

Yassine El Kheir

Speech, Audio & Multimodal AI Researcher

(+49) 157 36279520 • elkheiryassine0@gmail.com • Website • GitHub • LinkedIn • Google Scholar

Final-year PhD researcher (DFKI / TU Berlin) building speech and audio intelligence that generalizes in the lab and ships in production – across self-supervised representations, efficient architectures, and speech LLMs. Trained at Télécom Paris (top 5%), KTH, and TU Berlin, with research at NII Tokyo alongside the ASVspoof founders. Seeking **Research / Applied Scientist** roles from **November/December 2026**.

SELECTED HIGHLIGHTS

- **Competition:** 1st of 97 teams at the **ESDD 2026 ICASSP** grand challenge (Track 2) and **1st on Deepfake-Eval 2024**, surpassing all commercial systems.
- **Production:** SOTA deployed detector (Gretchen AI) and multilingual speech pipelines across **150+ languages** (CAMB.AI; IMAX, Comcast, Google Cloud).
- **Open source:** **DeepFense** – 455+ pretrained models and 12 benchmarks, adopted by the research community.
- **Research:** 20+ papers (Interspeech, ICASSP, ACL, EMNLP, NAACL); contributor to the **Fanar** Arabic-centric multimodal LLM training effort.

PROFESSIONAL EXPERIENCE

PhD Researcher & AI Scientist

DFKI / TU Berlin

2024 – Present

Berlin, Germany

- **Project Coordinator – Fake-O-Meter:** day-to-day coordinator of a multi-partner **BMBF** (German Federal Ministry of Education & Research) reference project, delivering public deepfake-detection tools for election integrity.
- **SSL representations & interpretability:** ran the first large-scale layer-wise analysis of speech SSL models (Wav2Vec 2.0, WavLM, HuBERT), showing early layers carry the discriminative signal and enabling $\approx 3\times$ faster inference (*NAACL Findings 2025*).
- **Efficient & state-space architectures:** designed BiCrossMamba-ST, among the first Mamba SSMs for speech, SOTA with **28% fewer parameters** than Transformer baselines (*Interspeech 2025*); parameter-efficient multi-scale adapters (*ICASSP 2026*).
- **Generalization under distribution shift:** spectral + SSL fusion for robustness to unseen generators (*WASPAA 2025*).
- **Competition:** led the DFKI-SLT system to 1st of 97 teams (Track 2) and 2nd (Track 1) at the **ESDD 2026 ICASSP** grand challenge (*results paper*).
- **Open source:** co-created **DeepFense**, the first modular, YAML-driven audio deepfake-detection framework (*Interspeech 2026*).

Visiting Researcher

NII (National Institute of Informatics)

Dec 2025 – Apr 2026

Tokyo, Japan

- Collaborated with the **ASVspoof founders** (Xin Wang, Junichi Yamagishi) on **speech LLMs** for audio understanding and reasoning.
- Adapted LLM-based speech evaluation with explicit **textual grounding** – models reason about speech in natural language and produce explainable feedback, moving beyond binary classification (*IEEE SLT 2026*).

Founding Research Scientist (part-time)

Gretchen AI

Jan 2025 – Present

Berlin, Germany

- Co-lead R&D for proprietary multimodal generative-media detection; built the detector ranked **1st on Deepfake-Eval 2024** ($\sim 83\%$ accuracy), outperforming all commercial systems.
- Own real-time inference and continuous retraining driven by biweekly drift monitoring; shipped enterprise-specific variants (media-tech, face-specific, region-specific).

Research Consultant, Speech AI (part-time)

CAMB.AI

Apr 2025 – Present

Remote

- Advised on ASR, TTS, and real-time-translation architecture for a production multilingual dubbing platform spanning **150+ languages** (clients incl. IMAX, Google Cloud, Comcast, NASCAR).
- Adapted SOTA research to production constraints; lead synthetic-speech quality, artifact, and robustness evaluation at scale.

Research Associate

Qatar Computing Research Institute (QCRI)

2022 – 2024

Doha, Qatar

- **Fanar LLM (training team):** contributor to Qatar's Arabic-centric multimodal LLM; co-designed MorphBPE, a morphology-aware tokenizer improving fertility for morphologically rich languages.
- **QVoice:** proposed the first end-to-end mispronunciation-detection system for Modern Standard Arabic (*Interspeech 2023*); grew it into a deployed **learning application** with 6+ follow-on papers.

- **Multilingual / accented speech recognition** (English, Mandarin, Arabic): SSL, end-to-end, and augmentation methods incl. **SpeechBlender** (*SLaTE 2023*) and an L1-aware framework for native-language transfer (*ICASSP 2024*).

