

Quality Estimation for Machine Translation

Or

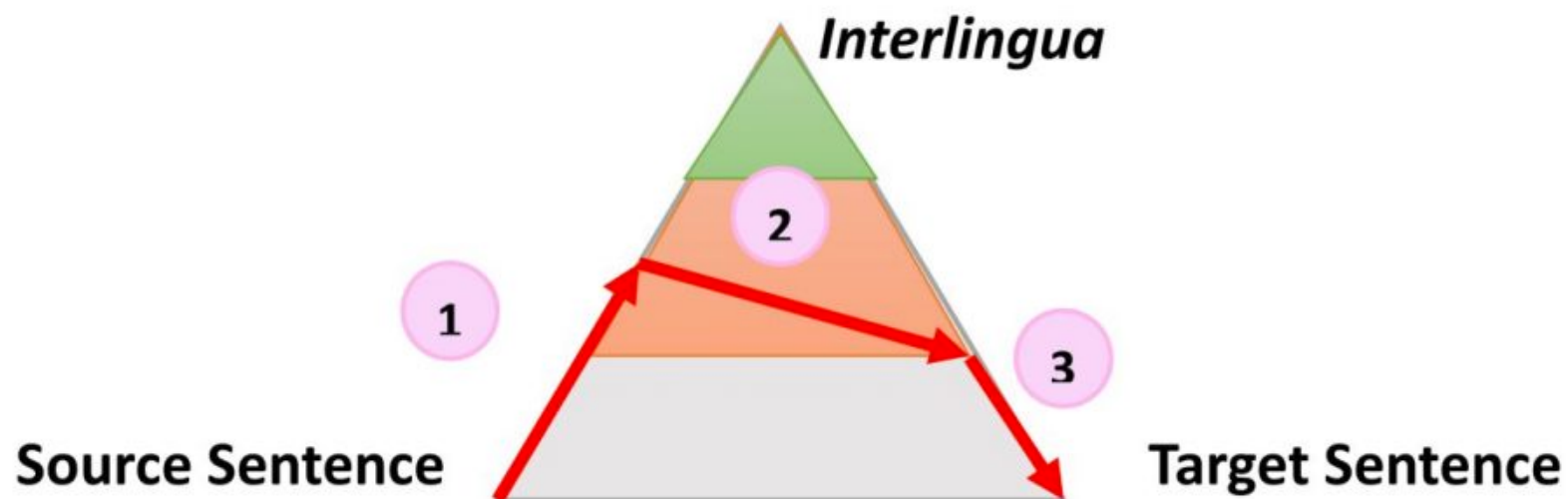
*“A tale of conducting applied research
with a strict deadline and no data or budget”*



CROSSLANG
TRANSLATION AUTOMATION

NLP at CrossLang

- Machine Translation built on top of NLP stack
- Multi-task learning
- Hyperparameter optimization
- Fundamental research: iMinds, imec, VLAIO
- Rest of presentation: example project



<http://blog.systransoft.com/how-does-neural-machine-translation-work/>




CROSSLANG
TRANSLATION AUTOMATION

Quality Estimation




- Task description



CROSSLANG
TRANSLATION AUTOMATION

[Questions](#) [Jobs](#) [Documentation](#) BETA [Tags](#) [Users](#)

Best evaluation method for real-time machine translation?



0



I'm aware that there are many different methods like BLEU, NIST, METEOR etc. They all have their pros and cons, and their effectiveness differs from corpus to corpus. I'm interested in real-time translation, so that two people could have a conversation by typing out a couple sentences at a time and having it immediately translated.

What kind of corpus would this count as? Would the text be considered too short for proper evaluation by most conventional methods? Would the fact that the speaker is constantly switching make the context more difficult?

