Constructing explainability: Monitoring the understanding of explanations

Hendrik Buschmeier (Universität Bielefeld) Petra Wagner (Universität Bielefeld)

The newly funded TRR 318 Constructing Explainability investigates the realization of transparent AI systems which are able to explain the underlying reasons for algorithmic decision processes. Unlike traditional approaches to explanatory AI (XAI), the main hypothesis of the TRR is that the human users must play an active role in reaching satisfactory explanations in a co-constructive process. The goal of the subproject A02 is to investigate and computationally model the multimodal signalling of different degrees of (non-)understanding in ongoing explanations, which (probably) shape the way that an explanatory dialogue unfolds.