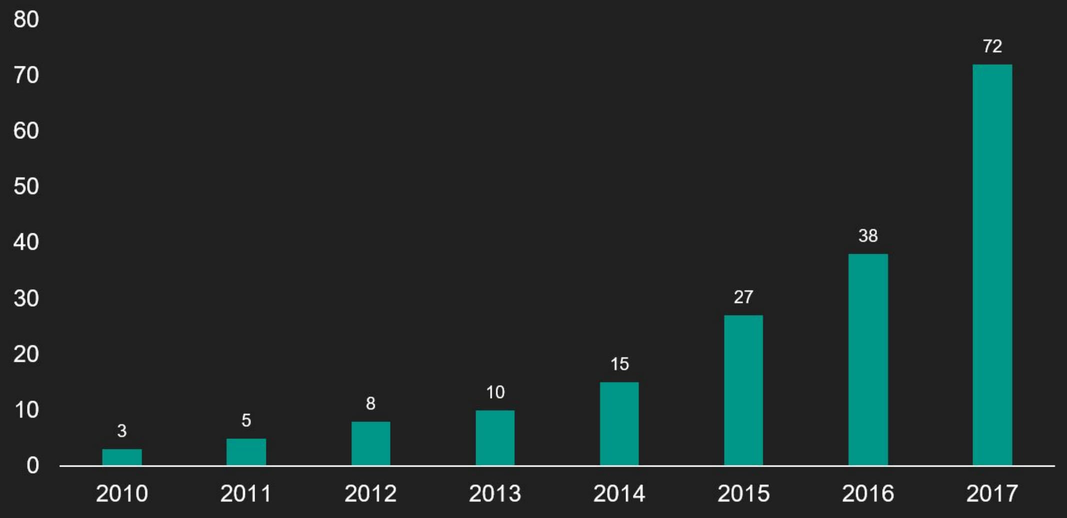
Towards Less Post-editing

Bill Lafferty, Memsource

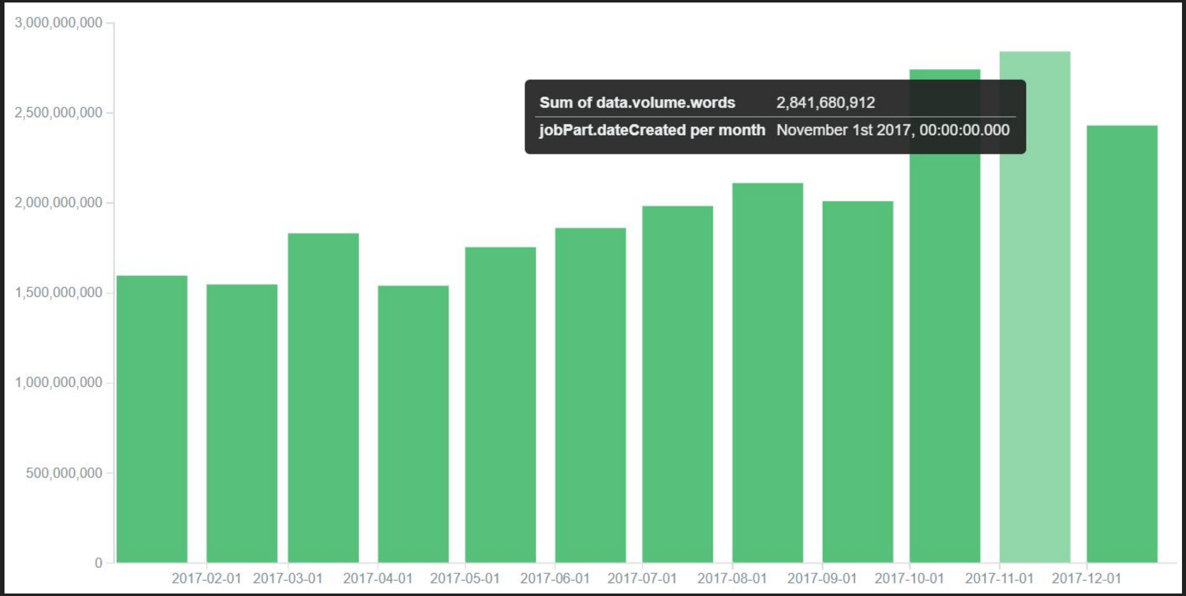
Memsource

- Founded in 2010
- Memsource helps global companies translate and manage translations
- Bootstrapped and profitable since 2013
- Based in Prague HQ

