Process Implementation in Post-editing to Ensure Quality in Translation

Gabriela Escarrá[12]

Universidad Nacional de La Plata

y Dolores R. Guiñazú[13]

Universidad del Salvador

Albany University in New York

✉ gabriela@langtic.com

✉ dolores@liftvalue.com

Recibido: 07/5/2021

Aprobado: 30/8/2021

Abstract

The translation profession is constantly evolving and becoming more technological to meet the needs of the increasing demand for translation and localization services. The focus should not be on when human linguists will become obsolete, but rather on the competences needed to face this challenge. Thus, the improvement in the quality of Artificial Intelligence-enabled Translation and Machine Translation (MT) are taking over the forefront of technological advances and have compelled translators to adapt to new environments and to learn and adjust their own processes

[12] **Gabriela Escarrá** is a Certified Legal Translator (UNLP), English and Literature Teacher (UNLP), and ISO 27001 Internal Auditor. She is a member of the ISO member body in Argentina. She has two certifications from TAUS. She works as a translator, editor, post-editor and transcreator, and has been lecturing on quality management, international quality standards, and post-editing worldwide.

[13] **Dolores R. Guiñazú** is a Certified Legal Translator and Copyeditor. Dolores is ISO 27001:2013 Internal Auditor and is an active member of IRAM (ISO) on Translation & Interpreting. She holds an MBA in marketing management She has two certifications from TAUS in post-editing and transcreation. She has been lecturing on quality management, international quality standards, and post-editing worldwide.

and requirements. To that purpose, it is imperative for Translation Service Providers (TSP) and linguists in general to learn and apply quality processes stipulated by ISO standards on Translation, Post-Editing, Terminography and Information Security Management System (ISMS).

Keywords*:* machine translation (MT), post-editing (PE), quality in translation, ISO 18578, artificial intelligence (AI)

Resumen

La profesión del traductor está en constante evolución y es cada vez más tecnológica para satisfacer las necesidades de la creciente demanda de servicios de traducción y localización. La atención no debe centrarse en cuándo la lingüística humana se volverá obsoleta, sino en las competencias necesarias para afrontar este desafío. Así, la mejora de la calidad de la traducción producida por inteligencia artificial (IA) y la traducción automática (TA), que están a la vanguardia de los avances tecnológicos, han obligado a los traductores a adaptarse a nuevos entornos y a aprender a ajustar sus propios procesos y requisitos. Para ello, es imperativo que los proveedores de servicios de traducción (PST) y los lingüistas en general aprendan y apliquen los procesos de calidad estipulados por las normas ISO sobre traducción, posedición, terminografía y sistema de gestión de seguridad de la información (SGSI).

Palabras clave: traducción automática (TA), posedición (PE), calidad en la traducción, ISO 18578, inteligencia artificial (IA)

Introduction

The translation profession has changed considerably over the last few decades. To meet the needs of the increasing demand for translation and localization services, the use of new technologies, including Computer-Assisted Tools (CAT) and Machine Translation (MT) has become imperative in many domains of texts, with a few exceptions like Poetry where Machine Translation (MT) can hardly be used. Thus, the mastery of these new technologies is becoming a *sine qua non* for staying competitive. Moreover, the new generation of translators should not only be trained on how to transfer the context of a text from the source language into the target language effectively and precisely but also on the necessary skills and processes for using the latest technological tools to provide the best possible quality in their translation assignments.

The compliance with requirements of Post-Editing Machine Translation (PEMT) output and the implementation of processes are as essential as applying the knowledge and skills in translation and the use of technologies related to translation. Thus, the application of quality standards by professional translators and Language Service Providers (LSPs) alike should be encouraged at all times. The purpose of this article is to broaden the knowledge on Post-Editing (PE) and the application of ISO quality standards and their recommended processes with the aim of providing the best translation quality, while ensuring a proper handling of terminography and confidentiality.

Discussion

Before deep diving into the use of ISO quality standards and the processes stipulated by them, it may be relevant to analyze the background of the advances in the field of new technologies, including Artificial Intelligence (AI) and Machine Learning (ML).

Even though we may tend to believe that Artificial Intelligence (AI) is something relatively new, it dates back to 1956, when John McCarthy coined this word and defined it as "The science and engineering of making intelligent machines". The study of John McCarthy stated that every aspect of learning or any other feature of human intelligence can in principle be so precisely described that a machine can be made to simulate it. This happened when McCarthy and some colleagues grasped the knowledge of machine intelligence and the ability it has to imitate human intellect.

