

How Close Are We to Removing Language Barriers in Healthcare

by Marco Trombetti*

Abstract

Language barriers create obstacles in providing high-quality healthcare and maintaining patient safety. Existing solutions, like interpreter services, cultural mediators, and smartphone apps, aren't able to overcome barriers adequately due to high costs, privacy implications, or poor quality. Machine translation offers a real solution once it becomes able to provide human-quality translations at reasonable prices. New research from Translated, a pioneer in AI-powered translation services, shows that this will soon be possible, thanks to the perfect symbiosis between professional translators and machine translation. Machine translation can radically impact the healthcare industry in three main areas: the translation of information for the general public, the translation of specialist publications such as scientific papers, patents, and disease reports, and seamless doctor-patient communication and the gathering of patient feedback.

Keywords

language barriers, machine translation, artificial intelligence, open source, innovation.

1. Introduction: The Availability of Universal Communication Tools and Their Impact on Healthcare

Allowing everyone to understand and be understood in their own language is one of the most significant challenges for humankind. Achieving this will facilitate

unprecedented collaboration between human beings. Imagine, for example, if scientific research were available in every researcher's language as soon as it was published. We are not far from that, as proven by the research we conduct at Translated, a global language services provider and pioneer of AI-powered translation services. This groundbreaking research was presented at the last

annual conference of the Association for Machine Translation in the Americas (AMTA) in Orlando.

By analyzing the edits made by 136,000 of the best professional translators in the world to 2 billion sentences processed by automatic translation software (machine translation), for the first time in history, we were able to quantify the speed at which we are approaching

* Computer scientist, entrepreneur, and investor.

the singularity in translation. The singularity is reached when the best-performing professional translators spend the same time correcting a translation produced by machine translation (MT) as they do correcting one completed by their peers.

In the medical field, breaking language barriers is even more impactful than in other areas. These barriers prevent patients from clearly understanding their clinical conditions, thus making it difficult for them to adhere to necessary therapy correctly. In addition, these barriers make it hard, if not impossible, for doctors to acquire the required consent to care and limit data and information needed for diagnoses and research. Today, many larger healthcare institutions offer interpreter services, but these come at significant costs. Thus, most organizations in the medical sector still rely on cultural mediators or DIY solutions, such as smartphone apps. The Covid-19 pandemic has made the need to address this problem more important than ever. According to the European Commission¹, the pandemic increased the demand for translation by 49% in the healthcare industry. Fortu-

nately, technological advances are pushing the healthcare community to look at machine translation as a means of overcoming language barriers, and universal communication tools are close enough to provide adequate support.

We expect that machine translation will have a radical positive impact on the healthcare industry. Potential applications are in three main areas:

- The translation of information for the general public.
- The translation of specialist publications such as scientific papers, patents, and disease reports. This will provide access to global research studies and real-world data (e.g. clinical trials and drug discovery).
- Seamless doctor-patient communication and the gathering of patient feedback collection, even from discussions happening on social media among users. Here, we can combine MT with automatic speech recognition (ASR) and text-to-speech (TTS) technologies to support spoken language.

However, machine translation will perform well in healthcare only when it pro-

vides translations as good as those done by translation professionals. At Translated, we have been monitoring MT quality since 2011, and recently, we decided to use the massive amount of data we have collected to measure how far we are from providing human-quality machine translation. When we reach the singularity in translation, we can integrate real-time automatic translation in almost every device at a very accessible cost.

2. A Brief History of Machine Translation

The concept of automatic translation was mentioned for the first time in the 9th century when an Arabic cryptographer introduced techniques for systematic language translation that are, incredibly, still relevant². However, the first public demonstration of machine translation was done in 1954 in the United States³. It was a small experiment, but it encouraged researchers to push forward. Early systems relied on bilingual dictionaries and rules stating how to translate words or phrases from a source language into a target language. Next, a statistical approach was developed: by analyzing large volumes of human translations,

machines started predicting the equivalence of a phrase in the target language. The phrase-focused pattern-learning and pattern-forecasting approach drove the first version of Google Translate in the early 2000s.