2017: A Growing Team



24 Billion Words in 2017

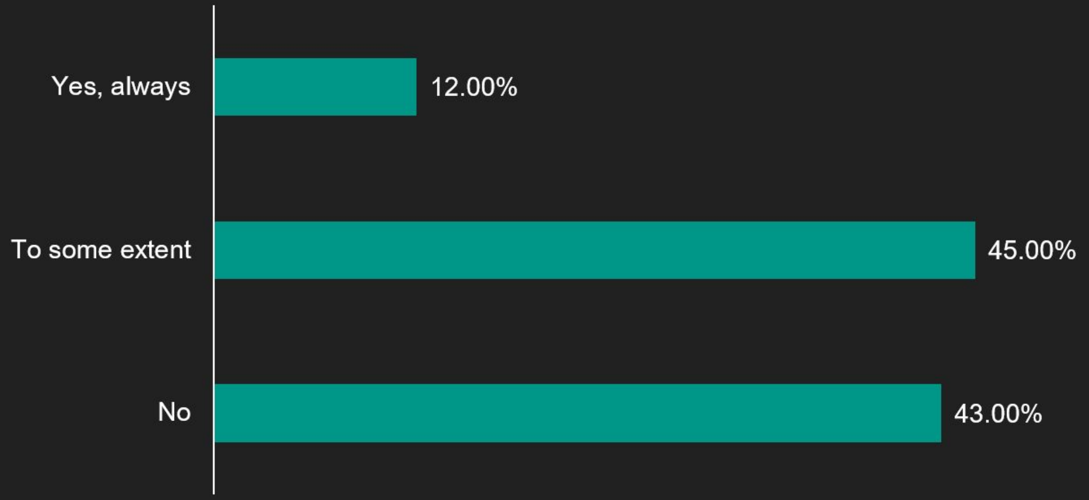


2017 - New AI Team

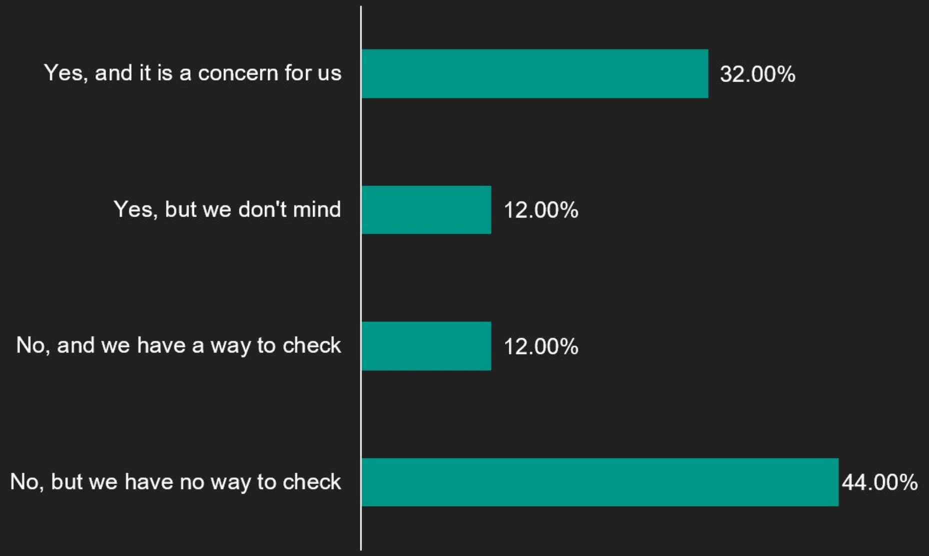
Back to 2012

- David Canek, Memsource CEO and founder presented at 2012 AMTA conference in San Diego
- Anyone attended? Raise hands
- At that time MT was picking up in the industry and David presented results from a survey Memsource and the GALA association ran among translation providers/LSPs

When using MT, are you able to measure its quality?



Do your translators use MT without you knowing about it?



