

WS 2026/2027

Lecture / Exercise

Fundamentals of ASR

Type	Schedule / Room	Start	Docent
V3			Priv.-Doz. Dr. Ralf Schlüter
Ü1	check RWTHonline	check RWTHonline	Zijian Yang

Content:

The aim of automatic speech recognition is to transform an speech audio signal into written text. An automatic speech recognition system consists typically of four components:

- **signal analysis:** extracts time-dependent features from the acoustic signal
- **acoustic model:** establishes link between acoustic features and written and or phonetic representation
- **language model:** covers syntactic, semantic language constraints
- **search process:** determines best recognized word sequence based on the statistical acoustic and language models.

The lecture gives an introduction into the individual components of an automatic speech recognizer, their modeling and training, as well as efficient decoding concepts.

Requirements:

knowledge of probability theory / statistics

References:

- Lecture notes of this lecture, together with comments/notes during lecture.
- L. Rabiner, B. H. Juang: *Fundamentals of Speech Recognition*, Prentice Hall, Englewood Cliffs, New Jersey, 1993.
- F. Jelinek: *Statistical Methods for Speech Recognition*, MIT Press, Cambridge, MA, 1998.
- R. Prabhavalkar, T. Hori, T. Sainath, R. Schlüter, S. Watanabe: "End-to-End Speech Recognition: A Survey," *IEEE Trans. on Audio, Speech and Language Processing*, Oct. 2023, ieeexplore.ieee.org/document/10301513

Subsequent courses planned:

In addition, our team offers the master-level lecture "Automatic Speech Recognition Search".

Others:

The lecture will be given in English.

For more information, please contact:

PD Dr. Ralf Schlüter, Lehrstuhl Informatik 6, Tel. 80-21612,

E-Mail: schlueter@ml.rwth-aachen.de