Artificial Intelligence in Technical Translation

Technical paper from the series
AI: Science over Fiction

bitkom

www.bitkom.org

Publisher

Bitkom e.V.
Federal Association for Information Technology,
Telecommunications and New Media e.V.
Albrechtstraße 10 | 10117 Berlin
P +49 30 27576-0
bitkom@bitkom.org
www.bitkom.org

Responsible Bitkom committee

AK Artificial Intelligence

Project management

Dr. Nabil Alsabah | Bitkom e.V.

Author

Christopher Kränzler | lengoo GmbH

Layout

Katrin Krause | Bitkom e.V.

Cover image

© Blake Silva | unsplash.com

1 Machine Translation is no substitute for a professional translator

When it comes to the current conversation on social change, the area of Machine Translation (MT) is becoming the epitome of a dystopian future, with highly qualified linguists and professional translators being replaced by Artificial Intelligence (AI). In the area of machine learning, MT is one of the few fields that is already sufficiently developed that the use of AI is already a reality. Nevertheless, in the professional language market – the domain of professional translators – MT is only used on a small scale to translate content. Rather, the technology, used with general or generic language models, currently serves to give people a basic understanding of texts written in a foreign language that they do not know. It covers situations in which one may have consulted a friend who is a native speaker, but certainly not a professional translator.

The past decade has seen rapid developments in MT. In particular, the emergence of Neural Machine Translation (NMT) is something of a quantum leap in the translation of language by machines. Nevertheless, the technology is far from able to replace professional human translators.

»Machine Translation (MT) is one of the few fields in the area of machine learning that is already well developed. Nevertheless, in the professional language market – the domain of professional translators – MT is only used on a small scale.«

The technology underlying this approach uses neural networks which can detect relationships between the source and target languages. The software uses data from previous translations to learn sentence structures and linguistic correlations. The approach aims to make the translated language fluent and easy to understand.

The quality of translation using neural networks, however, depends entirely on the quality of the data with which they are trained. To achieve high-quality translations, the neural networks need to be fed client- and subject-specific data, and even then, the way in which the technology is applied in the process is critical in achieving high language quality. In terms of professional linguistics, humans continue to play the most important role.

2 Generic MT is not suitable for professional use

Popular and, for the most part, free online tools such as Google Translate and DeepL use so-called generic MT models that are trained with the broadest possible data set. These tools can be a great help when it comes to translating texts in everyday language to obtain a general understanding (so-called »gisting«). For use in a professional environment, where the quality and consistency of the language used are the most important requirements for the text, a different picture emerges. For example, a study by the Swiss Confederation concluded that generic MT is not suitable for translating documents for publication and »clearly reaches its limits when translating technical texts.«[1] The study also concluded that although translations using DeepL produced fluid texts in the target language, analyses by professional translators revealed content errors that a layperson would be unlikely to notice.

Generic MT is therefore not suitable for standardized use in a professional context, particularly in the translation of technical texts, and without being edited by a professional translator.

1 Swiss Confederation (November 19, 2019). Report on DeepL test. Retrieved April 5, 2020, from ↗**https://uepo.de/wp-content/uploads/2019/12/schweiz_bericht_deepl-test_2019.pdf**

Instead, as already mentioned, the technology serves as a basis for a rough understanding of the content of texts.

The reason for this is obvious, as MT is like all other areas of machine learning; the quality of decisions, categorization, or in this case the translation itself, increases with the quality of the data used to train the model. The narrower the focus and the area of application, the better the language produced.

For professional translations, generic MT systems are not a great help. The effort involved in post-editing (PE) the translation does not lead to efficiency gains when compared to the use of rule-based productivity tools. These types of tools, such as Computer Aided Translation (CAT) tools, which use translation memory technology, enable professional translators to transfer previously translated terms or entire sentences to the target text more quickly, and simplify the application of specific terminology.

»To achieve high-quality translations, the neural networks need to be fed client- and subject-specific data. In terms of professional linguistics, humans continue to play the most important role.«

3 Customized models enable vast increases in productivity

For MT to be useful in a professional context, the models need to be adapted to specific needs and situations to increase the quality of the translation as much as possible. These customized MT models do not differ fundamentally in terms of the technology used, but rather in the selection and processing of the data used for training.

