

Title: ASR, Transformers and Explainable AI

Louis ten Bosch

Radboud University, Nijmegen, NL

<https://www.ru.nl/en/people/bosch-l-ten>

In this presentation we will discuss a number of ways in which deep networks (Speech Foundation Models, LLMs) can be used and interpreted, especially for the purpose of acoustic-phonetic analyses and automatic speech recognition.

I will discuss various deep network topologies and a number of methods that can be used to interpret latent representations. We will put quite some focus on Transformer layers, and go into some of the mathematical details.

Finally I will discuss a number of key papers and experiments, dealing with the quality of grapheme segmentation using a recent variant of Whisper, with structure discovery (ordering structure in embedding sequences related to regular expressions, in the Chomsky hierarchy), and with explainability using mechanistic interpretability (and the problems thereof).