2011: The Introduction of The Post-editing Analysis

- MT integration with CAT tools began around 2010
- Limited features supporting MT post-editing, e.g., to measure PE efficiency
- In 2011 Memsource launched the post-editing analysis

Post-editing Analysis

Introduction

In This Article

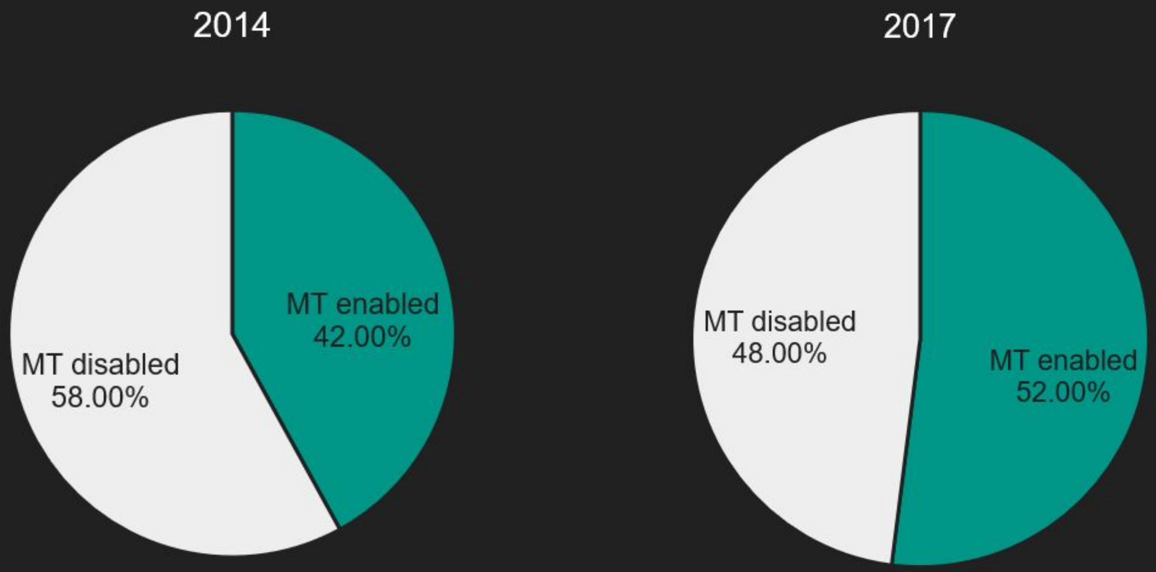
The post-editing analysis in Memsource Cloud extends the traditional translation memory analysis to also include machine translation and non-translatables (NT). It analyzes the MT and NT post-editing effort for each segment and compares the MT and NT output with the final post-edited translation (edit distance). Therefore, if the MT or NT output was accepted without further editing (the linguist did not need to change it at all), it would come up as a 100% match in the analysis.

If, on the other hand, the linguist changes the MT or NT output heavily, the match rate will be close to 0%. The score counting algorithm is identical to the one that we use to calculate the score of translation memory fuzzy matches. The only difference is that the post-editing analysis is based on the target. Therefore, the post-editing analysis must be, quite naturally, launched after the post-editing job has been completed.

A sample post-editing analysis:

	Segments				Pages				Words				Characters				%			
	TM	MT	NT	All	TM	MT	NT	All	TM	MT	NT	All	TM	MT	NT	All	TM	MT	NT	All
Net Rate	**4**	**3**	**1**	**8**	**0.21**	**0.07**	**0.01**	**0.29**	**65**	**18**	**1**	**84**	**325**	**108**	**21**	**454**				
All	8	9	3	20	0.38	0.23	0.04	0.64	113	60	3	176	580	359	64	1003	64.2	34.1	1.7	100
Repetitions	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0
101%	1	-	-	1	0.02	-	-	0.02	5	-	-	5	24	-	-	24	2.8	-	-	2.8
100%	2	9	0	11	0.02	0.23	0	0.25	6	60	0	66	33	359	0	392	3.4	34.1	0	37.5
95%–99%	1	0	3	4	0.17	0	0.04	0.21	47	0	3	50	263	0	64	327	26.7	0	1.7	28.4
85%–94%	1	0	0	1	0.05	0	0	0.05	18	0	0	18	80	0	0	80	10.2	0	0	10.2
75%–84%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50%–74%	2	0	0	2	0.11	0	0	0.11	36	0	0	36	160	0	0	160	20.5	0	0	20.5
0%–49%	1	0	0	1	0.01	0	0	0.01	1	0	0	1	20	0	0	20	0.6	0	0	0.6

MT Usage Started Picking Up

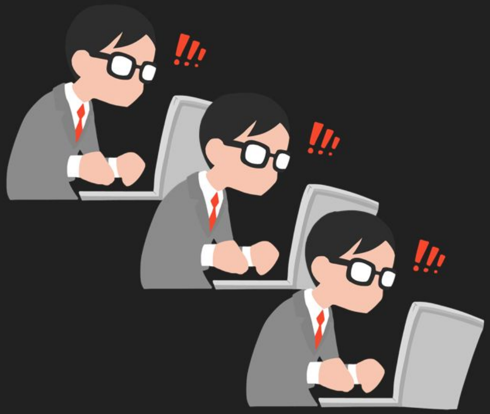


Post-editing Became the Norm

In 50%+ translation projects with machine translation enabled, MT post-editing became the preferred method of professional translation.

