



Topic of dissertation thesis

Academic year **2026/2027**

Title (Engl./Slovak)	Robust and Time-Invariant Speaker Voice Modelling Using Deep Learning / Robustné a časovo invariantné modelovanie hlasu hovoriaceho s využitím hlbokého učenia		
Institute	Faculty of Electrical Engineering and Information Technology		
Place	University of Žilina		
PhD. programme	telecommunications		
Supervisor	doc. Ing. Roman Jarina, PhD. Department of multimedia and information-communication technology		
Co-supervisor	Ing. Maroš Jakubec, PhD		
Study form	Internal		
Study duration	3 years (internal form)		
Study language	Slovak, English		
Start date	1.9.2026		
Research domain	Speech processing, Speaker recognition, Machine learning, Artificial intelligence, Biometrics		
Contact person	Phone: +421 41 513 2268	E-mail: maros.jakubec@uniza.sk	Web-page https://kmikt.uniza.sk/

Dissertation topic abstract

This dissertation topic focuses on modern AI methods for analysis of human voice and sound with emphasis on speaker recognition and long-term stability of voice representations. Human voice naturally changes over time due to aging, health, emotional state, and recording conditions, which makes reliable speaker modelling a challenging task. The research will explore new AI and deep learning approaches for building robust voice representations that remain reliable across different conditions and time periods. The expected outcome is improved methods for speaker modelling and sound analysis applicable in speech biometrics, human-machine interaction, and audio-based intelligent systems.

Extended information, research responsibilities and tasks of PhD. candidate

Sound and speech carry a large amount of information about a person, environment, and context. Modern artificial intelligence makes it possible to analyse this information much more deeply than before. In recent years, research has shifted toward large-scale audio foundation models trained on massive amounts of unlabelled sound data. These models learn general acoustic representations that can later be adapted to specific tasks such as speaker recognition. At the same time, there is growing interest in understanding how voice characteristics evolve over long periods of time, especially due to aging and natural physiological changes. This creates a need for speaker representations that are not only accurate in short-term tests, but also stable across years and across different recording conditions. Efficient neural architectures, representation learning, and advanced training strategies are therefore becoming key research directions in modern audio and speech processing.

The PhD research will focus on experimental work with speech and audio data, development of AI-based sound representations, and testing of new methods for robust speaker modelling. The work will include designing neural models, working with real and synthetic audio data, evaluation under different acoustic conditions, and systematic experimental validation. Results are expected to contribute to modern speech technologies, audio biometrics, and intelligent audio systems, together with scientific publications and collaboration within ongoing research activities.

The project provides access to speech datasets, experimental infrastructure, and opportunities for collaboration, scientific publications, and applied research activities.

Candidate profile

Required qualifications:

Education and Professional Background

- Master's degree (MSc, MEng, or equivalent) in Telecommunications, Computer Science, or a closely related technical field.

Technical Skills

- Intermediate programming skills in Python,
- Experience with data analysis, signal processing, and machine learning methods,
- Interest in working with speech and audio data.

Research and Personal Competences

- Strong interest in speech technologies, artificial intelligence, and data-driven methods,
- Strong analytical skills and solid mathematical foundation,
- Ability to communicate scientific results effectively (e.g., through publications, conference presentations, or patents),
- Proficiency in written and spoken English.

Project relevance

The dissertation topic is directly linked to the research project:

- [APVV-MVP-24-0389 — Understanding speaker voice changes due to aging for robust speaker voice representation](#), focused on analysis of age-related voice variability and development of robust speaker representations.