Quality prediction

Use good machine translations like 100% translation memory matches



Adam Bittlingmayer

ModelFront



Artur Aleksanyan
CTO
ModelFront



Boris Zubarev

ML LEAD

ModelFront

MISSION

Make quality translation faster and cheaper

HOW TO MAKE QUALITY TRANSLATION FASTER AND CHEAPER

Beyond better machine translation

POST-EDITING

You send 100% to humans, humans edit __% of segments.

POST-EDITING

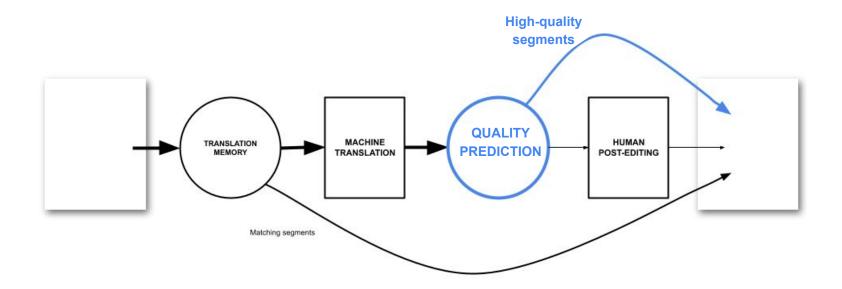
Humans don't edit ___% of segments.

POST-EDITING

Humans don't edit ___% of segments.

Can we predict which?

HOW IT WORKS



machinetranslate.org/hybrid-translation

"QUALITY ESTIMATION"

No human reference translation

source segment, target segment \rightarrow score

POST-EDITING WORKFLOW

RAW MACHINE TRANSLATION

POST-EDITING WORKFLOW

RAW MACHINE TRANSLATION

Same quality, faster and cheaper

POST-EDITING WORKFLOW

RAW MACHINE TRANSLATION

Same quality,

Better quality

faster and cheaper

(no "catastrophes")

POST-EDITING WORKFLOW

RAW MACHINE TRANSLATION

Same quality,

Better quality

faster and cheaper

(no "catastrophes")

90%

42%

QUALITY **THRESHOLD**

SKIP **HUMANS**

POST-EDITING WORKFLOW

RAW MACHINE TRANSLATION

Same quality,

Better quality

faster and cheaper

(no "catastrophes")

90% QUALITY **42**%

QUALITY SKIP THRESHOLD HUMANS

1%
QUALITY
THRESHOLD

99% SKIP HUMANS

Rol?

ROI

Skip humans for most of the good machine translations

Adoption

EVOLUTION

1. In-house

2. API providers

3. TMS integrations

ADOPTION

Customer support

Technical documentation

Product titles and descriptions

PROVIDERS

	ENGINES	CUSTOMIZATION	ACCESS
KantantQES	only KantanMT	✓	only KantanStream
Omniscien CS/QE	only Omniscien	✓	✓ API
Memsource QE	all		only Memsource
ModelFront	all	✓	✓ API, TMS integrations

TMS INTEGRATION

	UNILATERAL	OFFICIAL	
Memsource	•	•	Memsource QE
KantanStream		•	KantanQES
Translate5	•	•	ModelFront
Crowdin	•	•	ModelFront
XTM	•		
Lokalise	•		
SDL Worldserver	✓		
SDL TMS			
MemoQ			

Questions?