Machine Translation Guide

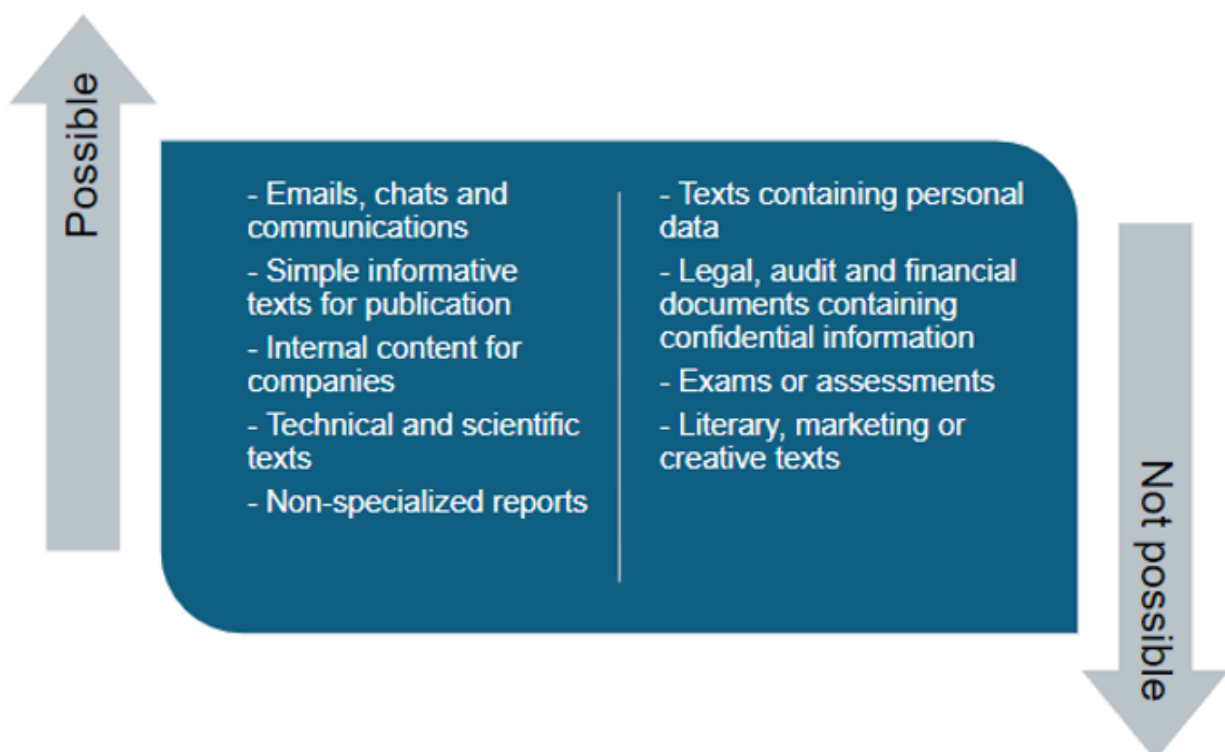**Lingua Viva |** Feb 5, 2024

Introduction

What is machine translation?

Machine translation is the process of using software or algorithms to automatically translate content from one language to another. This process is carried out through software or APIs that connect to Translation Management Systems (TMS) like Phrase or Trados.

Machine translation can work in different ways, depending on the type of system used. Some machine translation systems use predefined rules for grammar and syntax, while others use statistical or machine learning models to produce more accurate translations.

Although machine translation has advanced significantly in recent years, there are still some limitations in terms of its accuracy and ability to capture the context and intention of the source text. For this reason, machine translation differs from human translation, as the latter requires the knowledge and skill of a human translator to capture the intent of the source text, as well as the cultural and linguistic connotations and nuances, which can be difficult for a machine to convey.

Machine translation use cases

Possible

- Emails, chats and communications
- Simple informative texts for publication
- Internal content for companies
- Technical and scientific texts
- Non-specialized reports

- Texts containing personal data
- Legal, audit and financial documents containing confidential information
- Exams or assessments
- Literary, marketing or creative texts

Not possible

Translation of web content: Machine translation can be useful for translating web content like news, articles and blogs, as the content is generally informative and does not require a high degree of accuracy.

