Natural language and AI.
New perspectives for linguistic studies

In recent years, Artificial Intelligence (AI) has made significant strides across various disciplines. The integration of AI applications, particularly Large Language Models (LLM), into linguistic studies has opened new horizons for the analysis of natural language. From morphosyntax to semantics and variation linguistics, these technologies provide linguists with the opportunity to explore complex linguistic phenomena. This development has led to the automation of linguistic tasks such as text generation, translation, and corpus annotation to an extent that was previously unimaginable.

However, the application of AI in linguistic research also reveals challenges. One significant issue lies in the need to provide adequate training data for AI models, covering a wide range of linguistic phenomena and structures. Often, these data are incomplete, uneven, or even erroneous, which can compromise the reliability of AI systems. Another obstacle is the fact that AI models may inherit implicit biases from the existing data on which they are trained. This can result in distortions in the results and compromise the neutrality and objectivity of linguistic analyses. Furthermore, language variation seems to pose a challenge. AI models must be able to recognize and process this diversity appropriately. This is often difficult as the models may be constrained by certain linguistic patterns or norms.

The planned workshop will be a forum to discuss how and to what extent AI applications (such as ChatGPT, DialoGPT, Meena, BlenderBot, etc.) may be relevant for linguistic studies. By merging theoretical approaches in linguistics with modern AI methods, the potentials and challenges of these technologies for linguistic research will be explored.

The workshop will focus on – but will not be limited to – the following research questions:

- How can AI applications be used to investigate and compare grammaticality in different languages? Which applications come closest to native speaker intuition?
- To what extent can Large Language Models (LLM) be used for automatic corpus annotation and analysis to identify and understand linguistic variation?
- What role do semantic models and neural networks play in translating between languages with different syntactic structures?
- How can Natural Language Processing (NLP) techniques be used to examine the meaning and usage of language variation in different social contexts?

Linguists from various theoretical orientations are invited to participate in this working group and present empirical studies addressing current developments and future perspectives in the field of natural language processing and AI for linguistic studies. The language of the workshop is English.

Abstracts for 20-minute presentations (plus 10 minutes discussion) must be submitted in .doc and in PDF format. They should not exceed one page in length (excluding references), in a font size no less than 12pt, and with margins of 1 inch/2.5cm. Every author can submit at most one single-authored abstract and one co-authored abstract.

Abstracts are to be submitted via e-mail no later than September 01st, 2024 to:
catasso@uni-wuppertal.de
thomas.scharinger@uni-jena.de

Contact persons:
Nicholas Catasso (catasso@uni-wuppertal.de)
Thomas Scharinger (thomas.scharinger@uni-jena.de)