


 ADAPT

 Engaging Content
 Engaging People

DELA - Document-Level Machine Translation Evaluation

Sheila Castilho

Motivation

- Evaluation of document-level MT systems is still performed at the sentence level
- Sentence-level evaluation is not able to recognise the improvements of those systems

Problem Statement

- Overblown claims of human equivalence in NMT were disproved with document-level assessments
- Automatic MT metrics underestimate the quality of NMT and are not suitable for document-level MT evaluation
- Lack of context has a great effect on quality assessment

Potential Impact

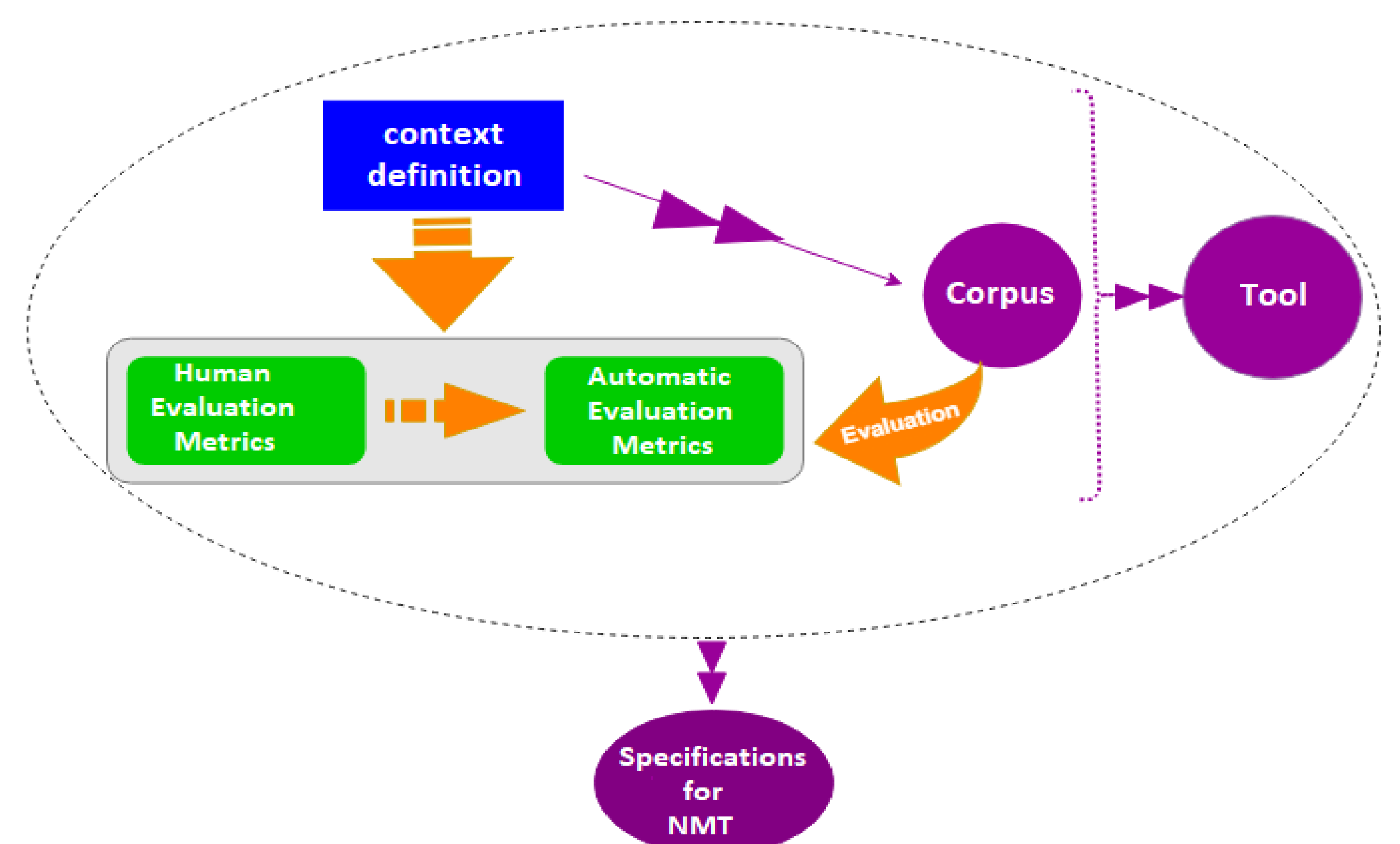
- Positively influence the MT and Translation field as the view on productivity and quality will shift (e.g. textual coherence features will be required in a translation)
- Positively influence the translation industry standards and impact translators' workflow

Novelty of Work

- Document-level evaluation will enable assessment of phenomena that cannot be assessed at sentence level (e.g. suprasentential context, textual cohesion, coherence types of errors, such as mistranslation of ambiguous words)

Methodology

1. Test context for document-level evaluation with non-literal language and ambiguity
2. Build a corpus for non-literal language and ambiguity
3. Identify new ways to evaluate document-level translations with humans (HEM) and automatic metrics (AEM)



Discussion

- Current MT evaluation metrics cannot recognise the improvements of document-level MT systems
- The research will overcome current limitations by implementing next-generation metrics that will be able to capture suprasentential context

Future Work

- Build a document-level translation quality assessment tool
- Define specifications for document-level NMT implementation

References

- Moorkens, J. Castilho, S., Gaspari, F. and Doherty, S. editors. 2018. Translation Quality Assessment: From Principles to Practice. Springer International Publishing.
- Toral, A. Castilho, S., Hu, K. and Way, A 2018. Attaining the unattainable? reassessing claims of human parity in neural machine translation. In Proceedings of the Third Conference on Machine Translation.