Translation of emails and chats: Machine translation can also be useful for translating emails and chats in real time when communication is the priority and there is not enough time to wait for human translation.

Translation of internal content: Machine translation can be useful for translating internal content, such as internal company documents or press releases, where accuracy is less essential and speed is important.

Translation of technical and scientific content: In some cases, machine translation can be useful for translating technical and scientific content. It can have sufficient or at least acceptable accuracy when it comes to translating terms which could be time-consuming for a translator. This is especially relevant if these terms are repeated throughout the project or across different texts.

Ethical and legal considerations for machine translation

There are several cases where it would not be ethical or legal to use machine translation, including the following:

Translation of legal and official documents: Machine translation is not appropriate for the translation of legal or official documents. These could include: contracts, wills, immigration documents, etc. In these cases a high degree of accuracy is required and any translation errors could have significant legal consequences. It is important to use a certified and skilled human translator to ensure the quality and accuracy of the translation.

Translation of confidential content: Machine translation is not appropriate for the translation of confidential content, such as confidential business information, internal company documents, personal data etc. Machine translation offers no guarantee of confidentiality and could compromise the security and privacy of the information.

Translation of culturally specific content: Machine translation is not appropriate for the translation of culturally specific content, such as literature, poetry, marketing and advertising, where cultural adaptation is essential to accurately convey the message. Machine translation cannot understand the cultural and linguistic context and could produce inappropriate or even offensive translations.

Translation without consent: It is important to note that machine translation should not be used to translate content without the consent of the author or owner of the content. Translation without consent could violate copyright and privacy rights.

Pre-editing

Pre-editing in machine translation (MT) refers to the process of modifying or optimizing a source text before it is input into a machine translation system. The goal is to improve the quality of the machine's output by making the text easier for the system to process.

Benefits of pre-editing

- **Improves MT Output**: Pre-editing helps to improve the fluency, accuracy, and readability of the machine-translated text.
- **Reduces Post-editing Effort**: By enhancing the quality of the initial translation, pre-editing can significantly reduce the time and effort required for post-editing.
- **Consistency**: Pre-editing ensures that consistent terminology and style are maintained throughout the text.

Key aspects of pre-editing

1. **Simplification of Language**: Removing long, convoluted sentences and unnecessary words.
2. **Use of Consistent Terminology**: Ensuring that terms are used consistently across the source text, using client glossaries.
3. **Avoiding Ambiguity**: Ambiguous words or phrases can lead to incorrect translations. Changing the verbs to the active voice where possible. Avoiding pronouns and repeating the noun to make it clearer.
4. **Avoiding Idiomatic Expressions**: Idioms, phrasal verbs and culturally specific phrases often don't translate well in MT. Pre-editing involves replacing these with more literal expressions.
5. **Correcting Grammar and Spelling Errors**: Typos, grammatical mistakes, or punctuation issues can confuse MT systems, leading to poor output. Pre-editing ensures the text is error-free.

Pre-editing in practice

MT output without pre-editing

Estimada Laura: ✕

Espero que te encuentres muy bien.

Disculpa la demora en mi respuesta, nos encontrábamos revisando nuestra disponibilidad para las fechas requeridas y, teniendo en cuenta que mencionas que es un trabajo similar al del año pasado en el que prestamos el servicio con dos intérpretes, te confirmo que contamos con disponibilidad de ellos apoyarlos en esta auditoría y adjunto te comparto sus hojas de vida.

Por otro lado, también te envío la cotización preliminar del servicio.

Quedo atenta a cualquier consulta que puedas tener.

Cordialmente,

Dear Laura:

I hope you are feeling very well.

Sorry for the delay in my response, we were reviewing our availability for the dates required and, considering that you mention that it is a similar job to last year in which we provided the service with two interpreters, I confirm that we have availability to support them in this audit and I share their resumes.

On the other hand, I am also sending you the preliminary quote for the service.

I remain attentive to any questions you may have.