By restricting the selection of training data to a specific area, for example, using only language data from the field of mechanical engineering, the models are better able to learn technical terms and subsequently use them in the actual translation. The increase in quality of language thus achieved is already obvious, but this approach still fails to produce translations that can be published without requiring further proofreading.

Nevertheless, customized models enable immense increases in productivity when they are applied as part of a workflow that includes subsequent proofreading by a professional translator. In such cases, the productivity gains can be measured by comparing the number of words a professional translator can translate to publishable quality per hour.

As part of a study by Lengoo – a language technology company focusing on technical translations with a client in the telecommunications industry – the number of words translated per hour was examined in an experiment. The experiment design compared the four different approaches detailed below. Approach D is the workflow applied at Lengoo, while approaches A, B, and C serve for comparison and correspond to other typical translation workflows:

A) Manual translation by a professional translator without using productivity tools

B) Manual translation by a professional translator using rule-based productivity tools (the CAT tool »MateCat« and a translation memory from the client)

C) Translation by a generic MT model (e.g. GoogleTranslate or DeepL), with subsequent post-editing by a professional linguist

D) Translation by a customized MT model trained on the client's translation memory, with subsequent post-editing of the translation by a professional linguist

The same 20,000 words of customer service documentation were translated from German into English following each of the four processes. The professional translators involved have comparable professional experience and linguistic ability in both the source and target languages. The texts were divided into individual packages of 1,000 words each. The time taken by each professional translator to complete the translation was recorded for each individual package.

On average, a professional translator in experimental approach A produced 250 words of publishable text in 60 minutes. In experimental approach B, professional translators produced an average of 450 words in 60 minutes. Speeds of 400 words per hour were achieved for experimental approach C, which involved post-editing of the content after it had been translated using a generic model. For experimental approach D, using MT with a customized model for the translation, which was then post-edited by a professional translator, the output increased to 1,300 words per hour.

The results of the study are clear: the use of MT with customized models can significantly increase productivity for professional translators. When compared to both translation using rule-based productivity tools and generic MT, increases of approximately 200% can be achieved.

»By restricting the selection of training data to a specific area, for example, using only language data from the field of mechanical engineering, the models are better able to learn technical terms and subsequently use them in the actual translation.«

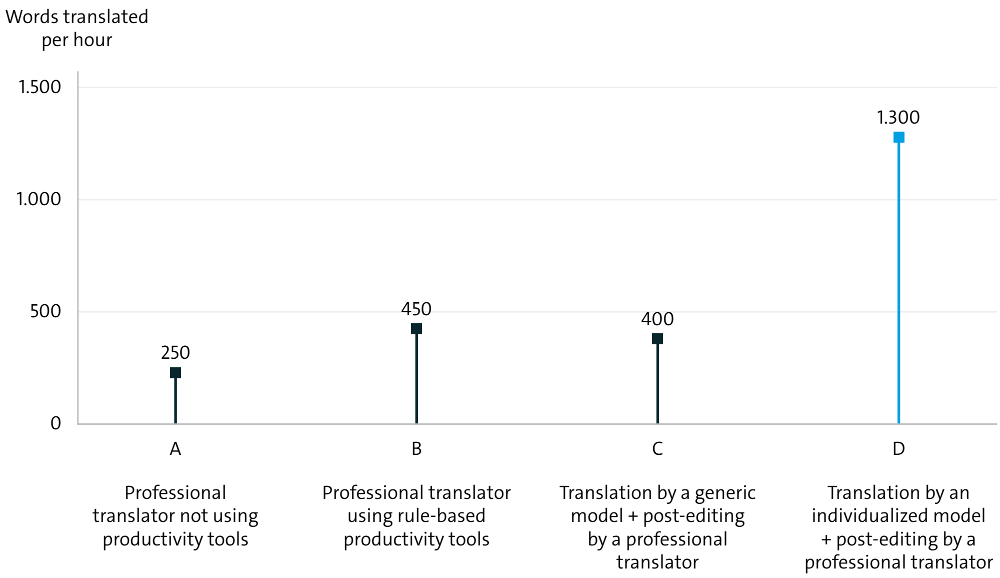

Abbildung 1: Comparison of translation speeds

The potential of customized NMT models is therefore not for the technology to replace human translators, but rather to support them in their work and considerably increase their productivity.

