

# APE-Quest

or how to be picky about machine  
translation

Sara Szoc  
Tom Vanallemeersch

## Translation workflow involving Machine Translation (MT)



eTranslation  
Machine Translation



Human translator



*MT outputs are manually checked and edited as needed*



APE-Quest

English source

French Translation

### MT can be impressively fluent

This privacy statement explains the reason for the processing of your personal data, the way we collect, handle and ensure protection of all personal data provided ...

La présente déclaration de confidentialité explique les raisons du traitement de vos données à caractère personnel, la façon dont toutes les données à caractère personnel fournies sont recueillies, traitées et protégées ...

### Privacy Statement (in-domain text)

### ... but failures can lead to frustrations

XXXX XXXX -- -- I am not aware **of this debt I would like this debt** validated or removed from my credit report.

XXXX XXXX — -- Je n'ai pas connaissance **de cette dette que je souhaiterais que cette dette** soit validée ou retirée de mon rapport de crédit.

### Consumer Complaints (out-domain text)





CROSSLANG  
TRANSLATION AUTOMATION



## Strategy 1 - MT Customization

### Generic data

We have read the new book.

Nous avons lu le nouveau livre.

They have described what happened there.

Ils ont décrit ce qui s'est passé là.

...

### Domain-specific data

I monitor my credit report more frequently now as we're attempting to buy our first house.

...

Je surveille mon rapport de crédit plus souvent maintenant car nous essayons d'acheter notre première maison.

...

→ Train new MT system



CROSSLANG  
TRANSLATION AUTOMATION

Introduction

## Strategy 2 - Automatic Quality Estimation (QE)

**Source:**

I am not aware **of this debt** I would like this debt validated or removed from my credit report.



MT

**Translation:**

Je n'ai pas connaissance **de cette dette que je souhaiterais que cette dette** soit validée ou retirée de mon rapport de crédit.

**Prediction of quality score:**

0.75

QE



**CROSSLANG**  
TRANSLATION AUTOMATION

Introduction

## Strategy 3 - Automatic Post-Editing (APE)

**Translation:** Je n'ai pas connaissance **de cette dette que je souhaiterais** **que cette dette** soit validée ou retirée de mon rapport de crédit.

0.75 ← QE

**Correction:** Je n'ai pas connaissance **de cette dette, je souhaiterais** **que cette dette** soit validée ou retirée de mon rapport de crédit.

APE



# Project Details

Connecting Europe Facility 2014-2020  
CEF TELECOM  
CEF-TC-2017-3, 2017-EU-IA-0151