Cordially yours,

Pre-editing suggestions

Estimada Laura: *Change to a comma*

Espero que te encuentres muy bien. *Espero que estés bien.*

Disculpa la demora en mi respuesta, nos encontrábamos revisando nuestra disponibilidad para las fechas requeridas y, teniendo en cuenta que mencionas que es un trabajo similar al del año pasado en el que prestamos el servicio con dos intérpretes, te confirmo que contamos con disponibilidad de ellos apoyarlos en esta auditoría y adjunto te comparto sus hojas de vida. *This is quite a long and complex sentence, so it could be simplified or split into shorter sentences.*

Por otro lado, *remove* también te envío la cotización preliminar del servicio.

Quedo atenta a cualquier consulta que puedas tener. *Por favor, déjame saber si tienes alguna pregunta.*

Cordialmente, *Saludos cordiales,*

MT output with pre-editing

Estimada Laura,

Espero que estés bien.

Disculpa la demora en mi respuesta, nos encontrábamos revisando nuestra disponibilidad para las fechas requeridas. Teniendo en cuenta que mencionas que es un trabajo similar al del año pasado en el que prestamos el servicio con dos intérpretes, te confirmo que contamos con disponibilidad de ellos apoyarlos en esta auditoría y adjunto te comparto sus hojas de vida.

También, te envío la cotización preliminar del servicio.

Por favor, déjame saber si tienes alguna pregunta.

Saludos cordiales,

Dear Laura,

I hope you are well.

Sorry for the delay in my response, we were reviewing our availability for the dates requested. Taking into account that you mention that it is a similar job to last year's, in which we provided the service with two interpreters, I confirm that we have the availability of them to support you in this audit and I am enclosing their resumes.

I am also sending you the preliminary quote for the service.

Please let me know if you have any questions.

Best regards,

Post-editing

Machine translation post-editing (MTPE) is the process of reviewing and editing a machine translation by a human translator before it is shared with the client. The aim of MTPE is to improve the quality and accuracy of the machine translation, correcting errors and adapting the text to the needs of the target audience and the context in which it will be used.

MTPE involves a careful analysis of the text produced by the machine translation system, whereby the human translator assesses the grammatical accuracy, consistency and relevance of the content for the target audience. In MTPE, the translator can change, add or delete words or phrases to improve the quality of the translation and to make it more accurate and natural.

MTPE is commonly used in the translation industry to improve efficiency and to reduce costs, as it enables human translators to work with machine-generated translations. This enables them to focus on reviewing and improving the translation, rather than having to start from scratch. Moreover, the MTPE can also improve translation consistency and quality in high volume projects, or those with a tight deadline.

In summary, machine translation post-editing is a process of reviewing and editing a machine translation by a human translator to improve the quality and accuracy of the translation, adapting it to the needs of the target audience and the context in which it will be used.

Procedure for post-editing into English

1. **Review source text**: Understand the original content thoroughly. Ensure that you grasp the meaning, context, and nuances before reviewing the machine-translated (MT) text.
2. **Evaluate machine translation**: Assess the accuracy and coherence of the machine-translated text. Aim to use as much of the raw MT output as possible while ensuring the translation is semantically correct and clear.
3. **Correct errors**: Address grammatical, semantic, and structural errors. Ensure that no information has been accidentally added or omitted. Pay close attention to pronouns, subjects of verbs, and verb tenses.
4. **Ensure consistency**: Verify that terminology and style are consistent throughout the text. Refer to the client's glossary to ensure the correct terms are used. Check for any untranslated terms, and ensure that proper nouns, acronyms, and other DoNotTranslate elements are handled appropriately.
5. **Enhance clarity and readability**: Simplify complex sentences, clarify ambiguous phrases, and improve flow. Avoid making changes that are purely stylistic unless necessary for clarity.
6. **Adapt for culture and language**: Adjust expressions, idioms, and cultural references to make them suitable for English-speaking audiences. Remove or edit any offensive, inappropriate or culturally unacceptable content.