4 Language models have to learn technical terminology and apply it consistently

The translation of highly complex content such as construction manuals, technical documentation for maintenance instructions, or pharmacological content such as product declarations or process descriptions requires a high degree of consistency and the precise use of specific terminology. These demands on the quality of the translation pose significant challenges for NMT technology as, unlike in rule-based systems, the integration of termbases and glossaries is rather complicated.

Models used for NMT are in no way designed to give priority to specific terminology in the translation or to translate previously translated content in exactly the same way as it was before. The current use of neural networks in MT in the context of so-called attention-based MT is based on the vectorization of individual words. This vectorization assigns each word to be translated a certain value, which is affected by the words immediately surrounding it. This means that the vectorization permits contextualization in the translation, which in turn, allows NMT models to produce language which flows as naturally as possible, i.e. a tone of voice with variety. It is, however, precisely this variation that stands in the way of the linguistic consistency which is so important when languages are used professionally.

»Language models are not designed to give priority to specific terminology in translation or to translate previously translated content in exactly the same way as it was before. Using ›terminology constraints‹ avoids this problem.«

Throughout the first decade of the 21st century, statistical MT was criticized for its linguistic monotony, unnatural flow of language, and frequent lack of logic. Today, it is precisely the advantages of this supposedly old technology that NMT lacks, which would allow its use in a professional linguistic context.

That said, the compatibility of glossaries with NMT is the subject of current research. The most promising method is to expand the training data set to include specific terminology and to mark the terms contained in a glossary. These marked terms will have a greater weight in the training and the model thus learns to give these terms preference in the translation.

This use of such »terminology constraints«[2] thus leads to a more accurate translation by NMT models without resorting to rule-based tools. Using this technique can result in significantly higher translation quality.

This technique is used by Lengoo for a client in the pharmaceutical industry and achieves remarkable results, as the example shows.

2 Dinu et al. (2019). Training Neural Machine Translation to Apply Terminology Constraints. Retrieved April 10, 2020, from ↗https://www.aclweb.org/anthology/P19-1294.pdf

Source text:
Repeatability and intermediate precision were determined using human samples and controls in accordance with the CLSI EP05 requirements (2 aliquots per run, 2 runs per day, 21 days).

Translation without terminology constraints:
Repetition precision and median precision were determined using human samples and controls in accordance with the requirements outlined in CLSI guideline EP05 (2 aliquots per run, 2 runs per day, 21 days).

Translation with terminology constraints:
Repeatability and intermediate precision were determined using human samples and controls in accordance with the CLSI EP05 requirements (2 aliquots per run, 2 runs per day, 21 days).

Abbildung 2: Comparison of translation by an NMT model before and after implementing terminology constraints

Using this technique leads to a significant increase in quality of language with a more appropriate use of technical terms and specific wording. Although the technique does require significant manipulation of the training data in order to change the weighting of the terms and sentence structure, the increase in translation quality justifies the increased effort.

5 Training with synthetic data can increase translation quality

The use of AI in translation, as in all other fields where AI is employed, is limited by two main factors: available computing power and the availability of high-quality data.

The availability of language data is not often guaranteed, especially in the field of natural language. In so-called low-resource language pairs, i.e. language pairs for which insufficient parallel data is available for training neural networks, data scientists apply techniques that can reduce the dependence on training data in both languages (target language and source language).

One of these methods uses so-called back translation to artificially synthesize training data. This synthesis involves the fully automated backward translation of high-quality text in the target language into the desired source language (i.e. in the opposite direction to a typical source-to-target translation). This uses translation models that are optimized on the basis of the decoder, the part of a translation model that is responsible for the production of language, that is to say, the target language.[3]

Based on the premise that the quality of the language used in monolingual data, i.e. texts already written in the target language, is higher and available in larger quantities, it is precisely

3 Language models use an encoder, which is responsible for understanding the source language, and a decoder, which enables the production of text in the target language.

those texts that are back-translated into a chosen source language using an already existing language model. The resulting parallel language data, in turn, form the basis for the subject-specific bilingual training of the model.

Language models trained using this method are able to produce significantly better translations. Researchers from Facebook and Google have found that models trained on synthetic data can achieve almost the same quality of language (83%) as those trained using truly parallel data.[4]

A fundamentally different approach to handling »low-resource« language pairs is to train and use multilingual language models. These models are not just able to translate from one source language into one target language, but can even combine several source and target languages in a single model.

