

The future of translation quality: Navigating AI, AutoMQM, and new translation methods

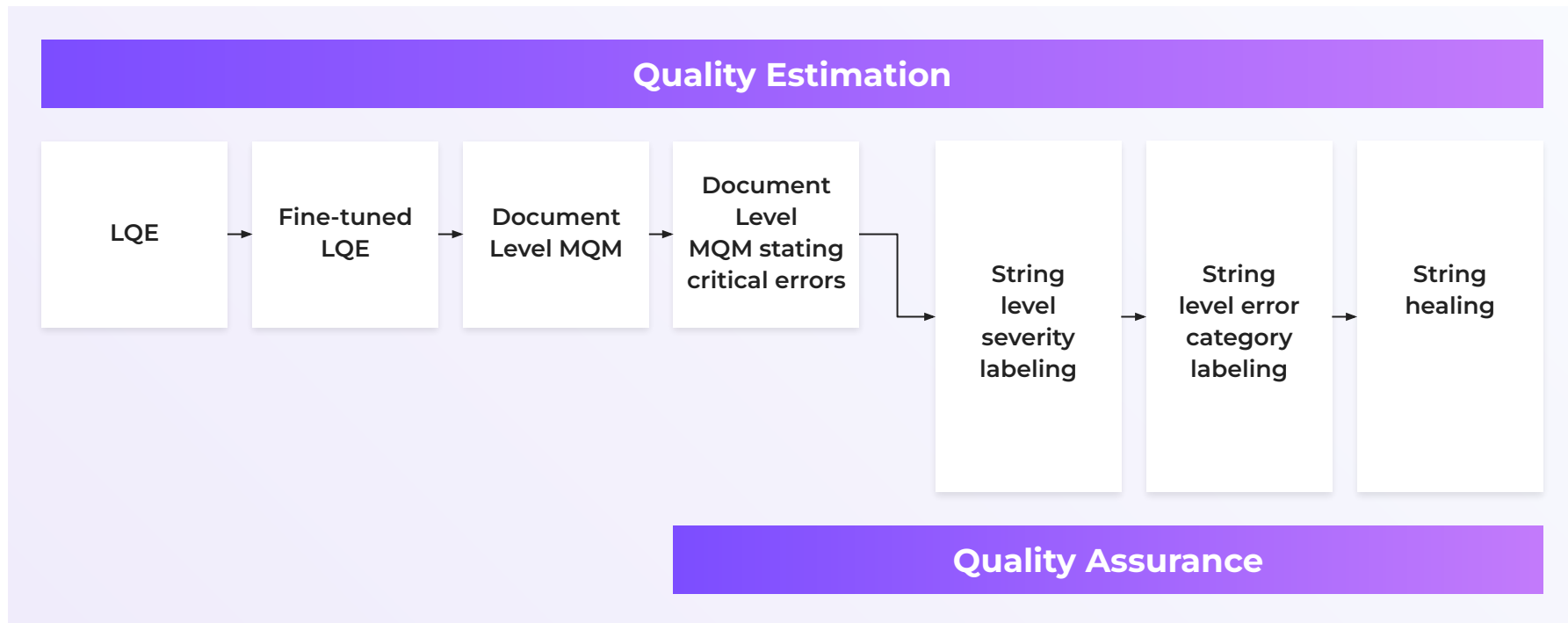


**Marina
Sánchez Torron**
Linguistic Engineer,
Smartling



Olga Beregovaya
VP, AI, Smartling

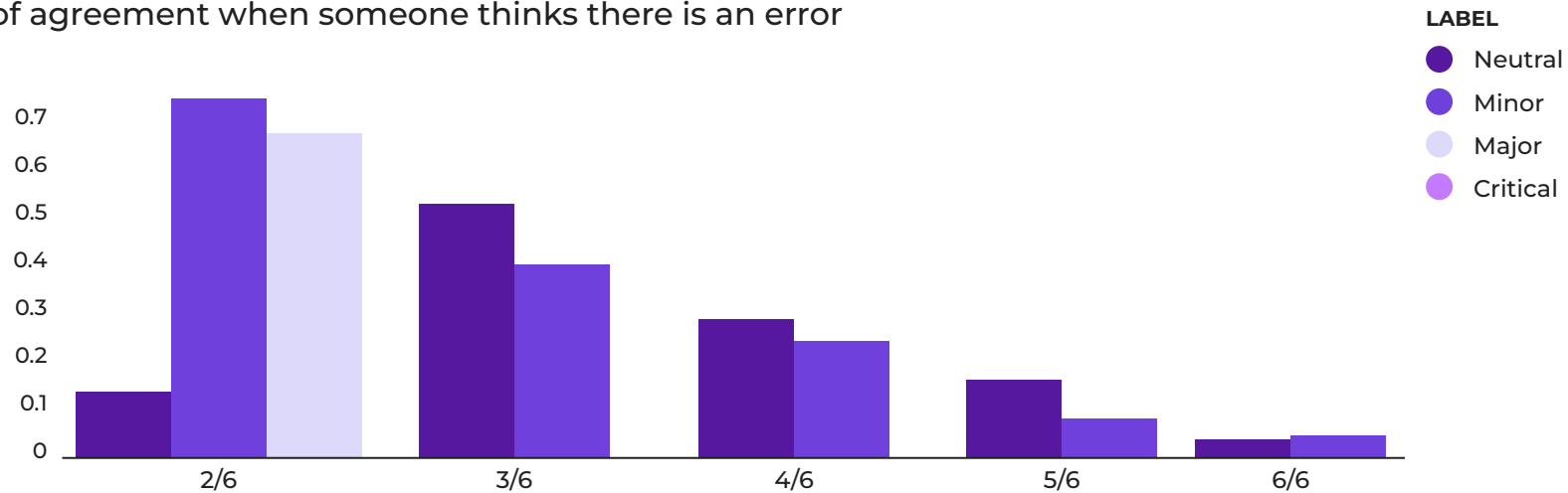
Vision for Automated MQM



More annotators = less agreement

An experiment in German

% of agreement when someone thinks there is an error



Semantic scoring: Beyond textual scoring

When textual scoring fails

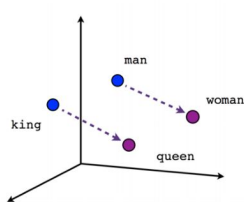
Let's eat, Grandmother!
Let's eat Grandmother

Ref	Hyp	Words	BLEU	PE Distance	TER %	CharacTER	ChrfScore
After lengthy deliberations, the jury found the defendant guilty	After lengthy deliberations, the jury found the defendant <u>not</u> guilty	11	89.27	0.06	0.1	0.06	95.75
Insert the oil extractor tube and pour gas into the engine	Insert the oil extractor tube and pour <u>oil</u> into the engine	12	93.75	0.05	0.09	0.05	87.75

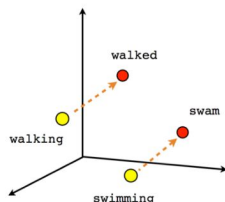
Semantic Scoring: Beyond textual scoring 2

Word embeddings

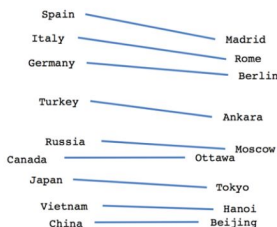
1. Numerical representation of words
2. Words that are similar are plotted close to each other