From TEP...

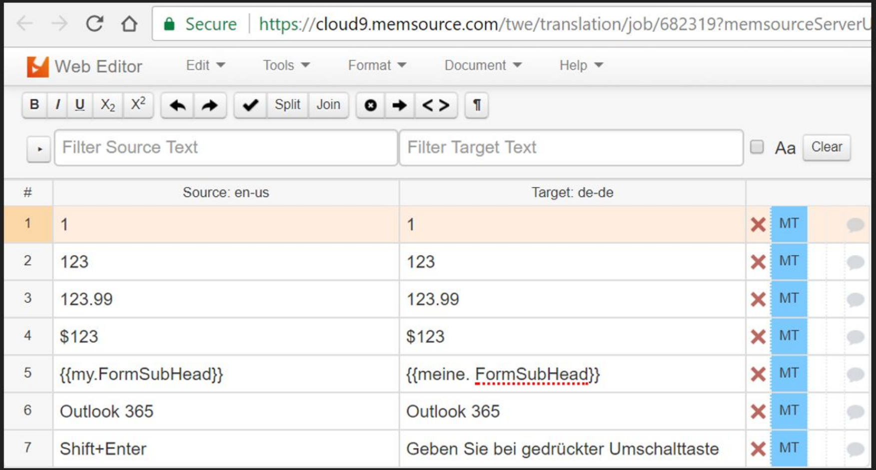
- Translation
- Editing
- Proofreading



To Post-editing MT

- Machine translation
- Human post-editing

Translator Post-editing MT



Less Post-editing

- Machine translation
- Automated quality estimation
- Human post-editing

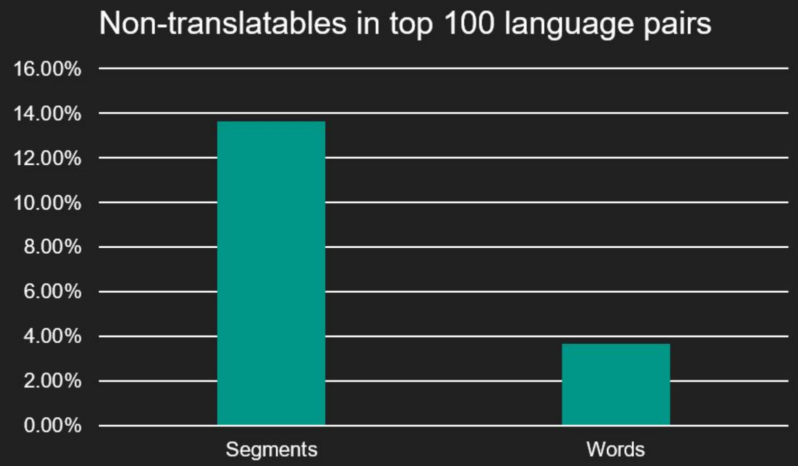
Automated Quality Estimation

- A very hard problem.
- We decided to first apply our approach to segments that DO NOT HAVE TO BE TRANSLATED: non-translatables

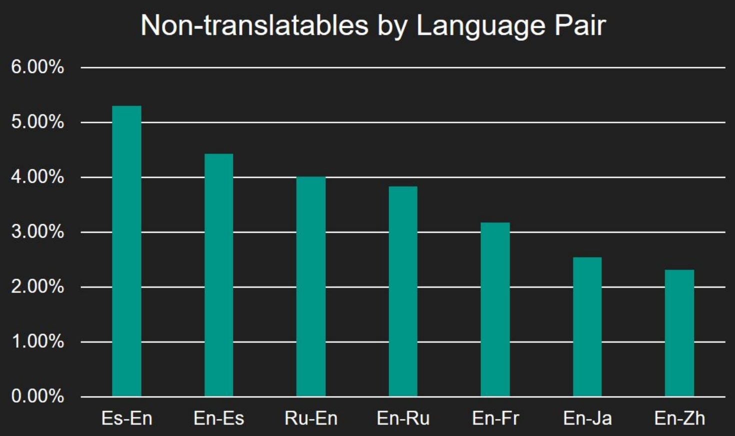
Examples of Non-translatables

- A segment that is simply copied from source to target.
- Some Examples:
  - *123*
  - *Labs.Core.Actions.IGetValueOptions*
  - *Memsource*
  - *Eva Smith*
  - *51 x 55 mm*