[nlp](#) [precision-recall](#) [machine-translation](#)

[share](#) [edit](#) [flag](#)



asked May 12 at 17:27

[Nicholas Hassan](#)
126 ● 9

Duplicate of [stackoverflow.com/questions/43943674/...](#) – [alvas](#) May 12 at 17:52

[add a comment](#)

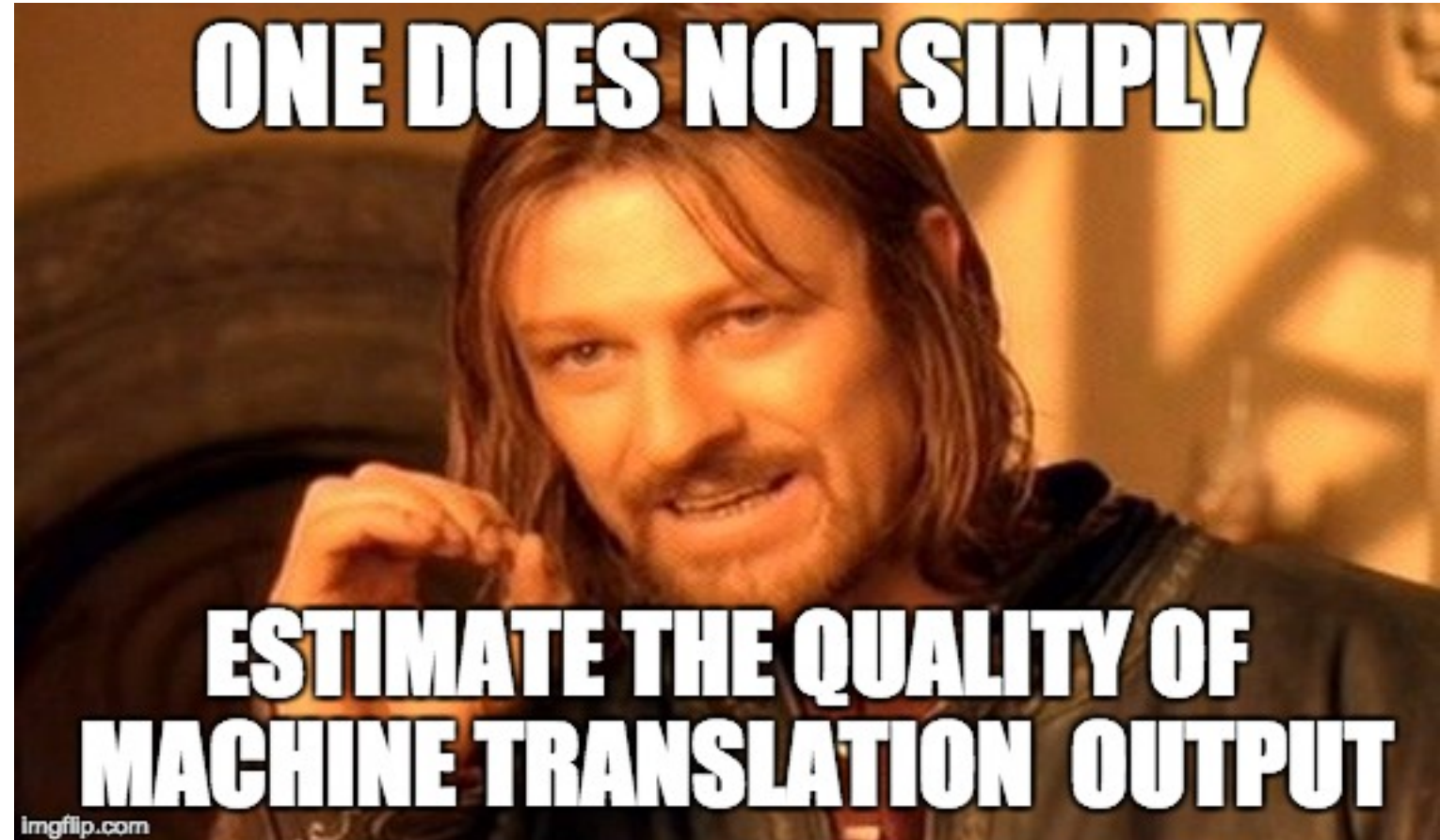
[start a bounty](#)

Constraints

- Research: get Post-Edited data, find better algorithms, develop features
- Industry: sorry, no data, budget or time available. Please make it scalable to run in production.



