



Mind your terms!

Terminology: headache for readers and post-editors

LESZNYÁK Ágnes, PhD

*European Commission,
Directorate General for Translation*

TSS2020 online, 2 July 2020

Machine Translation ≠ Machine Translation

Quality of output depends, among others, on

- type of technology used (rule-based, statistical, neural)
- type of texts used for training the engine (general or special language)
- language pairs

- machine translation service provided by the European Commission
- integrated into digital services + stand-alone services
- neural technology



Free of charge to

- employees of public administrations,
- university staff and students, and
- small & medium-sized enterprises

} in the EU

User guide to the SME Definition

Can I use it?



enterprise = ‘any entity engaged in an economic activity, irrespective of its legal form’



“the self-employed, family firms, partnerships and associations or any other entity that is regularly engaged in an economic activity may be considered as enterprises”



economic activity = ‘the sale of products or services at a given price, on a given/direct market’

User guide to the SME Definition

Can I use it?



“The category of micro, small and medium-sized enterprises consists of enterprises which:

- employ fewer than 250 persons; and
- have either an annual turnover not exceeding EUR 50 million or an annual balance sheet total not exceeding EUR 43 million.”



[User guide to the SME Definition](#)





- controlled access
- access point hosted in the European Commission Data Centre
- translation engines run on cloud servers based in Europe
- source document deleted after translation
- makes no claim on the IPR of documents submitted to it or to the output of its system

Welcome to the registration page for eTranslation , the European Commission's safe and secure machine translation system.

To use eTranslation, you need an EU Login ID.

You need to register!

Create an EU Login account and an eTranslation profile

<https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eTranslation+for+SMEs#eTranslationforSMEs-first>



Connecting Europe eTranslation

Languages covered



eTranslation now offers Chinese! Try it under the “General Text” domain (“Advanced options”).

eTranslation supporte maintenant le Chinois ! Testez-le en sélectionnant le domaine «General Text» (« Options avancées »).

eTranslation bietet jetzt auch Chinesisch an! Probieren Sie es aus unter „General Text“ („Erweiterte Optionen“).

How can I use it?

- translate text - type or copy and paste your snippet of text
- translate documents - upload documents individually or in batch
 - specific output formats for computer-aided translation: tmx and xliff
 - up to 5 documents into several languages in one go
 - size limitation for documents: 10 MB or 300 000 sentences
 - translation stored in personal workspace for 24 hours, or sent by e-mail on request

What is it for?

- trained using translations produced by the translators of the EU institutions over the past decades
 - over 1 billion sentences in the 24 official EU languages
- ➔ produces best results translating EU policy documents

Quality warning

„The output provided by the service is raw machine translation intended to facilitate understanding of foreign language texts. The European Commission makes no claims as to its accuracy or stylistic quality and accepts no liability for possible errors. Nor does it make, implicit or explicit, guarantees regarding the availability, uptime and usefulness of the eTranslation service. Any texts intended for publication or where accuracy is paramount should be edited by human translators. Documents produced by this system contain no indication that they were machine translated.”

Quality issues

- unpredictable quality
 - quality varies widely from one segment to another
- typical errors

Typical errors

- deceptive fluency
- mistranslation due to literal translation (elliptical sentences, non-literal meaning)
- omissions, insertions, repetition
- grammar
- spelling

Typical errors

- incorrect or inconsistent terminology
- context-independent translation of vocabulary
- translation of proper names into fictional words

'Ars Experandi' → 'Ars Excavandi'

Why is terminology a problem for NMT?

- no efficient technical solution available yet to
 - enforce the use of custom terminology from glossaries and
 - consider context beyond the segment level

A way forward for eTranslation

- adapt to text types that are typically used in specific contexts
 - "EU Formal Language" – a formal, legalistic style
 - "General Text" – more colloquial
 - specific domains – specific vocabulary and style geared towards the given domain or policy area

domains available will depend on the language pair chosen

EU Formal Language

General Text

Legacy MT@EC

Court of Justice Case Law

Cultural

Deutsche Bundesbank

Deutsche Bundesbank neural test

IP Case Law

Ministère des Finances (France)

Public Health

Valtioneuvoston Kanslia

EU Formal Language



Let us look at an example
together!

- Please read the pdf file *Case T-669_19*. This English text is the source text for the translations we will work with.
- Highlight all the occurrences of the following terms:
 - *applicant*
 - *European Union Intellectual Property Office*
 - *trade mark at issue*
 - *infringement*
 - *requirement of evidence*
- Please read *Rechtssache T-669-19 EU formal / Affaire T-669 EU formal* and find the equivalents of the listed terms in the translated text.

We are splitting up into breakout rooms in a minute...

In the breakout rooms

- Discuss whether the highlighted terms are translated consistently and correctly.
- To decide about correctness, please refer to the official translations *Rechtssache T-669_19 / Affaire T-669_19* (pdf files)
- In 10 minutes we will all meet to draw some conclusions.
- (If you are very quick and have some time left, check out the files with 'case law' and 'general text ' in their names.)

Different engines, different terms

ORI EN	FR OJ	FR EU	FR CL	FR GENERAL
applicant	Partie requérante	Demandeur Partie requérante La requérante	Requérante Partie requérante	Demandeur Requérant/La requérante
European Union Intellectual Property Office	Office de l'Union européenne pour la propriété intellectuelle		L'Office de l'Union européenne pour la propriété intellectuelle	Office de la propriété intellectuelle de l'Union européenne
trade mark at issue	Marque litigieuse concernée		Marque litigieuse	Marque en cause
infringement	violation		violation la violation	infraction violation
requirement of evidence	l'exigence de preuve	l'exigence de la preuve	l'exigence des preuves	

Different engines, different terms

ORI EN	DE OJ	DE EU	DE CL	DE GENERAL
applicant	Klägerin	Antragsteller Kläger vor dem Gericht Klägerin		Antragsteller Kläger vor dem Gericht Klägerin
European Union Intellectual Property Office	Amt der Europäischen Union für geistiges Eigentum			Amt für geistiges Eigentum der Europäischer Union
trade mark at issue	streitige marke			
infringement	Verstoß	Verletzung Verstoß	Verletzung Verstoß	Verletzung Verstoß
requirement of evidence	Beweiserfordernis	Erfordernis von Beweismitteln		Beweispflicht

Conclusions

- use different engines for different purposes
- as a reader
 - be aware of these weak points and develop a certain level of distrust
- as a translator
 - use termbases (glossaries) in your CAT tool (e.g. [IATE](#))
 - filter for terms, replace terms (with caution)
 - use automatic term checks

Thank you for your attention!



© European Union 2020

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.