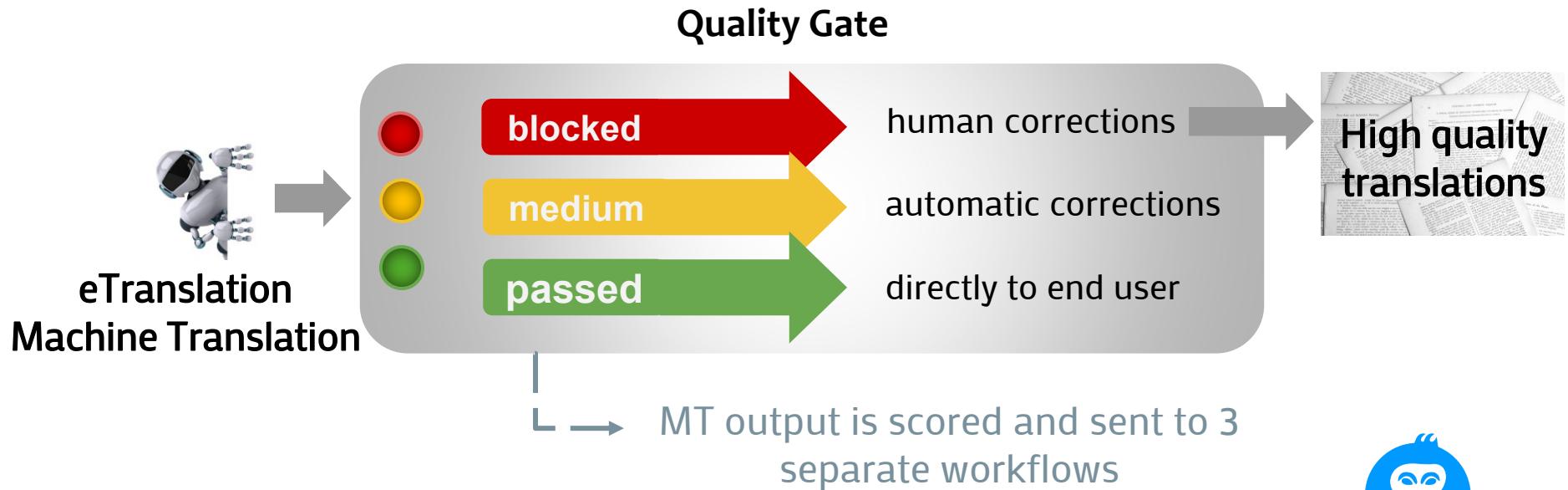
2018-2020



Automated Postediting and Quality Estimation

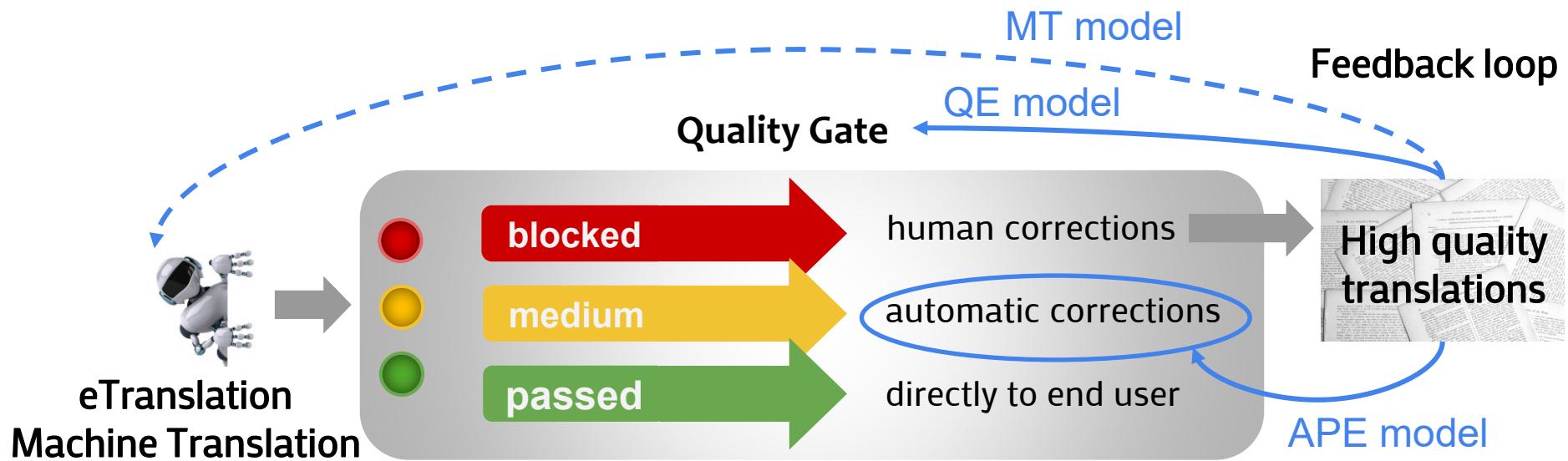


## APE-Quest: integrating MT, QE, and APE



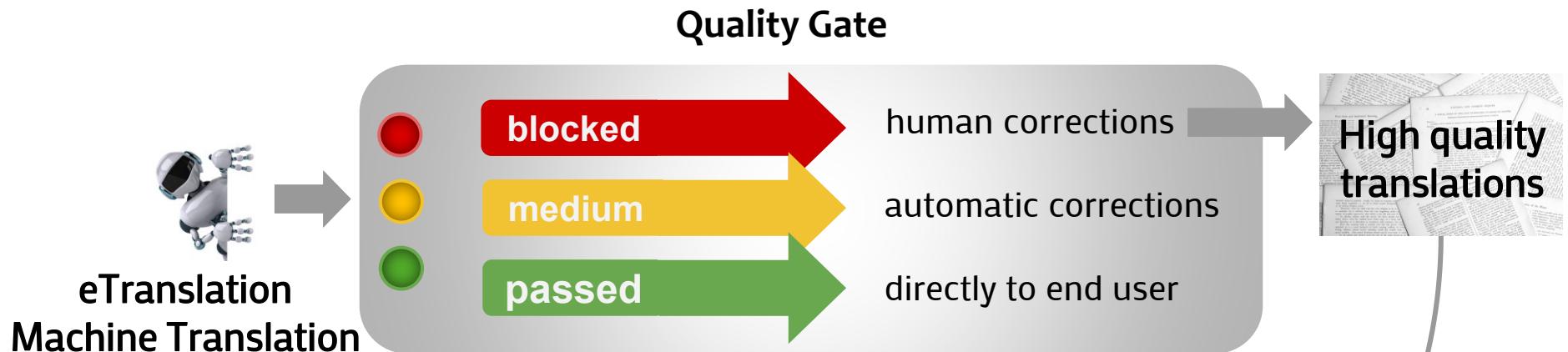
**APE-Quest**

## Re-using corrections to improve core components



APE-Quest

## Making translations publicly available (e.g. research)



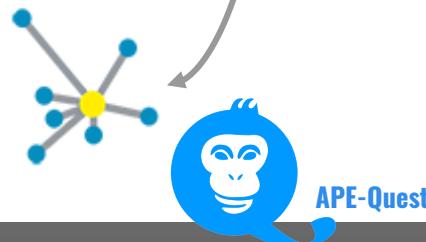
Post-editing corpus English to Dutch/French/Portuguese, legal domain ⓘ

APE-QUEST posteditition tuples

10k sentences per language pair

<https://ape-quest.eu>

European Language  
Resource Coordination  
Connecting Europe Facility  
ELRC-SHARE

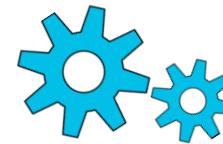


## The MT component

eTranslation MT system / other online system / in-house system  
→ based on deep learning

Experiments include English > Dutch, French, Portuguese

Generic data	
We have read the new book.	Nous avons lu le nouveau livre.
They have been monitoring their credit report more frequently now as we're attempting to buy our first house.	Il sont en train de surveiller leur rapport de crédit plus souvent maintenant car nous essayons d'acheter notre première maison.
Domain-specific data	
I monitor my credit report, more frequently now as we're attempting to buy our first house.	
...	



## The QE component

Translation quality estimation at the sentence level

Prediction of quality score:  $1 - \text{HTER}$  score  
→ Human Translation Edit Rate

1 = Everything was changed



0 = Nothing was changed



**OpenKiwi**  
By Unbabel

[github.com/Unbabel/OpenKiwi](https://github.com/Unbabel/OpenKiwi)

**TransQuest**

[github.com/TharinduDR/TransQuest](https://github.com/TharinduDR/TransQuest)



**CROSSLANG**  
TRANSLATION AUTOMATION