CROSSLANG
TRANSLATION AUTOMATION



Bootstrapping

- Good enough approach
- Fall back to "abandoned" protocol
- Tweak protocol by lack of data

DOMAIN	Dom1	Dom2	Dom3	TOTAL
DE-EN	800	800	-	1,600
EN-DE	800	800	800	2,400
EN-ZH	-	800	800	1,600
EN-ES	-	800	800	1,600
EN-PT	-	800	800	1,600
EN-FR	-	800	800	1,600
EN-RU	-	800	800	1,600
EN-IT	-	800	-	800
EN-JP	800	800	800	2,400



CROSSLANG
TRANSLATION AUTOMATION

Results

DOMAIN	DE-EN MAE/MRSE		EN-DE MAE/MRSE		EN-ZH MAE/MRSE		EN-ES MAE/MRSE		EN-PT MAE/MRSE		EN-FR MAE/MRSE		EN-IT MAE/MRSE		ALL MAE/MRSE	
Dom1	0.65	0.88	0.68	0.88	-		-		-		-		-		0.73	0.97
Dom2	0.54	0.86	0.94	1.16	0.79	1.06	0.63	0.98	0.77	0.99	0.54	0.76	0.62	0.87	0.76	1.03
Dom3	-		0.80	1.05	0.68	0.95	0.54	0.85	0.86	1.10	0.63	0.95	-		0.79	1.03
LANG	0.63	0.90	0.80	1.03	0.70	0.97	0.52	0.83	0.76	1.02	0.55	0.80	0.62	0.87	0.77	1.04
BULK	0.77								1.04							