Different research and contradictory findings fueled our fear, as translators of the future of AI. Human intelligence is not competing with another form of non-human intelligence. We have to understand the benefits of AI for our professional and personal lives. Different research and contradictory findings have been fueling our fears in terms of Artificial Intelligence (AI) overcoming Human Intelligence (HI). However, Artificial Intelligence (AI) can be considered as another tool for human translators.

There has been a turn into human augmented intelligence revolution. The world has shifted from the Internet of Things (IoT) to Artificial Intelligence of Things (AIoT). As professional translators, we must be ready to embrace this revolution**.** We have to consider "human augmented intelligence" to understand how we can be supported and equipped to work better. If we can become augmented professionals, we may not need to perform repetitive tasks and we will have more time to focus on the creative aspect of the translation process.

The application of AI in languages has resulted in the rise of different types of Machine Translation (MT): Rule-Based Machine Translation (RBMT), Example-Based Machine

Translation (EBMT), Neural Machine Translation (NMT), and Hybrid Machine Translation (HMT).

Rule-Based Machine Translation (RBMT) was developed several decades ago. It was the first practical approach to Machine Translation (MT) in which the rules attempt to define correspondences between the structure of the source language and that of the target language. RBMT engines do not deal well with slang or metaphorical texts. The main setback of this system is that it is time-consuming.

Statistical Machine Translation (SMT) works by training the translation engine with a very large volume of bilingual (source texts and target texts, i.e., translations) and monolingual corpora. The system looks for statistical correlations between source texts and translations. This engine is the core of systems used by Google Translate and Bing Translator, and is the most common form of MT in use today.

In Example-Based Machine Translation (EBMT), a sentence is translated by analogy. It consists in the use of a bilingual corpus with parallel texts. A number of existing translation pairs of source and target sentences are used. When a new source sentence has to be translated, the examples are retrieved so as to find similar ones in the source, then the target sentence is generated by imitating the translation of the matching examples. This approach may result in a high-quality translation when very similar examples are found. On the contrary, when there is no similar example, the quality is very low.

Neural Machine Translation (NMT) is based on the paradigm of machine learning. This is the newest approach to Machine Translation (MT). Neural Machine Translation (NMT) uses neural networks that consist of nodes. The nodes can hold single words, phrases, or longer

segments and relate to each other in a web of complex relationships based on bilingual texts used to train the system. The Neural Machine Translation (NMT) system is continuously learning and adjusting to provide the best output. This system requires a lot of processing. Neural Machine Translation (NMT) is an end-to-end learning approach for automated translation. There is still a lot of potential in this translation system.

Hybrid Machine Translation (HMT) is a system that combines different approaches of Machine Translation. The Hybrid Machine Translation (HMT) system is based on two main types of Machine Translation (MT): Rule-Based Machine Translation (RBMT) and Statistical Machine Translation (SMT). Nowadays, almost all Machine Translation Systems (MTS) choose hybrid approaches.

Post-editing and How to Measure Machine Translation (MT) Quality

Though there are some quality parameters to measure translation quality, there is no absolute measure so far. The standard practice is to work with the final human eye to review, edit, and enhance texts. The most reliable method of Machine Translation (MT) quality assessment requires human evaluators to score each sentence.

Adequacy and fluency are widely used metrics for human scoring of Machine Translation (MT). The final human eye is expensive and time consuming. However, nowadays it is the best approach to ensure the quality of translations.

Some automatic methods of assessment measure similarity of Machine Translation (MT) and Human Translation (HT) output, for instance: WER (Word Error Rate) in reference to the translation at word level; PER (Position independent error rate) the word error is calculated by

considering each sentence as a separate unit of words, independent of the word order; Bilingual Evaluation Understudy (BLEU) takes into account precision, and not word error rate.

ISO 17100:2015 on Translation and ISO 18587:2017 on Post-editing

The International Organization for Standardization (ISO*)* along with translation professionals have timely addressed the latest developments in the translation field by issuing two translation-related international standards, ISO 17100:2015 /AMD 1: 2017 on Translation Services, and ISO 18587:2017 on Post-Editing of Machine Translation Output, which establishes the framework and requirements for post-editors of Machine Translation (MT) output.

ISO 17100 was based on EN15038 but, among other things, it incorporates the requirement for translations to be subject to revision by a second linguist; it pays particular attention to the qualification of all parties involved in the production process of a translation; and, it clarifies that the responsibility for the preparation does not just rest on the Language Service Provider (LSP), but also on the client. Finally, in comparison to EN 15038, ISO 17100 requires having a process for handling client feedback, which may serve as an indicator of the client's satisfaction and as a basis for future improvement.