Today, Google Translate and the most advanced machine translation engines rely on deep learning-based neural network models to learn and predict final outputs. This is a deeper, more reliable form of pattern detection and forecasting. In this kind of system, the translation is produced by a single sequence model trained to predict one word at a time, considering the entire source sentence and the translation that has already been provided.

In 2017, a consortium comprised of Translated, the Fondazione Bruno Kessler, the University of Edinburgh, and TAUS introduced the first adaptive machine translation, ModernMT⁴. It was initially a research project backed by the European Union⁵ that later became open-source software and a commercial service powered by Translated. In this new model, MT learns in real-time from the translator's corrective feedback without retraining the translation model.

The idea for adaptive machine translation dates back to a previous research project conducted by Translated, the Fondazione Bruno Kessler, the University of Edinburgh, and Le Mans University, sponsored again by the European Union⁶. The initial idea was to create a tool to edit machine translation results and manage localization workflows. The research objective was an MT system that could learn from translators' corrections and automatically improve over time. The MT component of the solution was later separated to focus on the editing tool, which was finally released as open-source software in 2014. The European Commission included the project amongst those with the highest potential for innovation funded by the Seventh Framework Program. Translated further refined the research prototype and created a commercial version, Matecat⁷, a free-to-use computer-assisted translation tool and MT editing software that the company adopted as its exclusive production tool. With Matecat and ModernMT, Translated is pushing hard for a perfect symbiosis between human creativity and machine intelligence: by removing

redundant tasks, AI allows professionals to focus on the nuances of language, improving the quality of the translation. This synergy gives linguists better suggestions while MT keeps learning. Together, they become more efficient, adaptable, and cost-effective every day.

3. About the Process and the Data Collected

In 2011, Translated standardized and settled on a highly reliable metric to evaluate MT quality accurately. We call it Time to Edit (TTE): this is the average time per word required by the best-performing professional translators to check and correct MT-suggested translations. This makes it possible to switch from automated estimates still in use in the industry to measurements of human cognitive effort, reassigning the quality evaluation to persons traditionally responsible for the task: professional translators. We have been tracking Time to Edit for almost a decade, collecting over 2 billion edits on sentences effectively translated by 136,000 professional translators worldwide working across multiple subject domains, ranging from literature to tech-

nical translation and including fields in which MT is still struggling, such as speech transcription. The linguists were selected for the specific jobs they completed using proprietary AI called T-Rank⁸, which gathers work performance and qualification data on over 300,000 freelancers who have worked with the company over the last two decades. The AI considers over 30 factors, including résumé match, quality performance, on-time delivery record, availability, and expertise in domain-specific subject areas.

Working in Matecat, translators check and correct translation suggestions provided by the MT engine of their choice. The data was initially collected using Google's statistical MT (2015-2016), then Google's neural MT, and most recently by Modern MT's adaptive neural MT, introduced in 2018, which quickly became the preferred choice amongst almost all our translators. Translated has been collecting the average time to edit a word continuously for over seven years.

To refine the sample, we only considered the following:

- Completed jobs delivered at a high level of quality.

- Sentences with MT suggestions that had no match from databases of previously translated segments of text.
- Jobs in which the target language has a vast amount of data available along with proven MT efficiency (English, French, German, Spanish, Italian, and Portuguese).

From the resulting pool of sentences, we removed:

- Sentences that didn't receive any edits since they did not provide information about TTE, and sentences that took more than 10 seconds per word to be edited, as they suggest interruptions and/or unusually high complexity. This refinement was required to make TTE comparison possible across multiple years.
- Work on adaptation to locale, i.e. translations between variants of a single language (e.g. British English to American English), since these are not representative of the problem at hand.
- Large customer jobs, since they employ highly customized language models and

translation memories in which TTE performance is far better than average.

Time to Edit is impacted by two main variables other than MT quality: the evolution of the editing tool and the quality delivered by the translator. The influence of these two factors can be considered negligible when considering the long-run trend of improvement we observed.

4. A Surprising Linear Trend Approaching the Singularity in Translation

When plotted graphically, the TTE data shows a surprisingly linear trend. Our initial hypothesis to explain this is that every unit of progress towards closing the quality gap requires exponentially more resources than the previous unit, and we accordingly deploy those resources: computing power (doubling every two years), data availability (the number of words translated increases at a compound annual growth rate of 6.2% according to Nimdzi Insights), and the efficiency of machine learning algorithms (computation needed for training, 44x improvement from 2012-2019, according to OpenAI⁹).