Multilingual models enable the translation of n source languages into n target languages, with all combinations being possible. During the training phase, greater emphasis is placed on improving the encoders and decoders than is the case for bilingual language models. As a result, it is possible to train the model based on high-quality data in either a target or source language to improve the language quality in several language pairs simultaneously.

»AI research has developed methods to address the scant availability of high-quality data.«

This method leads to »significantly better translations«[5], based on the transfer of learning within the model. As a specific example, Lengoo is observing that the language quality for all possible language pairs within a multilingual language model increases when the training has been improved for just one target language. Research into the human learning process and the human brain suggests that this transfer of learning should be particularly noticeable for languages belonging to the same family, such as Spanish and Italian. Even so, researchers have found that such improvements within a model can also be seen for language pairs that are not part of the same language family. However, no satisfactory answer has yet been provided as to why this is the case. Despite dramatic advances in MT research in recent decades, many important questions still remain unanswered.

Using both techniques described above enables language models to be trained even in language pairs for which no parallel data is available. This makes NMT technology available for a wider range of applications. That said, MT cannot yet be used for language pairs where there is not even sufficient parallel data for a generic initial training of language models.

4 Edunov, Ottetal (2018). Understanding Back-Translation at Scale. Retrieved April 10, 2020, from ↗**https://arxiv.org/pdf/1808.09381v2.pdf**

5 Kocmi and Bojar (2019). Transfer Learning across Languages from Someone Else's NMT Model. Retrieved April 10, 2020, from ↗**https://arxiv.org/pdf/1909.10955.pdf**

6 Process integration is crucial for MT to be used successfully

The most important determinants of success in terms of the integration of AI in translation are how well the solution can be integrated into existing processes, and how well it is adapted to the needs of clients and professional translators. The main questions are:

1. **How much efficiency is gained in the process when employing AI as compared to the same process without AI?**

In terms of client-specific MT, the experiment described in ↗Section 3 clearly demonstrates how high this efficiency gain can be. Only if the increase in efficiency compared to existing solutions and software – i.e. the status quo – is high does it make sense to introduce AI into existing processes. In many other areas, however, rule-based software is still better suited.

2. **How high is the acceptance of the technology among all those involved?**

The introduction of new technology is rarely a guaranteed success, since human beings are and remain the decisive factor. The level of acceptance of new technology is a key driver for the potential success or failure of new technology. Therefore, users should already play a key role in the development phase, and the benefits for users should be communicated clearly and unambiguously. Professional translators have been using technology to support their work for decades, helping them increase consistency of language and making it easy to find and use previously translated content. The introduction of NMT is fundamentally transforming the work of linguists; from translators who produce content themselves to proofreaders who check and stylistically improve previously translated content. This transformation must be managed, and professional translators trained in the use of the new technology.

3. **How well can the solution be integrated into existing processes?**

In addition to its acceptance by end users, the integration into already existing ecosystems plays a major role in the success of any technology. In the specific case of NMT, this means, for example, that companies that order translations from service providers are not required to adapt to new file formats or processes. Rather, the service provider needs to be able to integrate into the existing environment and the already proven processes on the customer's side. This allows the potential of the technology to be fully exploited and increases acceptance.

4. **Are the companies and individuals involved »prepared« to set up feedback loops, to collect the necessary data on an ongoing basis and to thus guarantee the continuous training of the AI models?**

Depending on the training data available, NMT can deliver great efficiency gains with just the initial training of the models. The great added value lies, however, in the continuous collection of high-quality data and regular training cycles. The feedback loop required for this, i.e. the constant feeding of data into the models, is key to the successful integration of AI into existing business processes. The resulting AI learning effect, which results in the continuous improvement of quality, means that the technology can deliver real added value. The one-time training of an AI system is simply not enough. Particularly in dynamic environments such as linguistics, AI systems need to constantly learn and develop. In time, efficiency can be increased even further, as is shown clearly in Figure 3.

