



Seminar Hora Informaticae

Institute of Computer Science, Prague

Tuesday, April 16, 2024, 14.00 – 15.30 (2 – 3:30 PM) CEST

Meeting Room 318, Address: Pod Vodárenskou věží 2, Prague 8

ZOOM Meeting ID: 954 7823 4977 , Passcode: 712564

ZOOM: <https://cesnet.zoom.us/j/95478234977?pwd=dXoyekFHbDJ0MkNrTjVVS3F2STZqUT09>



**Jindřich Matoušek, Department of Cybernetics, Faculty of Applied Sciences, UWB Pilsen:
Czech Speech Synthesis in the AI Era.**

In this talk, we will introduce the current trends in the field of text-to-speech synthesis (TTS) and focus on Czech speech synthesis. TTS technology aims to read arbitrary text automatically, without human intervention, but still in a quality that is indistinguishable from human speech. We will show that significant progress has been made recently and that current approaches based on machine learning and artificial intelligence (especially deep neural networks) are already approaching this goal. We will present some popular neural speech models and architectures. The talk will be complemented by examples of practical applications of Czech speech synthesis: personalized speech synthesis for people at risk of losing their voice, automatic reading of textbooks for visually impaired students, or the synthetic voice of Karel Gott, which we have created within the framework of the Czech Radio project Gott Forever.

Due to the relation of the topic to the pronunciation and intonation of the Czech language, the lecture will be given in Czech.

References:

- (1) Tan, X. Neural Text-to-Speech Synthesis. Artificial Intelligence: Foundations, Theory, and Algorithms (AIFTA). Springer Singapore, 2023. DOI: 10.1007/978-981-99-0827-1.
- (2) Kunešová, M., Matoušek, J., et al. Ensemble of Deep Neural Network Models for MOS Prediction. In: ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023. DOI: 10.1109/ICASSP49357.2023.10095676.
- (3) Example of a practical application of Czech speech synthesis: Project Gott Forever.

Jindřich Matoušek (<https://www.zcu.cz/cs/Employees/person.html?personId=17268>) works at the Faculty of Applied Sciences of the University of West Bohemia in Pilsen, Department of Cybernetics, and at the New Technology for Information Society (NTIS) Centre, where he has been teaching and researching speech technologies, especially speech synthesis. His recent research focuses on end-to-end generative speech synthesis models and automatic evaluation of synthetic speech quality. He has extensive experience in basic and applied research projects and is the author or co-author of more than 160 scientific publications, two patents, and more than 40 applied results.

HORA INFORMATICAЕ (meaning: TIME FOR INFORMATICS) is a broad-spectrum scientific seminar devoted to all core areas of computer science and its interdisciplinary interfaces with other sciences and applied domains. Original contributions addressing classical and emerging topics are welcome. Founded by Jiří Wiedermann, the seminar is running since 1994 at the Institute of Computer Science of the Czech Academy of Sciences in Prague.

<https://www.cs.cas.cz/horainf>