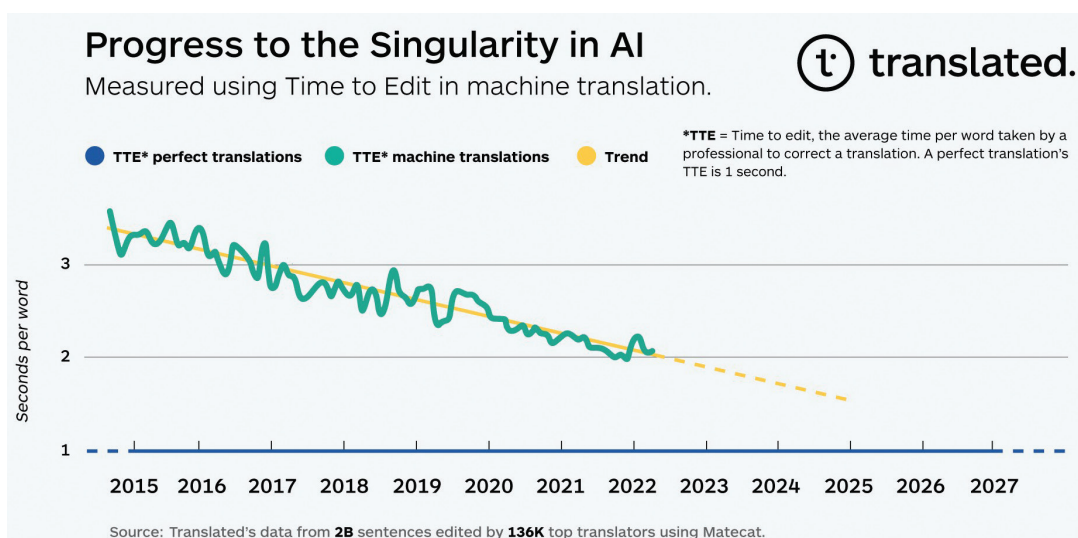


Fig. 1.

5. Conclusion: How Close We Are to Breaking Language Barriers

If progress in machine translation quality continues with the current trend, in about six years the best-performing professional translators will spend the same time correcting a translation produced by machine translation as they do correcting one completed by their peers. The exact date when we will reach the singularity in

translation could vary somewhat, but the trend is clear. We are therefore close to being able to provide real-time, universal, accessible translation tools that will break the language barriers, allowing us to improve clients' health outcomes, lowering the risk of death.

From a research point of view, the evidence Translated has provided about the progress in MT quality is quite possibly the most compelling evidence of success at scale seen

in both the MT and AI communities in general. Indeed, many AI researchers think that solving the language translation problem is equivalent to producing artificial general intelligence (AGI). Translated's discovery has thus quantified, for the first time in history, the speed at which we are approaching the singularity in artificial intelligence – the hypothetical future point in time at which artificial intelligence transcends human intelligence.

Notes

1. <https://blogs.ec.europa.eu/emt/covid-19-how-has-it-affected-the-world-of-translation/>.
2. DuPont Q. (2018), *The Cryptological Origins of Machine Translation: From al-Kindi to Weaver*, «Amodern», Issue 8, Translation-Machination.
3. <https://www.semanticscholar.org/paper/The-first-public-demonstration-of-machine-%3A-the-%2C-7-Hutchins/ado14a7b7a3142e6f17eedddf4218489b56ab18e>.
4. <https://modernmt.com>. More information on the project at <https://slator.com/business-academia-join-forces-launch-next-gen-machine-translation/>.
5. EU Horizon 2020 Innovation Action (2015-2017).
6. MateCat, acronym of Machine Translation Enhanced Computer Assisted Translation, is a 3-year research project (11/2011 – 10/2014) funded by the European Union's Seventh Framework Programme for research, technological development and demonstration.
7. <https://matecat.com>.
8. <https://translated.com/t-rank>.
9. <https://openai.com/blog/ai-and-efficiency/>.

