



Automated Post-Editing (APE) and Automated Quality Estimation (AQE)

Overview

Table of Contents

Automated Post-Editing (APE) and Automated Quality Estimation (AQE) 3

APE definition	3
Supported languages	3
Input	3
What APE does	3
Output	3
Key success factors	4
Examples of Client data	4

AQE Score Definition 5

QE Definition	5
Supported languages	6
Input	6
What it does	6
Output	6
Examples of AQE scores	6

Appendix 7

Supported languages	7
Languages by post-editing capability	8

APE definition

The Automated Post-Editing (APE) service uses Large Language Models (LLMs) to automatically refine and correct machine-translated text. It enhances MT output by addressing general linguistic issues as well as client-specific correction requirements, when customized.

The service is offered in two variants:

- Basic APE: focuses on general language corrections and improvements
- Customized APE: incorporates client-specific guidelines, terminology, and correction patterns

Supported languages

APE supports the 30 major world languages (see Appendix), but additional languages can be considered.

Input

APE now works integrated into translation workflows.

- Supports XLIFF files only
- Is applied to segments pre-translated with MT and tagged as MT segments
- The segmentation used is the one in the XLIFF file

What APE does

The Basic APE version analyzes the MT segment and attempts to fix common issues. It checks the source segment if necessary.

- Corrects clear errors in grammar, syntax, punctuation, and lexical choice
- Explicitly avoids unnecessary changes or rephrasing

The Customized APE solution performs changes based on client data sources:

- Glossary: corrects terms according to a provided glossary list
- DNT: corrects terms according to a provided DNT list
- Style guide instructions: modifies the text to conform guidelines provided
- Specific client corrections: modifies the text to apply client specific corrections
- More details and examples are provided in the section Examples of Client data

Output

- The XLIFF file with modified segments
- Target MT segments are overwritten with the APE modified text

- The original MT segment is stored as an alternative in the XLIFF file (only in Octave XTM workflow)

Key success factors

To ensure successful integration of the APE component into the translation workflow, the following aspects are essential.

Initial Onboarding

- Define APE outcome expectations (quality goals): Establish a common understanding of what the APE component is expected to improve and what aspects will remain unchanged.
- Collect client data to support APE expectations, including:
 - Glossaries (terminology)
 - Do-Not-Translate (DNT) lists
 - Style guides (a simplified selection suitable for AI use)
 - Client-specific corrections

Regular Updates

- Maintain and update client data, especially:
 - Glossaries
 - DNT lists

Examples of Client Data

Data type	Description and examples
Glossary	Terminology list exported from the Termbase. Preferably it should include domain specific words and compound words. It is advisable to avoid common general vocabulary words. E.g.: <i>Prefer:</i> allocation of assets allocation d'actifs <i>Avoid:</i> allocation allocation
DNT List	Proper names are included in the glossary as names that do not change and include proper names and acronyms. E.g.: American Red Cross American Red Cross NASA NASA
Style Guides	Style guides must be AI-oriented, they must be simple and clear. E.g.: • Ensure the use of formal tone in all contexts. • When degree symbols have both Celsius (°C) and Fahrenheit (°F) versions, use the Celsius first, followed by the Fahrenheit unit in parentheses. • Translate "&" as Portuguese "e" when surrounded by spaces and not being part of a Proper Name. • Maintain capitalization as in source English.

Data type	Description and examples
Client-specific corrections	These may include recurrent errors detected in the MT output or just client-specific output changes. E.g.: • In percentages, ensure non-breaking space between a number and the % character. • When a sentence in passive voice is used in the source text, turn it into active voice whenever possible.

AQE Score Definition

QE Definition

The Quality Estimation (QE) score is a numerical prediction of translation quality of a translation. The score is based on a Large Language Model assessment of grammatical accuracy, semantic precision, and overall readability.

QE helps measure the overall quality of machine translation (MT) or automated post-editing (APE) outputs without the need for a reference human generated translation.

Quality is evaluated at two different levels:

1. SEGMENT LEVEL: each individual segment receives its own score.
2. DOCUMENT LEVEL: a single weighted score summarizes the quality of the entire document.

Quality scores are customizable for each client and should help clients decide about the quality grades for each content type along with their content publishing strategy. The output quality score ranges from 0 to 100, and thresholds are defined on a client basis.

- Ready For Publishing: Very good quality. No errors found.
- Almost Ready for Publishing: Good quality. Errors relate to possible different linguistic choices.
- Light Review Required: Acceptable but needs review. Occasional grammatical or terminology errors.
- Needs Attention: Below target. Requires attention, errors may impact the understandability.
- Full Review Required: Poor. Requires review as errors may impact comprehension.