ISO 18587:2017, which is based on ISO 17100, requires the same degree of professionalism for post-editors as for translators. It is intended to be used by post-editors, Translation Service Providers (TSPs) and their clients. It specifies the requirements for the Post-Editing output (including terminology, formatting, consistency, style) and post-editors' competences to ensure the same quality of human translation. As stated by this quality standard, post-editors need to be thoroughly trained on the use the post-editing tools, the recognition of

common Machine Translation (MT) errors, the assessment of Machine Translation (MT) output to determine whether it is worth spending time and effort on a post-editing project; and the difference between the full and light post-editing processes and the final outcome.

Moreover, after performing all necessary edits on Machine Translation (MT) output, ISO 18587 states that post-editors should ensure the best quality of post-editing tasks by using the available tools to perform quality checks at the end of the project and providing feedback on Machine Translation (MT) output to contribute to the improvement of MT systems. In this way, both Human Translation (HT) and Machine Translation (MT) errors cannot only be detected and modified during the post-editing process but also avoided in future tasks.

ISO 12616:2002 on Translation-oriented Terminography

According to the Collins dictionary, *terminography i*s a compilation of the terminology used in a specific field.

ISO 12616:2002 is the Translation-oriented Terminography standard. The three main objectives of this standard are:

1. To record and systematize terminology.
2. To use terminology consistently over time.
3. To deal more efficiently with different languages.

Terms (unique words for a certain field) differ from words (all the units used when writing a document). All the terms are words, but not all the words are terms. Specially for long projects, having a reliable term base helps linguists to prepare a consistent and seamless final translation work.

ISO/IEC 27000 Family on Information Security Management System (ISMS)

Information security and confidentiality is a very serious issue in the context of using Machine Translation (MT). The practice of collecting, analyzing and sharing private data has raised confidentiality concerns. Many governing bodies have incorporated privacy concerns into regulation and legislation. Perhaps the most extensive effort to date is the European Union's General Data Protection Regulation (GDPR).

As stated by ISO 17100, Translation Service Providers (TSPs) shall have an infrastructure to ensure a safe and confidential handling, storage, retrieval, archiving and disposal of all relevant data and documents. Moreover, for a deep understanding of the requirements for an Information Security Management System (ISMS), Translation Service Providers (TSPs) can refer to the ISO/IEC 27000 Family, mainly ISO/IEC 27001:2013 and ISO/IEC 27014:2020 on information security, cybersecurity and privacy protection.

Conclusion

As it has been shown in this paper, the translation profession has been transformed by the latest technological advances. Therefore, with the advent of Machine Translation (MT) and new technologies, quality does not only depend on the professional linguistic skills and competences but also on the awareness of the importance of proper handling of terminography and confidentiality and the compliance with ISO standards.

References

Atlam, H. F., Walters, R.J., and Wills, G. B.(July 2018). Intelligence of things: Opportunities & challenges, 10.1109/CIOT.2018.8627114

Doherty, S. (2016). The impact of translation technologies on the process and product of translation. *International Journal of Communication*10, pp 947–969.

El-Tawil, A. (2014). *Standards & quality book*. World Scientific.

Hamel, G. (2020). *Humanocracy*. HBR Press.

Hendrik Kockaert (2015) *Handbook of terminology*. John Benjamins Publishing Company.

ISO 12616 (2002). Translation-oriented terminography.

ISO 17100 (2015). Translation services requirements for translation services. International Standardization Organization.

ISO 18587 (2017). Translation services –Post-editing of machine translation output – Requirements. International Organization for Standardization.

ISO/IEC 27000 Family. Information security management –International Organization for Standardization.

ISO/IEC 27014 (2020). Information Security, Cybersecurity and Privacy Protection – Governance of Information Security.

ISO/IEC 27001 (2013). Information Technology, Security Techniques, Information Security Management Systems – Requirements.

Kohen P. (May 2020). Neural machine translation. https://bit.ly/304si5I

McCarthy, J. (1996). Defending AI research: A collection of essays and reviews. CSLI Publications.

Madankar, M. (2016). Information retrieval system and machine translation: A review. ELSEVIER.

Nunes Vieira, L.; Alonso, E. (May 2018). The use of machine translation in human translation workflows: Practices, perceptions and knowledge exchange. https://bit.ly/3bQpKtU

Van der Meer, J. (February 2016). The future does not need translators. (consulted 6/18/2021) TAUS, https://bit.ly/30h4Wcv.