References

- Turchi M., Negri M., Farajian A., Federico M. (2017), *Continuous Learning from Human Post-edits for Neural Machine Translation*, Proceedings of the 20th Annual Conference of the European Association for Machine Translation, n. 108, June 2017.
- Bertoldi N., Cattoni R., Cettolo M., A.F., Federico M., Caroselli D., Mastrostefano L., Rossi A., Trombetti M., Hermann U., Madl D. (2017), *MMT: New Open Source MT for the Translation Industry*, Proceedings of the 20th Annual Conference of the EAMT.
- Farajian M.A., Turchi M., Negri M., Bertoldi N., Federico M. (2017), *Neural vs. Phrase-Based Machine Translation in Multi-Domain Scenario*, Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, Volume 2.
- Bentivogli L., Bisazza A., Cettolo M., Federico M. (2018), *Neural versus Phrase-Based MT Quality: an In-Depth Analysis on English-German and English-French*, «Computer Speech and Language», Volume 49, May 2018.
- Bentivogli L., Bisazza A., Cettolo M., Federico M. (2016), *Neural versus Phrase-Based Machine Translation Quality: a Case Study*, Proceedings of Conference on Empirical Methods in Natural Language Processing, November 1-5, 2016.
- de Souza J.G.C., Federico M., Sawaf H. (2015), *MT Quality Estimation for E-Commerce Data*, Proceedings of MT Summit XV (AMTA 2015), October 30 – November 3, 2015, pp. 20-29.
- de Souza J.G.C., Turchi M., Negri M., Anastasopoulos A. (2014), *Online and Multitask learning for Machine Translation Quality Estimation in Real-world scenarios*, Italian Computational Linguistics Conference (CLIC-it 2014), December 9-10, 2014.
- Federico M., Negri M., Bentivogli L., Turchi M. (2014), *Assessing the Impact of Translation Errors on Machine Translation Quality with Mixed-effects Models*, Conference on Empirical Methods in Natural Language Processing (EMNLP 2014), October 25-29, 2014, pp. 1643-1653.
- de Souza J.G.C., Turchi M., Negri M. (2014), *Machine Translation Quality Estimation Across Domains*, 25th International Conference on Computational Linguistics (COLING 2014), August 23-29, 2014, pp. 409-420.
- Girardi C., Bentivogli L., Farajian M.A., Federico M. (2014), *MT-EQuAl: a Toolkit for Human Assessment of Machine Translation Output*, 25th International Conference on Computational Linguistics (COLING 2014), August 23-29, 2014, p. 120-123.
- Federico M., Bertoldi N., Cettolo M., Negri M., Turchi M., Trombetti M., Cattelan A., Farina A., Lupinetti D., Martines A., Massidda A., Schwenk H., Barrault L., Blain F., Koehn P., Buck C., Hermann U. (2014), *The MateCat Tool*, 25th International Conference on Computational Linguistics (COLING 2014), August 23-29, 2014, pp. 129-132.
- Turchi M., Negri M., Federico M. (2014), *Data-driven Annotation of Binary MT Quality Estimation Corpora Based on Human Post-editions*, «Machine Translation», Volume 28, Issue 3-4, December 2014, pp. 281-308.
- de Souza J.G.C., González-Rubio J., Buck C., Turchi M., Negri M. (2014), *FBK-UPV-UEdin participation in the WMT14 Quality Estimation shared-task*, 9th ACL Workshop on Statistical Machine Translation (WMT 2014), June 26-27, 2014, pp. 322-328.
- Turchi M., Anastasopoulos A., de Souza J.G.C., Negri M. (2014), *Adaptive Quality Estimation for Machine Translation*, 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014), June 22-27, 2014, pp. 710-720.
- Turchi M., Negri M., Federico M. (2013), *Coping with the Subjectivity of Human Judgements in MT Quality Estimation*, Proceedings of the 8th Workshop on SMT, August 8-9, 2013, pp. 240-251.

de Souza J.G.C., Esplà-Gomis M., Turchi M., Negri M. (2013), *Exploiting Qualitative Information from Automatic Word Alignment for Cross-lingual NLP Tasks*, Proceedings of ACL 2013, August 4-9, 2013.

MateCat: Machine Translation Enhanced Computer Assisted Translation, Project Presentation, Proceedings of the European Conference on Machine Translation (EAMT), May 28-30, 2012.