While QE scores are a fairly reliable benchmark, it's worth remembering that they're only an automated estimation.

Supported languages

The AQE component supports the same languages as the APE component.

Input

AQE now works integrated into translation workflows.

- Supports XLIFF files only
- Is applied to the same segments sent to the APE workflow
- The segmentation used is the one in the XLIFF file

What it does

The AQE component:

- Analyzes an input segment according to the provided scale of error types and provides a numerical value.
- The input segment may be a raw MT segment or an APE modified segment.

Output

- Quality Estimation score at segment level
- A quality score

Examples of AQE scores

100	Very good quality. No errors found.
95–99	Good quality. Errors relate to possible different linguistic choices.
90–94	Acceptable but needs review. Occasional grammatical or terminology errors.
80–89	Below target. Requires attention, errors may impact the understandability.
–80	Poor. Requires review as errors may impact comprehension.

	Source	MT	QE Score	Human translation
VERY GOOD QUALITY	Top Manufacturing & Delivery Performance	Máximo rendimiento de fabricación y entrega	100	
GOOD QUALITY Minor style or fluency issues.	Understanding how to leverage automation effectively is the first step to achieving your digital transformation goals.	Entender cómo hacer uso de la automatización de manera efectiva es el primer paso para lograr sus objetivos de transformación digital.	97	Comprender cómo aprovechar la automatización de forma eficaz es el primer paso para alcanzar sus objetivos de transformación digital.

	Source	MT	QE Score	Human translation
ACCEPTABLE Occasional grammar or terminology errors.	Why do you need permanent downhole monitoring?	¿Por qué necesita una monitorización permanente del downhole ?	94	¿Por qué necesita una monitorización permanente de fondos de pozo ?
BELOW TARGET Requires attention.	Martin is a registered Professional Engineer with ABLATA, and holds a B.Sc. in Electrical Engineering from the University of Calgary as well as a Project Management Certificate from Mount Royal University.	Martin es Ingeniero Profesional registrado en ABLATA, y cuenta con un B.Sc en Ingeniería Eléctrica de la Universidad de Fluoroy y un Certificado de Gestión de Proyectos de la Universidad Monte Real .	70	Martin es ingeniero profesional registrado por ABLATA y posee una licenciatura en Ingeniería Eléctrica por la Universidad de Calgary , así como un certificado en Gestión de Proyectos por la Universidad Mount Royal .
POOR Requires review, as errors may impact comprehension.	Matt West	Matt Oeste	50	Matt West

Appendix

Supported languages

APE supports the 30 major world languages (but additional languages can be considered).

- | | |
|-----------------------|----------------|
| 1. English | 16. Polish |
| 2. Spanish | 17. Turkish |
| 3. French | 18. Vietnamese |
| 4. German | 19. Ukrainian |
| 5. Chinese (Mandarin) | 20. Hebrew |
| 6. Portuguese | 21. Czech |
| 7. Japanese | 22. Finnish |
| 8. Korean | 23. Romanian |
| 9. Italian | 24. Hungarian |
| 10. Russian | 25. Indonesian |
| 11. Arabic | 26. Malay |
| 12. Dutch | 27. Thai |
| 13. Swedish | 28. Bulgarian |
| 14. Norwegian | 29. Greek |
| 15. Danish | 30. Slovak |

Languages by Post-editing capability

Group 1	Group 2	Group 3	Group 4
Excellent fluency, style control, and error-correction. Safe for demanding post-editing tasks.	Very good quality; still strong for professional post-editing in most domains.	Good everyday quality; sometimes a bit less consistent with rare vocabulary, very formal style, or low-resource domains.	Usable for drafts and light post-editing, but quality can jump around, especially for idioms and low-frequency terms.
1. English 2. Spanish 3. French 4. German 5. Mandarin 6. Portuguese (PT+BR) 7. Japanese 8. Korean 9. Italian 10. Russian	11. Arabic 12. Dutch 13. Swedish 14. Norwegian 15. Danish 16. Polish 17. Turkish 18. Vietnamese 19. Ukrainian 20. Hebrew 21. Czech 22. Finnish	23. Romanian 24. Hungarian 25. Indonesian 26. Malay 27. Thai 28. Bulgarian 29. Greek 30. Slovak 31. Croatian 32. Serbian 33. Catalan 34. Cantonese 35. Farsi 36. Hindi	37. Bengali 38. Urdu 39. Tamil 40. Telugu 41. Marathi 42. Gujarati 43. Malayalam 44. Swahili 45. Tagalog 46. Lithuanian 47. Latvian 48. Estonian 49. Slovenian 50. Basque