EDUCATION

- 2024 – 2027 **Ph.D., Computer Science** — TU Berlin / DFKI. Thesis: *Robust Audio Deepfake Detection: from SSL Models to Audio LLMs* (advisors: S. Möller, T. Polzehl).
- 2021 – 2022 **M.Sc., Machine Learning** (double degree) — KTH Royal Institute of Technology, Stockholm.
- 2019 – 2022 **Diplôme d'Ingénieur** (B.Sc. + M.Sc.) — Télécom Paris. GPA 3.93/4.00, top 5%.

SELECTED PUBLICATIONS

20+ publications • [410]+ citations, h-index [9] (Google Scholar) • first author unless noted.

Speech LLMs & Generative-Media Detection

- 2026 **Y. El Kheir** et al., “DeepFense: A Unified, Modular Framework for Robust Deepfake Audio Detection”, *Interspeech (Long Track) 2026*.
- 2026 **Y. El Kheir** et al., “Textual Acoustic Grounding for Generalizable LLM-Based Deepfake Voice Detection”, *IEEE SLT 2026* (under review).
- 2026 Y. Xin, **Y. El Kheir** et al., “Evidence Subspace Projection: Measuring How Much Evidence Explains Deepfake Detection in Self-Supervised Speech Models”, *Interspeech 2026*.
- 2026 **Y. El Kheir** et al., “Parameter-Efficient Multi-Scale Adapter for Synthetic Speech Detection”, *IEEE ICASSP 2026*.
- 2026 **Y. El Kheir** et al., “DFKI-SLT for ESDD 2026: BiCrossMamba-ST with Attentive SSL Fusion”, *IEEE ICASSP 2026*.
- 2025 **Y. El Kheir**, T. Polzehl, S. Möller, “BiCrossMamba-ST: Speech Deepfake Detection with Bidirectional Mamba Spectro-Temporal Cross-Attention”, *Interspeech 2025*.
- 2025 **Y. El Kheir** et al., “Two Views, One Truth: Spectral and SSL Fusion for Robust Speech Deepfake Detection”, *IEEE WASPAA 2025*.
- 2025 **Y. El Kheir** et al., “Layer-wise Analysis of SSL Models for Audio Deepfake Detection”, *Findings of NAACL 2025*.

Multilingual Speech Recognition & Pronunciation

- 2026 **Y. El Kheir** et al., “IQRA 2026: Interspeech Challenge on Automatic Pronunciation Assessment for MSA”, *Interspeech 2026*.
- 2024 **Y. El Kheir** et al., “Beyond Orthography: Recovery of Short Vowels and Dialectal Sounds in Arabic”, *ACL 2024*.
- 2024 **Y. El Kheir** et al., “L1-Aware Multilingual Mispronunciation Detection Framework”, *IEEE ICASSP 2024*.
- 2023 **Y. El Kheir** et al., “Automatic Pronunciation Assessment — A Review”, *Findings of EMNLP 2023*.
- 2023 **Y. El Kheir** et al., “SpeechBlender: Speech Augmentation Framework for Mispronunciation Data Generation”, *SLaTE 2023*.
- 2023 **Y. El Kheir** et al., “QVoice: Arabic Speech Pronunciation Learning Application”, *Interspeech 2023*.

LLMs & Speech Representation Learning

- 2025 E. Asgari, **Y. El Kheir** et al., “MorphBPE: A Morphology-Aware Tokenizer for Morphologically Rich Languages”, *ACL 2026*.
- 2024 **Y. El Kheir** et al., “Speech Representation Analysis Based on Inter- and Intra-Model Similarities”, *XAI-SA Workshop @ IEEE ICASSP 2024*.

TECHNICAL SKILLS

Core	Python, PyTorch, CUDA, Bash, Git, Docker, SLURM/HPC.
Speech & NLP	HuggingFace (Transformers, PEFT, Datasets), SpeechBrain, ESPnet, fairseq, Kaldi.
Models	SSL: Wav2Vec 2.0, HuBERT, WavLM, MERT, EAT. LLMs: Qwen, LLaMA, LLM-Foundry.
Languages	Arabic (native), English (professional), French (professional), German (intermediate).

AWARDS, COMMUNITY & SERVICE

Awards	1 st (Track 2) & 2 nd (Track 1), ESDD 2026 Challenge @ ICASSP (97 teams) • 1 st , Deepfake-Eval 2024 • WASPAA 2025 Travel Grant • ArabicNLP 2025 Grant.
Organizing	Founder & lead organizer, <i>Iqra'Eval</i> — the first shared task on Arabic pronunciation assessment (ArabicNLP 2025; 2 nd edition accepted as an Interspeech 2026 Challenge).
Open source	DeepFense – audio deepfake-detection framework • MoulSot – 1,500 hr Moroccan Darija corpus; leading the first ASR & TTS effort for the dialect.
Mentoring	Supervising 3 MSc students (TU Berlin); mentor, HackAI Morocco 2025 & 2026.
Service	Reviewer – ACL, EMNLP, NAACL, EACL, COLING, Interspeech, ICASSP, ArabicNLP (2024–2026).