

EUGEN SEDLAR

linkedin.com/in/eugen-sedlar | github.com/egi03

EDUCATION

- Faculty of Electrical Engineering, Computer Science and IT (FERIT)** Osijek, Croatia
MSc Computer Science — Data Science & AI/ML track *Oct. 2025 – Nov. 2027*
- Faculty of Electrical Engineering, Computer Science and IT (FERIT)** Osijek, Croatia
BSc Software Engineering — summa cum laude *2022 – 2025*

EXPERIENCE

- ML Engineer** Oct. 2025 – Mar. 2026
PrivateSynapse.ai *Remote*
- Built production Audio RAG system achieving 83% answer correctness (RAGAS) across a 10k+ audio query dataset, using Whisper, Pinecone, GPT-4o, and LangChain/LangGraph.
 - Designed end-to-end pipeline for automated call-centre QA: audio ingestion, transcription, semantic chunking, vector retrieval, and LLM-graded evaluation.

PROJECTS

- AudioAgent-QA** | *LangGraph, LangChain, ChromaDB, NeMo, FastAPI, Docker* 2025
- Built an audio Q&A agent with speaker diarization and emotion recognition, achieving 83% answer correctness (LangGraph ReAct, 40-question custom eval, Gemini 2.0 Flash judge).
 - Designed 3-mode inference pipeline (fully local GPU / hybrid / cloud) with Nvidia Parakeet TDT ASR (3.2% WER) and semantic chunking improving retrieval MRR by 12% over fixed chunking.
- Dialogo** | *faster-whisper, Piper TTS, FastAPI, aiortc, Silero VAD, TypeScript* 2026
- Engineered real-time multilingual voice AI (EN/HR/DE/ES/FR) achieving sub-800ms time-to-first-audio via async LLM-to-TTS chunking — LLM streams to TTS while generating the next sentence.
 - Replaced WebSocket transport with WebRTC (UDP, aiortc) and implemented VAD-triggered filler audio, eliminating perceived latency on consumer hardware (RTX 4070 Super, 12 GB VRAM).
- Fall Detection System** | *BiLSTM, MediaPipe, PyTorch, Python* | *Paper: SST 2026* 2025
- First-authored paper (accepted SST 2026) on real-time skeleton-based fall detection; BiLSTM ensemble achieved AUC 0.897 under LOSO cross-validation (URFD) and 93.9% cross-dataset sensitivity on unseen UP-Fall benchmark.
 - Deployed edge-optimised inference pipeline with 15 biomechanical features, 5-state alarm FSM, and ~12 FPS CPU-only inference on AMD Ryzen 7 7700; 261 unit and integration tests.
- Music Recommendation System** | *TensorFlow, Django, Spotify API, Last.fm API* 2024
- Built hybrid recommendation engine combining neural collaborative filtering (TensorFlow NCF, 50-dim embeddings, negative sampling 4:1) with content-based Last.fm tag similarity and 3-strategy weighted ensemble.
 - Delivered full-stack Django platform with Spotify OAuth, offline evaluation (Precision@K, NDCG@K), and 135 passing tests across recommender, social, and API layers.

COMPETITIONS & AWARDS

- 1st Place — EU Colours Reinforcement Farming Hackathon** 2025
PPO-based RL agent; 100% win rate throughout entire competition
- 2nd Place (National) — IEEEExtreme** 2023
Competitive programming; 2nd place nationally

TECHNICAL SKILLS

Languages: Python, TypeScript, SQL, Bash, Kotlin, C#
ML/AI: PyTorch, TensorFlow, LangChain, LangGraph, Whisper, Hugging Face, RAGAS, scikit-learn
Audio/Speech: faster-whisper, NeMo (Parakeet, TitaNet), Piper TTS, StyleTTS2, Silero VAD, aiortc
Infrastructure: GCP, Terraform, Docker, FastAPI, PostgreSQL, ChromaDB, Pinecone, Redis, Git
Web: Next.js, Django, Prisma, Tailwind CSS

CERTIFICATIONS

- Google Cloud — Generative AI Leader (GAIL) 2026
- Deep Learning Specialization — DeepLearning.AI / Coursera 2026
- Mathematics for ML and Data Science — DeepLearning.AI / Coursera 2025
- Machine Learning Specialization — DeepLearning.AI / Coursera 2024