Automatic and Manual Metrics for Operational Translation Evaluation

Workshop Programme

08:45 – 09:30 – Welcome and Introduction by Workshop Organizers

09:30 – 10:30 – Talks Session 1

Joke Daems, Lieve Macken and Sonia Vandepitte, *Two Sides of the Same Coin: Assessing Translation Quality in Two Steps Through Adequacy and Acceptability Error Analysis*

Leonid Glazychev, *How to Reliably Measure Something That's Not Completely Objective: A Clear, Working and Universal Approach to Measuring Language Quality*

Mihaela Vela, Anne-Kathrin Schumann and Andrea Wurm, *Translation Evaluation and Coverage by Automatic Scores*

Arle Lommel, Maja Popović and Aljoscha Burchardt, *Assessing Inter-Annotator Agreement for Translation Error Annotation*

10:30 – 11:00 – Coffee break

11:00 – 13:00 – Talks Session 2

Marianne Starlander, *TURKOISE: A Mechanical Turk-based Tailor-made Metric for Spoken Language Translation Systems in the Medical Domain*

Caitlin Christianson, Bonnie Dorr, and Joseph Olive, *MADCAT Evaluation Approach: Operational Accuracy of MT applied to OCR*

Ekaterina Stambolieva, *Continuous Operational Evaluation of Evolving Proprietary MT Solution's Translation Adequacy*

Lars Ahrenberg, *Chunk Accuracy: A Simple, Flexible Metric for Translation Quality*

Michael Carl and Moritz Schaeffer, *Word Transition Entropy as an Indicator for Expected Machine Translation Quality*

Doug Jones and Tamas Marius, *A New Multiple Choice DLPT-STAR Comprehension Test for MT* and *Standardized ILR-Based* and *Task-Based Speech-to-Speech MT Evaluation*

Lena Marg, *Rating Evaluation Methods through Correlation*

Federico Gaspari, Antonio Toral, Arle Lommel, Stephen Doherty, Josef van Genabith and Andy
Way, *Relating Translation Quality Barriers to Source-Text Properties*

13:00 – 14:00 – Lunch break

14:00 – 15:00 – Hands-On Session 1

15:00 – 16:00 – Hands-On Session 2

16:00 – 16:30 – Coffee break

16:30 – 17:30 – Hands-On Session 3

17:30 – 18:00 – Discussion, Potential for Future Collaboration, Next Steps, and Conclusion