In this video, we are going to see the motivation, history, and applications of machine translation.

As I see it, we do not really need to motivate anything, but still, we are going to say some words about it. Machine translation is interesting from the commercial and academic point of view.

For example, if we take the commercial point of view, we have the Internet, where we have a lot of languages interacting. We have users native in different languages, and we have content in different languages. On the Internet, it's not possible to perform a human translation, so we need a machine translation, automatic machine translation.

From the academic point of view, machine translation is a complex problem that involves many natural language processing problems, like parsing, word sense disambiguation, named entity recognition, transliteration.

MT involves also many machine learning problems, like working with large databases, noise, optimization, search, and so on.

And involving so many problems, machine translation is not going to be solved in the next years.

Just to see an example of commercial success, we can name the Language Weaver company that was founded by Kevin Knight and Daniel Marcu in 2002, and then it was sold in 2010 to the larger company, STL, for $42.5 million.

Here we see an example of machine translation, just to see or read about the quality that we can expect. Nowadays, we cannot use machine translation for medical care, as it requires 100% quality, but we can use it in touristic domains and in translation of news, news translation, and other examples.

OK, machine translation was first born, we could say, in 1954 with the first MT system from English to Russian. This was an experiment proposed by the IBM guys and Georgetown University, and it could translate around 300 words in English, and it contained like six rules. So it really was a very small experiment, but in the historical context, the cold war, it was a very important achievement.

Here we are going to see a little bit the progression of machine translation in history. So from 1947 to 1954, there were the information theory foundations proposed by Weaver and others. These foundations were very important to propose later the statistical machine translation systems that we have.

From 1954-1966, this was the so-called "Decade of Optimism" for machine translation. And people were very happy with the results that they were obtaining using large dictionaries and rules.

Then, from 1966 to the 1980s, there was a little bit of pessimism, because the ALPAC Report was released that stated that machine translation was a very complex problem that could never be achieved. After that, the research stopped in the United States and continued in Europe and Canada.

So it was during the 80s that there was a variety of systems, including Rule-Based and Interlingua, we are going to talk more in detail about them later, in the next video. And in the 90s, there appeared the

Statistical Machine Translation systems that nowadays are very popular, and there are a lot of people working on them. And finally, in the 2000s, there appeared open-source MT software that allowed for the field to progress very quickly.

So now, what is going on in MT? MT is consistently improving with resources and computational power. It offers a reasonable quality when resources are available. And really, we are moving to hybrid architectures with the statistical and linguistic knowledge to cover what data does not cover.

Just to name some MT-related challenges, we can see the speech-to-speech translation. That is an application that translates from a source speech into a target speech. The process is the following: we have the source speech that is transformed into a source text by automatic speech recognition system, then the source text is translated into the target text by using a machine translation system, and finally, the target text is transformed into the target speech by a TTS, a text-to-speech synthesizer.

OK, the applications that we can name particularly are the simultaneous translation, the film translation, the speech message translation. All of them use speech-to-speech translation. And as you can imagine, the quality, well, it's not perfect, obviously. We are concatenating the errors from the automatic speech recognizer, the machine translation system, and the synthesizer.

Other applications are the ones that go from source speech into a target text that are like TV and video subtitles.

Other applications related to MT are the ones like cross-language information retrieval that allows to search for hotels in Paris in French-language pages, or bars in Moscow from Russian sources. Also, the computer-aided human translation, that helps human translators to do their work more efficiently. And then we have the communication via email, chat, and the translation on handheld devices.

In the next video, we are going to see more in detail the MT approaches that are used nowadays.