APE-Quest Components

## The QE component

Example used for training the component:

HTER = 1 / 24:  
1 substitution  
Post-edited sentence = 24 words

Input sentence: I am not aware **of this debt** I would like **this debt** ...

MT output: Je n' ai pas connaissance **de cette dette que** je  
**souhaiterais que cette dette** ...

Human post-edition: Je n' ai pas connaissance **de cette dette ,** je  
**souhaiterais que cette dette** ...



APE-Quest

## The APE component

### Neural MT is hard to correct automatically



- Overcorrecting APE systems

**Training source:** Je n'ai pas connaissance de cette dette que je souhaiterais ...

**Training target:** Je n'ai pas connaissance de cette dette , je souhaiterais que ...



## The APE component

### Neural MT is hard to correct automatically



Cautious system: CopyCAT Networks (Ive et al., 2019)  
[github.com/ImperialNLP/CopyCat](https://github.com/ImperialNLP/CopyCat)

**Training source:** I am not aware of this debt I would like this debt ... +

Je n'ai pas connaissance de cette dette que je souhaiterais que ...

**Training target:** Je n'ai pas connaissance de cette dette , je souhaiterais que ...



# PoC Implementation of the Quality Gate

CROSSLANG  
TRANSLATION AUTOMATION

▼ urgent\_t...t.docx (15) > en-US > fr-FR

PREVIEW ▾ ✓ 🔍 🔎 T

193 COMMUNICATION STUDIES

194 Which role did the tabloid press play in "Brexit"? Comment la crise des réfugiés est-elle décrite dans les médias ?

195 How is the refugee crisis portrayed in the media?

What is the business model of Netflix? Quel est le modèle d'affaires de Netflix ?

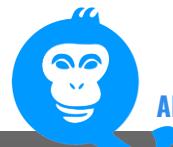
What is the business model of Netflix? Quel est le modèle d'affaires de Netflix ?

