

Pre-conference workshop:

Exploring Powerful Tools to Ensure Robust and Reproducible Results in Corpus Linguistics

Convenors: Martin Schweinberger (The University of Queensland, Australia),
Joseph Flanagan (University of Helsinki),
Gerold Schneider (University of Zurich)

Wednesday August 18th, 2021

10:00 - 10:25: Start - Introduction: Joe (Gerold, Martin)

Case Study 1 (chair: Martin)

10:30 - 11:00 Paper 1: Michael Haugh and Simon Musgrave:

Looking for a good laugh: A combinatorial approach to identifying pragmatic phenomena in spoken corpora

Technical Platforms (chair: Joe)

11:10 - 11:25: Martin Schweinberger:

Language Technology and Data Analysis Laboratory

11:25 - 11:40: Gerold Schneider:

Text Crunching Centre

11.40 - 12:30: Lunch Break

Case Studies 2 (chair: Martin)

12:30 - 13:00: Paper 2: Claus Weihs and Sarah Buschfeld:

Variable importance in ensembles from undersampling: generalizability meets robustness

13:05 - 13:35: Paper 3: Sean Wallis:

How do we understand significance?

13:40 - 14:10: Paper 4: Maud Reveilhac and Gerold Schneider:

Detecting Stance from Corpora of Short Texts

14:10 - 14:30: Break

Discussion (chairs: Gerold, Joe, Martin)

14:30 - 16:00: Panel Discussion about Reproducibility and Robustness by Stefan Gries, Stefan Evert, Gerold Schneider, Joseph Flanagan, Martin Schweinberger, Michael Haugh, Simon Musgrave, Claus Weihs, Sarah Buschfeld, Sean Wallis, and Maud Reveilhac

16:00 - 16:15: EoB