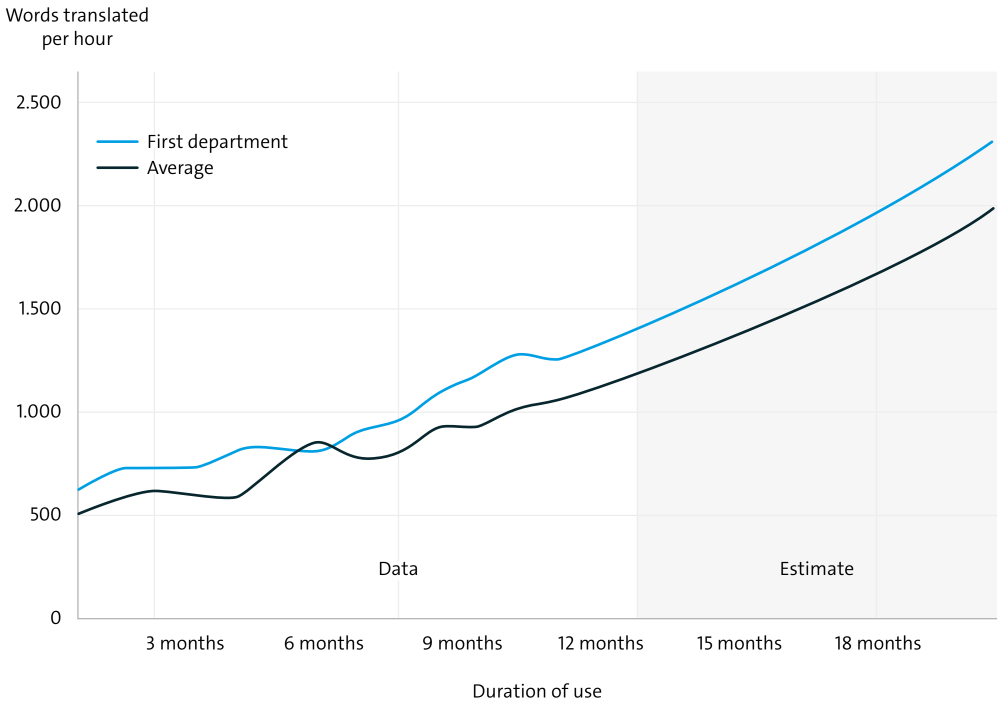

Abbildung 3: Increase in translation speed over time with continuous training of the models

The figure shows how many words per hour professional translators can translate when supported by continuously trained NMT models. As time progresses, it becomes very clear how efficiency increases each time the model is trained. The reason for this is simple: the better the translation by the NMT model gets, the less effort the professional translator needs to spend revising it. The two curves provide an internal comparison of data from a telecommunications company in Switzerland: the green curve shows the increase in efficiency for the first department, in which an NMT model was used, while the blue curve shows the average for all of the company's departments.

»The great added value of MT lies, however, in the continuous collection of high-quality data and in regular training cycles. The feedback loop required for this is key to the successful integration of AI into existing business processes.«

Implementing a process including an AI system requires a high level of expertise in data science and software development. In addition to the use of language models, NMT also uses specially developed software that is specifically designed for post-editing MT content. This can collect additional metadata to improve the linguists' performance, such as the number of words translated per hour as shown in Figure 3.

Innovations based on AI must be ready for use in the real world before they can be introduced. However, since companies often have little or no experience with AI, it is advisable to introduce the technology in collaboration with a full-service partner with the necessary technological expertise, and which can accompany the client each step of the way.

Especially when it comes to fine tuning and optimizing models and processes, working with an external service provider makes it possible to do just that: namely to facilitate the interaction between humans and machines.

The use of custom-trained NMT has immense potential as a means of raising productivity levels for professional translators. Using generic MT in place of human professional translators is not an acceptable alternative. Rather, the use of AI in the translation field is becoming a symbol of successful cooperation between humans and machines.

7 Outlook

The global demand for translation has more than doubled over the last ten years[6]. There are simply not enough professional translators in the world today to cope with the growing amount of content. It is thus obvious that the use of AI in the translation field could fill this supply gap, particularly in light of how rapidly general language and automatic translation applications such as Google Translate and the German company DeepL have developed in recent years. A closer look at the existing solutions shows that the use of AI in translation already offers great advantages in the professional language market, but also places very complex demands on developers.