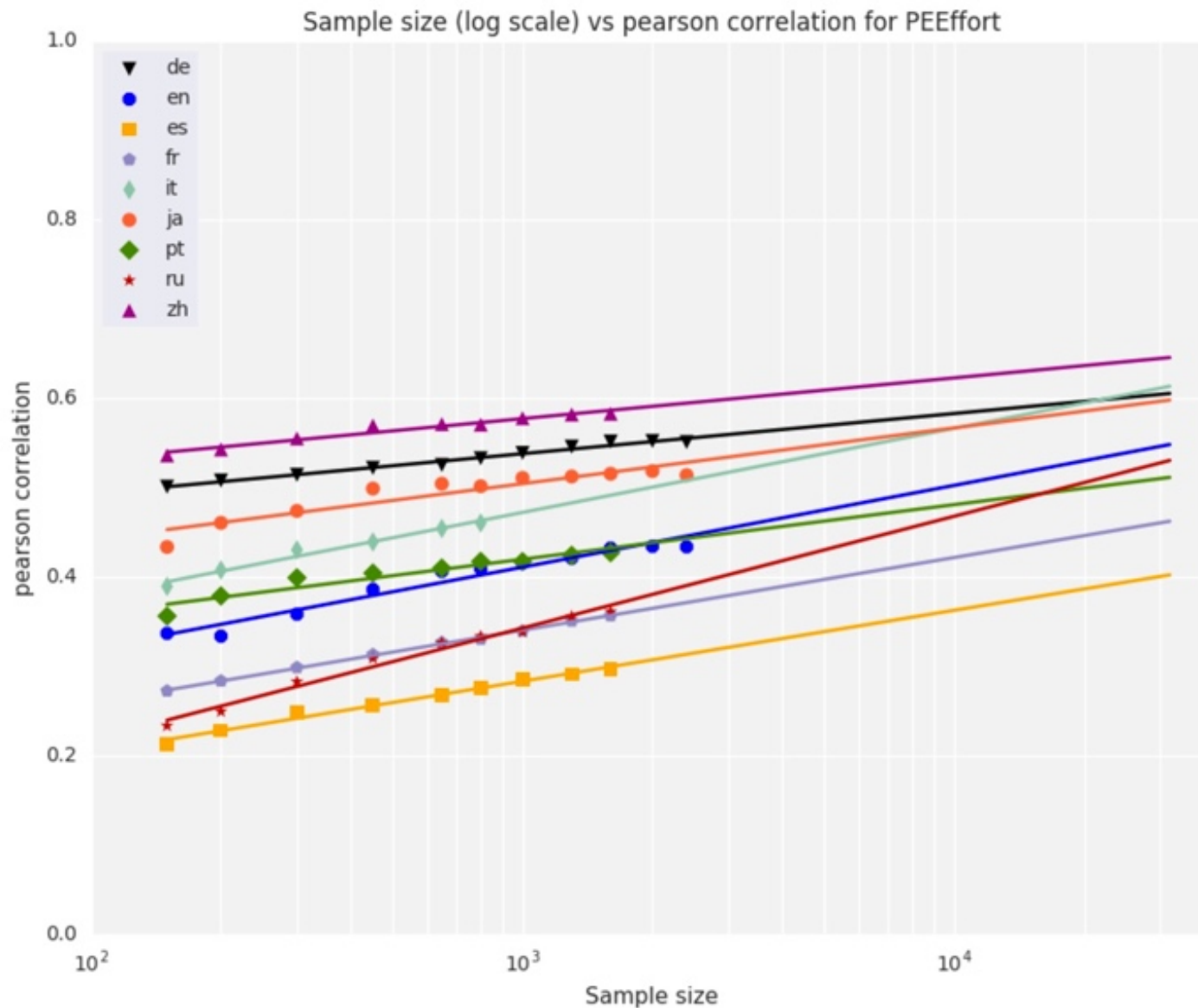
CROSSLANG
TRANSLATION AUTOMATION

Validation

- How bad/good did we do?
- What is the required data set size to do it "the right way"