Translation Matches (No MT) (3) TM Search Glossary

What is the business model of Netflix?

Source: QE existing translation 2019-11-21 55.55555

MT output /  
Human post-edition /  
APE output



APE-Quest

## Project Evaluations

Examine the applicability and benefits of introducing a Quality Gate into the MT workflow

### Source (English)

Privacy statement if I'm a trader

PROTECTION OF YOUR PERSONAL DATA

1

### Target (Dutch)

Privacyverklaring als ik ondernemer bent

### Translation Quality



Publishable



Not publishable

Use Case 1 (Assimilation)	Use Case 2 (Dissemination)
Consumer Complaints (informal)	Privacy Statement (formal)
Content needs to be understood	Content needs to be published
Is the translation Acceptable / Not Acceptable ?	Is the translation Publishable / Not Publishable ?



APE-Quest



CROSSLANG  
TRANSLATION AUTOMATION

APE-Quest Evaluations



## Quality predictions English > Dutch

Evaluate the performance of the Quality Gate workflow based on different values of QE score thresholds

Main Results:

Threshold	Assimilation		Dissemination	
	Time (%)	Quality (%)	Time (%)	Quality (%)
Traditional	100.00	97.67	100.00	97.89
QE < 0.90	64.18	83.87	33.02	94.74
QE < 0.80	8.80	59.16	0.75	90.53
MT-Only	0.0	54.07	0.0	90.53

- High quality gains compared to MT only workflow
- Cost / time savings compared to traditional workflow
- Less impact for dissemination due to high quality in-domain MT output



## Quality predictions English > Dutch

## Oracle scores

Threshold	Assimilation		Dissemination	
	Time (%)	Quality (%)	Time (%)	Quality (%)
Traditional	100.00	97.67	100.00	97.89
QE < 0.90	64.18	83.87	33.02	94.74
QE < 0.80	8.80	59.16	0.75	90.53
MT-Only	0.0	54.07	0.0	90.53

	Assimilation		Dissemination	
	Time (%)	Quality (%)	Time (%)	Quality (%)
	100.00	97.67	100.00	97.89
	90.04	97.09	62.00	96.84
	80.44	94.04	33.90	91.58
	0.0	54.07	0.0	90.53

Oracle scores show that there is still much room for improving QE predictions



APE-Quest

# Conclusions

- ❑ Neural MT is able to produce high quality output, but can fail dramatically on out-of-domain content
- ❑ Quality Gate of APE-Quest (PoC implementation) routes translations to different workflows based on the estimated translation quality (QE score)
- ❑ Evaluation results evidence the potential of the Quality Gate for reducing cost/time while preserving high translation quality
- ❑ However, the application of APE to neural MT remains difficult

# Future work

- Post-editions and evaluations for additional domains
- Elaborate other APE strategies, for instance automatic recognition and modification of incorrect terminology in MT output





@ape\_quest

[www.ape-quest.eu](http://www.ape-quest.eu)

Van den Bogaert, J., Depraetere, H., Szoc, S., Vanallemeersch, T., Van Winckel, K., Everaert, F., Specia, L., Ive, J., Khalilov, M., Maroti, C., Farah, E., & Ventura A. (2019). *APE-QUEST: an MT Quality Gate*. In Proceedings of Machine Translation Summit XVII Volume 2: Translator, Project and User Tracks (pp. 110-111).

Van den Bogaert, J., Depraetere, H., Szoc, S., Vanallemeersch, T., Specia, L., Ive, J., & Khalilov, M. (2020). *APE-QUEST: a Quality Gate for Routing MT*. In Proceedings of EAMT 2020 (pp. 473-474).

Szoc, Sara, Heidi Depraetere (2020). *Quality Estimation*, In: Jörg Porsiel (Hg., 2020): Maschinelle Übersetzung für Übersetzungsprofis. Berlin: BDÜ Fachverlag. Sammelband mit Beiträgen in deutscher und englischer Sprache. 384 Seiten, ISBN: 9783946702092.

Ive, J., Specia, L., Szoc, S., Vanallemeersch, T., Van den Bogaert, J., Farah, E., Maroti, C., Ventura, A. & Khalilov, M. (2020) *A Post-Editing Dataset in the Legal Domain: Do we Underestimate Neural Machine Translation Quality?* In Proceedings of LREC 2020.

# For more info ...