomes ever more refined. By feeding the systems a huge amount of already translated text, the system can train itself to make continual improvements in its translations. This is also why many translation systems ask users to rate translations – it helps the system learn. Because Deep Learning takes place without further human input, it also gets called Artificial Intelligence, which is why a lot of people are talking about AI translation.

At Jonckers, our WordsOnline system uses NMT in combination with your translation memory to provide a content management system powered by an AI. The AI learns which translation fragments are likely to need human review and shows just those to our linguists. This provides a continuous publishing and localization approach that is fully automated and data-driven with a fully integrated AI empowered language community.

The WordsOnline cloud-based platform uses further programming to give real-time insight, so that customers can review and filter business statistics such as; how many words are completed per language, per division, per service etc. This is handy to ensure deadlines are met, especially with multi-million word translation projects. Jonckers provides the Client Dashboard as an integrated module in WordsOnline.

So, while we may not yet have the universal translators from Star Trek, we can now definitely thank the robots, make that the AIs, in whatever language we please, for making machine translation so much easier and better. How about Portuguese? “Muitoobrigado, tecnologia!”

PLACE YOUR FIRST ORDER ON TRANSLATE NOW, TO UPLOAD YOUR DOCUMENT AND PAY ONLINE.
DISCOVER OUR INDUSTRY-FIRST TRANSLATION SUBSCRIPTION SERVICE TO REDUCE ADMIN AND COSTS.

Click here to get in touch if you have any questions
via our chatbot on wordsonline.com