7. **Handle structural elements:** Preserve formatting and structure for lists, tables, headers, and body text. Ensure translations match the original layout.
8. **Maintain contextual accuracy**: Ensure the translation accurately reflects the intended context and meaning.
9. **Proofread**: Check for spelling, punctuation, and formatting errors. Follow basic rules regarding spelling, punctuation, and hyphenation to ensure the text is error-free.
10. **Quality check**: Perform a final review to meet quality standards and audience expectations.

Tips for post-editors

- **Achieving "good enough" quality:** Focus on producing a translation that is semantically correct and accurate. Avoid unnecessary changes that are solely stylistic unless they significantly enhance clarity or readability.
- **Use as much of the MT output as possible:** Leverage the raw MT output where it is accurate, making only essential corrections. This helps maintain efficiency while ensuring quality.
- **Check locale-specific conventions:** Always check and adapt measurements, punctuation, and date formats to match the target locale's conventions. Ensure that commas and full stops are used correctly in numbers (e.g., 1,000 vs. 1.000), and dates are formatted according to the target audience's norms (e.g., MM/DD/YYYY vs. DD/MM/YYYY).
- **Review grammar and syntax:** Correct any grammatical errors, including agreement (e.g., singular vs. plural), negations, and word order to maintain accuracy. Focus on clarity without unnecessary sentence restructuring. Watch for and eliminate any duplicates or omissions in the translation. Any errors can significantly impact the perceived quality of the translation.
- **Focus on context and meaning:** The primary goal is to ensure that the translation conveys the intended meaning and context of the source text accurately and clearly.

Post-editing exercises

These examples include some common mistakes that machine translation can make. Write your correction in the Target EN correction and an explanation of what you have changed. You can look up terms and organizations to help you if needed. Use **/j/** to signify joining segments together if needed.

1.

Source ES	Target EN	Target EN correction	Explanation
3.000.000 Acuerdos	3.000.000 Agreements		

2.

Source ES	Target EN	Target EN correction	Explanation
Ella recibió su título en lenguas modernas	She received her title in modern languages		

3.

Source ES	Target EN	Target EN correction	Explanation
La responsable del equipo les permitió marcharse antes.	The team leader allowed them to leave early.		
Se fue poco después.	He left shortly afterwards.		

4.

Source ES	Target EN	Target EN correction	Explanation
La formación como profesional se fundamenta principalmente en tres áreas, articuladas a través de la formación aplicada y experiencial con sentido crítico y de compromiso social.	Training as a professional is based mainly on three areas, articulated through applied and experiential training with a critical and socially committed sense of social commitment.		

5.

Source ES	Target EN	Target EN correction	Explanation
ecoturismo, pesca artesanal, piangua y economía circular.	ecotourism, artisanal fishing, piangua and circular economy.		

6.

Source ES	Target EN	Target EN correction	Explanation
Mecanismos financieros	Financial mechanisms		
sostenibles de largo plazo.	sustainable in the long term.		

7.

Source ES	Target EN	Target EN correction	Explanation
O C É A N O	O C E A N O		
PA C Í F I C O	P A C C I F I C O		
M A R C A R I B E	M A R C A R I B E		
C O L O M B I A	C O L L O M B I A		

8.

Source ES	Target EN	Target EN correction	Explanation
Llevamos más de 15 años comprometidos con la conservación de los océanos y el bienestar de las familias que viven del mar.	For more than 15 years we have been committed to the conservation of the oceans and the well-being of the families that live from the sea.		

9.

Source ES	Target EN	Target EN correction	Explanation
Piloto de medidas de reducción de riesgos de desastres basados en ecosistemas (Eco-RRD) para disminuir la vulnerabilidad de comunidades locales, protegiendo ecosistemas estratégicos.	Pilot of ecosystem-based disaster risk reduction measures (Eco-RRD) to reduce the vulnerability of local communities, protecting strategic ecosystems.		

10.

Source ES	Target EN	Target EN correction	Explanation
Gracias al generoso apoyo de:	Thanks to the generous support of:		
Blue Action Fund, KfW, el Fondo Mundial para los Arrecifes de Coral	Blue Action Fund, KfW, the Global Coral Reef Fund		

Answers

There is more than one possible answer to the exercises, but here are some examples of how the machine translation could be improved.