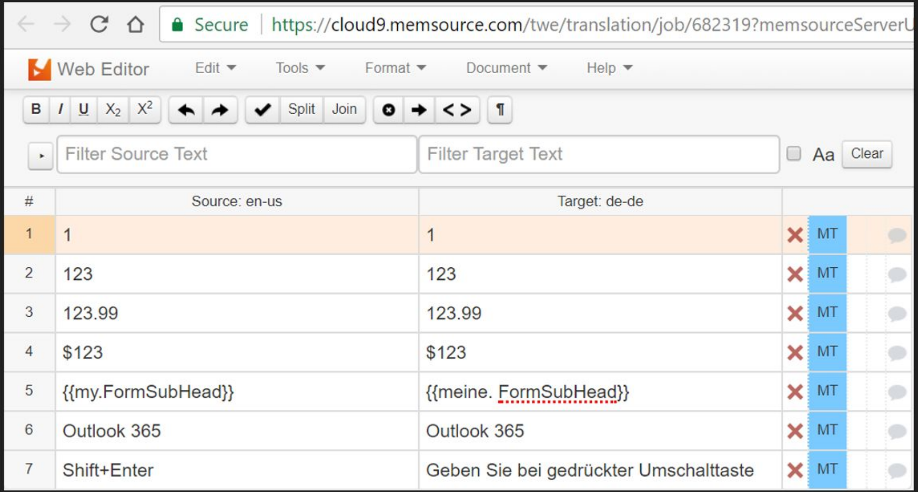
14% of Segments Are Non-translatable



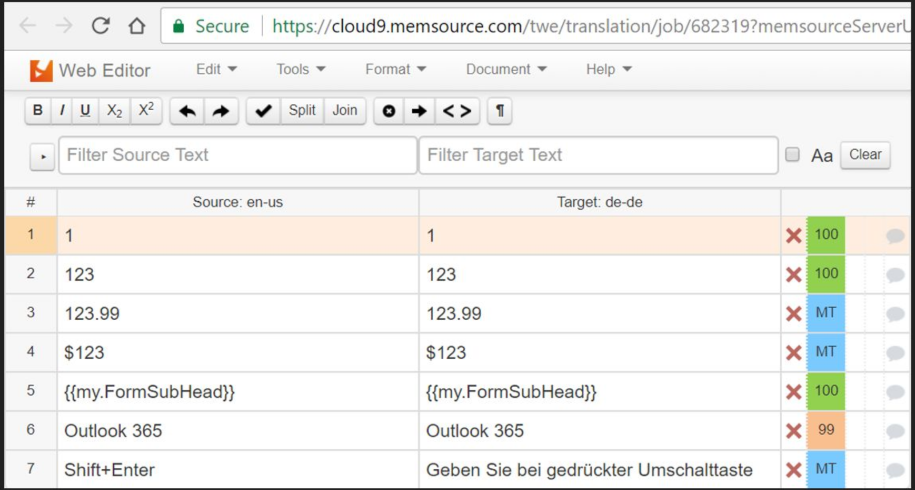
Differences between Language Pairs



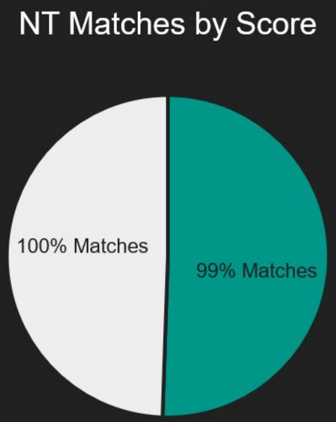
Traditional MT Post-editing



AI-powered MT Post-editing



Score Distribution



Accuracy

Overall accuracy very high:

- ~98% of accuracy of non-translatable segment identification
- ~95% of accuracy of 100% non-translatable match category