CROSSLANG
TRANSLATION AUTOMATION

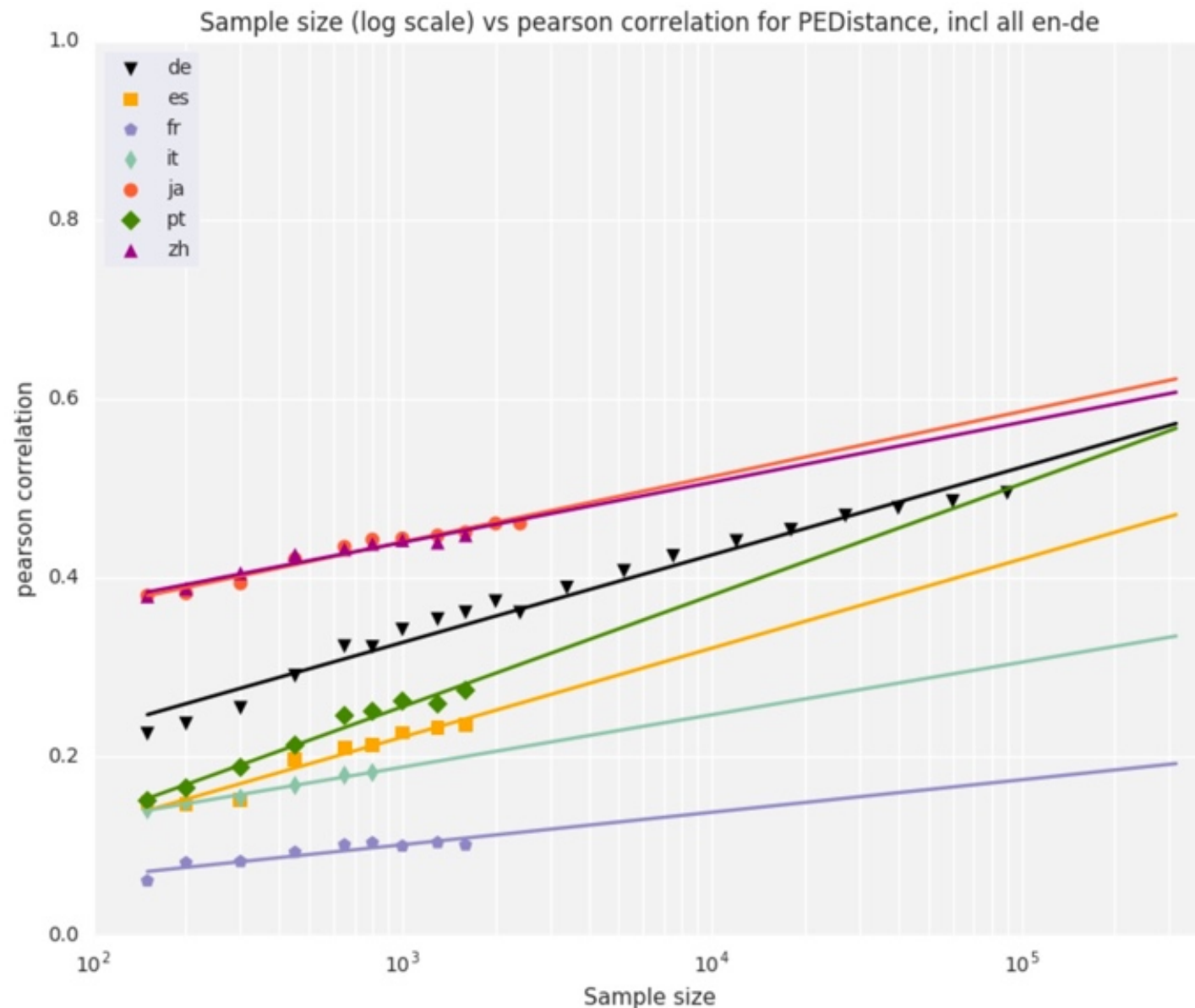


Validation

- How bad/good did we do?
- What is the required data set size to do it "the right way"

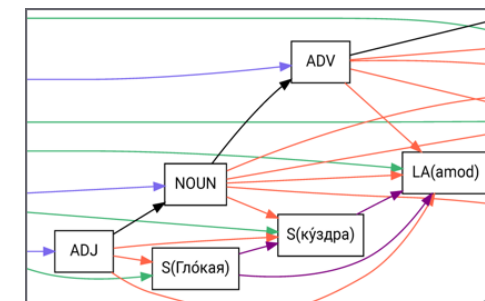


CROSSLANG
TRANSLATION AUTOMATION

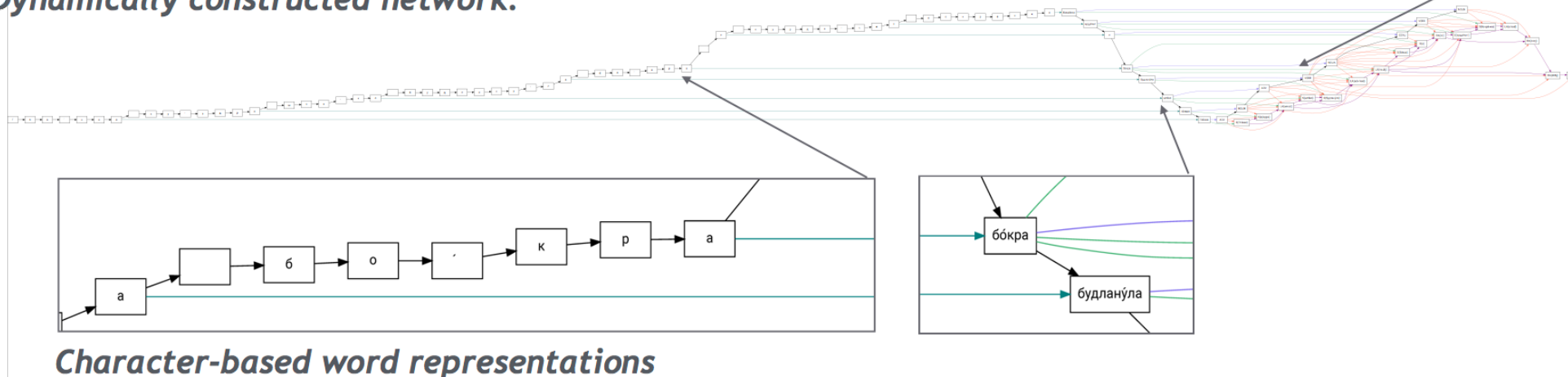


Further feature development

ParseySaurus analysis:



Dynamically constructed network:



CROSSLANG
TRANSLATION AUTOMATION

<https://github.com/tensorflow/models/tree/master/syntaxnet>

Questions?

- joachim@crosslang.com



CROSSLANG
TRANSLATION AUTOMATION