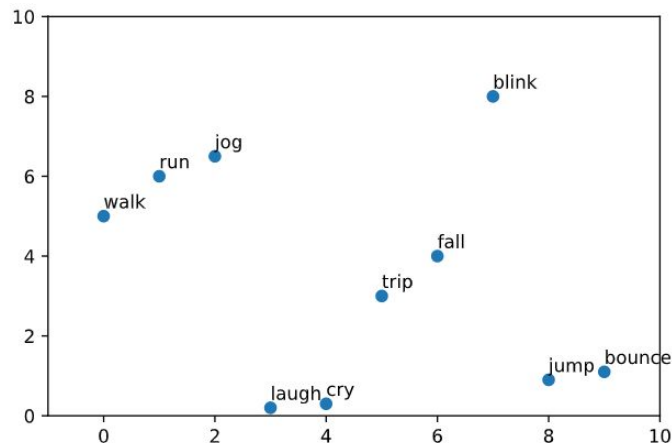
Male-Female



Verb tense



Country-Capital



Semantic Scoring: Beyond textual scoring 3

Sentence Embeddings

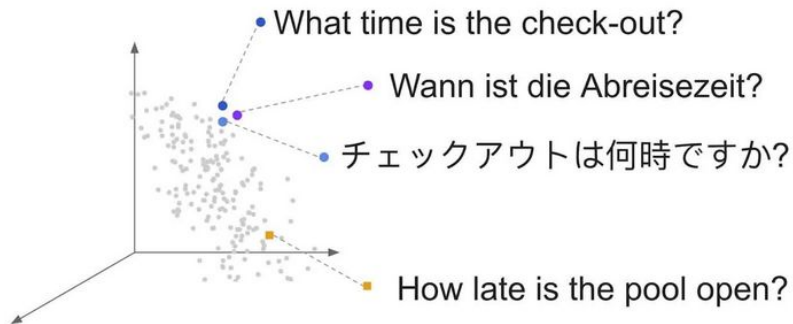
1. Expand concept of word embeddings to sentences
2. Consider source and target

"Lion is the king of the jungle."



"The tiger hunts in this forest."

"Everybody loves New York."



Semantic Scoring: Beyond textual scoring 4

Sentence Embeddings

1. Lower scores indicates source and target may not match semantically
2. Erroneous information for user
3. Information may not adhere to be brand tone or style

EN	FR	distiluse-base-multilingual-cased-v1_cosine_sim	laser_cosine_sim	LaBSE_cosine_sim
Hi Alyssa, We understand your concern that you did not receive an item from this restaurant.	Bonjour Florence, Nous comprenons votre inquiétude de ne pas avoir reçu un article de ce restaurant.	0.691071272	0.862825155	0.809140086
Hi Alyssa, We understand your concern that you did not receive an item from this restaurant.	Quelle est l'heure de départ	-0.094488777	0.279111385	0.174875483
I apologize for the inconvenience it has caused.	Je m'excuse pour le désagrément que cela a causé.	0.915562093	0.938958287	0.918583035

