

The Human Factor: Discourse Biases in Humans and Machines

Florian Kankowski (Bielefeld University)

Do modern large language models mirror humans in the way they use language -- and, if not, how can we explain why they diverge? This question is a main impetus behind my PhD project.

In this talk, I will give a cross-sectional overview of my dissertation project about the study of language use in LLMs. Starting with the background of the discourse phenomenon at the heart of the project, Implicit Causality, I will motivate the project and its goals, and present results from a recently (conditionally) accepted paper analyzing Implicit Causality discourse biases in a range of LLMs. From there, I will discuss the open questions that arose from this paper, sketching out the further (planned) work of this PhD project (and beyond), and how I will try to relate it back to linguistic creativity and citizen science.