The term »Artificial Intelligence« often suggests an independent factor that is equal to human intelligence and on the same level. However, in the field of translation in particular it is clear that the use of machine learning – the most important component of AI – should be used above all to support humans. So, perhaps the terminology is, in fact, wrong, and we should rather be referring to »Augmented Intelligence«, i.e. the extension of human intelligence through technology.

»According to UNESCO, about 2,400 of the world's 7,000 existing languages are currently in danger. Language models can support the active learning of these languages.«

Although there are huge resources of available language data for the major world languages such as English and Spanish, as well as many other European and Asian languages, this is not the case for the vast majority of languages around the world. By using techniques such as fully automated back translation, it is possible to apply MT here, too. The potential of AI in this field, though, has a different starting point.

UNESCO concludes that about 2,400 of the world's 7,000 independent languages are currently in danger, as only between a few hundred to some thousands of native speakers remain.[7] For these languages, language models can be used to support the remaining native speakers, promote the active learning of the language, and hence preserve the language in the long term.

6 Statista ↗**https://www.statista.com/statistics/257656/size-of-the-global-language-services-market/**

7 Moseley, C. 2010. Atlas of the World's Languages in Danger, 3rd edn. Paris, UNESCO Publishing

In the field of translation, AI will not replace professional human translators. Rather, it will become established as a supporting technology that makes a professional translator's work easier and increases productivity; a way of working in which humans and machines can work hand in hand to achieve better results than they would working alone.

8 Bibliography

1. Dinu et al. (2019). Training Neural Machine Translation To Apply Terminology Constraints. Retrieved April 10, 2020 from ↗**https://www.aclweb.org/anthology/P19-1294.pdf**

2. Edunov, Ott et al (2018). Understanding Back-Translation at Scale. Retrieved April 10, 2020 from ↗**https://arxiv.org/pdf/1808.09381v2.pdf**

3. Kocmi and Bojar (2019). Transfer Learning across Languages from Someone Else's NMT Model. Retrieved April 10, 2020 from ↗**https://arxiv.org/pdf/1909.10955.pdf**

4. Moseley, C. 2010. Atlas of the World's Languages in Danger, 3rd edn. Paris, UNESCO Publishing

5. Swiss Confederation (November 19, 2019). Report DeepL test. Retrieved April 5, 2020 from ↗**https://uepo.de/wp-content/uploads/2019/12/schweiz_bericht_deepl-test_2019.pdf**

6. Statista ↗**https://www.statista.com/statistics/257656/size-of-the-global-language-services-market/**

Author

Christopher Kränzler
As founder and CEO of Lengoo, a language tech company, Christopher Kränzler has made it his mission to shape the future of enterprise translations. Lengoo has developed a technology for neural machine translation based on a highly innovative training approach for artificial intelligence. The specialized translations, backed by AI, unite the precision of human creativity with the immense advantages of artificial intelligence. After graduating with a Bachelor's Degree in Business Engineering from Karlsruhe Institute of Technology (KIT), he went on to achieve a Master's Degree in the field of Data Science from Columbia University in New York. Christopher Kränzler has been a member of the board of directors of Bitkom e.V. since June 2019. Bitkom e.V. is Germany's digital association.

Bitkom represents more than 2,700 companies of the digital economy, including 2,000 direct members. Through IT- and communication services alone, our members generate a domestic annual turnover of 190 billion Euros, including 50 billion Euros in exports. The members of Bitkom employ more than 2 million people in Germany. Among these members are 1,000 small and medium-sized businesses, over 500 startups and almost all global players. They offer a wide range of software technologies, IT-services, and telecommunications or internet services, produce hardware and consumer electronics, operate in the digital media sector or are in other ways affiliated with the digital economy. 80 percent of the members' headquarters are located in Germany with an additional 8 percent both in the EU and the USA, as well as 4 percent in other regions of the world. Bitkom promotes the digital transformation of the German economy, as well as of German society at large, enabling citizens to benefit from digitalisation. A strong European digital policy and a fully integrated digital single market are at the heart of Bitkom's concerns, as well as establishing Germany as a key driver of digital change in Europe and globally.

Federal Association for Information Technology, Telecommunications and New Media e.V.

Albrechtstraße 10
10117 Berlin
P +49 30 27576-0
F +49 30 27576-400
bitkom@bitkom.org
www.bitkom.org

bitkom