1.

Source ES	Target EN	Target EN correction	Explanation
3.000.000 Acuerdos	3.000.000 Agreements	3,000,000 Agreements	In English a comma is used to separate numbers rather than a decimal point.

2.

Source ES	Target EN	Target EN correction	Explanation
Ella recibió su título en lenguas modernas	She received her title in modern languages	She received her degree in modern languages	Título in this context would be better translated as degree.

3.

Source ES	Target EN	Target EN correction	Explanation
La responsable del equipo les permitió marcharse antes.	The team leader allowed them to leave early.	The team leader allowed them to leave early.	'Responsable' is feminine so the pronoun in the next segment should be 'she' rather than 'he'.
Se fue poco después.	He left shortly afterwards.	She left shortly afterwards.	

4.

Source ES	Target EN	Target EN correction	Explanation
La formación como profesional se fundamenta principalmente en tres áreas, articuladas a través de la formación aplicada y experiencial con sentido crítico y de compromiso social.	Training as a professional is based mainly on three areas, articulated through applied and experiential training with a critical and socially committed sense of social commitment.	The professional training course focuses on three main areas, which are put into practice through applied and experiential training with a critical and socially committed approach.	There is a repetition of 'socially committed / social commitment', which has been removed. Some other elements have also been edited to sound more natural in English.

5.

Source ES	Target EN	Target EN correction	Explanation
ecoturismo, pesca artesanal, piangua y economía circular.	ecotourism, artisanal fishing, piangua and circular economy.	ecotourism, artisanal fishing, shellfish picking, and circular economy.	Piangua has been left in the source language, so a translation has been added.

6.

Source ES	Target EN	Target EN correction	Explanation
Mecanismos financieros	Financial mechanisms	Long-term sustainable /j/ financial mechanisms.	Adjective placement is typically before the noun in English, so the segments have been joined and the adjectives put before the noun.
sostenibles de largo plazo.	sustainable in the long term.		

7.

Source ES	Target EN	Target EN correction	Explanation
O C É A N O	O C E A N O	PACIFIC OCEAN	The source text has spaces in, which has led to it not being translated. In the correction the text has been translated and 'pacific ocean' has been joined together to correct the adjective placement.
PA C Í F I C O	P A C C I F I C O		
M A R C A R I B E	M A R C A R I B E	CARIBBEAN SEA	
C O L O M B I A	C O L L O M B I A	COLOMBIA	

8.

Source ES	Target EN	Target EN correction	Explanation
Llevamos más de 15 años comprometidos con la conservación de los océanos y el bienestar de las familias que viven del mar.	For more than 15 years we have been committed to the conservation of the oceans and the well-being of the families that live from the sea.	For over 15 years, we have been committed to ocean conservation and the well-being of families who make a living from the sea.	Some stylistic changes have been made, such as changing 'live from the sea' to 'make a living from the sea'.

9.

Source ES	Target EN	Target EN correction	Explanation
Piloto de medidas de reducción de riesgos de desastres basados en ecosistemas (Eco-RRD) para disminuir la vulnerabilidad de comunidades locales.	Pilot of ecosystem-based disaster risk reduction measures (Eco-RRD) to reduce the vulnerability of local communities.	Pilot of ecosystem-based disaster risk reduction measures (Eco-DRR) to reduce the vulnerability of local communities.	The acronym has been changed from the Spanish acronym (RRD) to the English one (DRR).

10.

Source ES	Target EN	Target EN correction	Explanation
Gracias al generoso apoyo de:	Thanks to the generous support of:	Thanks to the generous support of:	The machine translation is a direct translation, but the organization's name in English is the Global Fund for Coral Reefs or Global Funds for Coral Reefs (GFCR).
Blue Action Fund, KfW, el Fondo Mundial para los Arrecifes de Coral	Blue Action Fund, KfW, the Global Coral Reef Fund	Blue Action Fund, KfW, the Global Fund for Coral Reefs	

Further reading

- *Traducción automática,* Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez
- *Practical Guide to Machine Translation*, Dimitar Shterionov
- Post-editing machine translation course, Trados