Xiaoqin Feng (冯小琴)

XqFeng-Josie xiaoqin.feng@usc.edu

Curriculum Vitae

Google Scholar

Introduction

I am a Master's student at the University of Southern California (USC) interested in Generative AI, LLM-based agents, human – AI collaboration, and AI for science. My work focuses on building reliable, deployable agentic systems, with experience spanning speech and multimodal modeling, data pipelines, and evaluation. I aim to bridge rigorous research with production-ready systems.

Education

M.S. in Artificial Intelligence **Sep. 2025 – Present**
University of Southern California USC *Los Angeles, California, United States*

M.S. in Software Engineering **Sep. 2016 – May. 2019**
Beijing University of Technology BJUT; overall grade: 86 (max. 100) between "Very Good" and "Good" *Beijing, China*

B.S. in Computer Science **Sep. 2012 – Jul. 2016**
Southwest Minzu University SMU; overall grade: 3.66 (max. 4.0), "Top 5" of 154 students *Chengdu, China*

Professional Experience

Part-time Consultant **Aug. 2025 – Present**
[Mobvoi] *Los Angeles, US*

- **Consultant**: Advised on speech model optimization, multi-agent systems, and real-time information tracking to support product decisions.

Tech Lead **May. 2023 – Jul. 2025**
[Mobvoi] *Suzhou, China*

- **Agents**: Built a production tool-using agent framework for speech/LLM applications, including multi-agent orchestration, tool calling, safety guardrails, and runtime monitoring for reliable and user-friendly task execution. Moyin
- **LLMs & Audio**: Led R&D to integrate LLMs with audio models and pipelines (ASR, TTS, and prompt generation), enabling speech-centric features and improving overall user experience in production. DupDub
- **Data Engineering**: Designed and deployed scalable multimodal data pipelines (text/audio/image) for training and evaluation, covering data collection, versioning, and automated quality checks; supported speech datasets at million-hour scale to accelerate iteration cycles.
- **Evaluation**: Built an end-to-end evaluation system for multimodal features (offline benchmarks + regression testing), establishing quality gates to support continuous optimization and stable releases.
- **Tech Lead**: Owned the roadmap and cross-functional execution across PM/Eng/Research, driving launches from prototype to production; mentored 3–5 engineers and interns.

Senior Speech Algorithm Engineer **Jul. 2019 – May. 2023**
[Mobvoi] *Beijing, China*

- **Speech Synthesis (TTS)**: Developed multilingual and controllable TTS algorithms, including style/emotion/prosody control and G2P for pronunciation accuracy.
- **NLP Frontend**: Built a unified speech-text frontend covering text normalization (TN), polyphone disambiguation, prosody prediction, emotion/style modeling, and audio tokenization for downstream speech systems.
- **System Engineering**: Designed and maintained production TTS/NLP services, supporting advanced capabilities such as controllable TTS and long-form (paragraph-level) synthesis.
- **Mentor**: Working with 3 interns (annual) in NLP and Text-to-speech(TTS)

Algorithm Research(Intern) **Aug. 2018 – Dec. 2018**
TAL AI Lab *Beijing, China*

- **Graph-based Learning**: Developed a Deep Knowledge Tracing (DKT) pipeline using graph embeddings and distance metrics to evaluate student knowledge progression.
- **Key results**: Co-authored paper accepted at *AIED 2019*; awarded 'Outstanding Intern' of the year.

Expertise and Skills

- **Agents & LLMs**: Tool-using agents; agent orchestration; task decomposition; LLM integration; multimodal interaction design (OpenAI Agents, MCP).
- **NLP & Speech**: Text classification; sequence modeling; multi-stage training; multilingual TTS; controllable TTS (style/emotion/prosody); G2P; text normalization; prosody modeling.
- **Data & Evaluation**: Large-scale multimodal dataset construction; data pipelines and quality checks; evaluation for multimodal generation (offline benchmarks, regression testing).
- **Software Engineering**: System design; production services; deployment and monitoring; tooling for training/evaluation.
- **Programming**: Python (8+ yrs), C++ (4+ yrs), Bash/Shell, Git, Docker.
- **Languages**: Mandarin (native); English (basic / conversational).

Teaching Experience

Lecture - Embedded System Design Practice **2018**
Beijing University of Technology *Beijing, China*
Supported instruction and labs for 80 M.S. students annually.

Company - Speech & NLP **2021-2025**
Mobvoi *Beijing, China*
Mentored 3 interns on speech/NLP projects (annual program).

Publications

Proceedings

- Xinsheng W, **Xiaoqin F**, et al. *Spark-TTS: An Efficient LLM-Based Text-to-Speech Model with Single-Stream Decoupled Speech Tokens.* arXiv preprint, 2025. [PDF]
- Dongsheng W, **Xiaoqin F**, Zeming L, et al. *2M-NER: Contrastive Learning for Multilingual and Multimodal NER with Language and Modality Fusion.* Applied Intelligence, 2024. [PDF]
- Wenjiang C, **Xiaoqin F**(† equal contribution), Yunlin C, et al. *Multi-granularity Semantic and Acoustic Stress Prediction for Expressive TTS.* APSIPA ASC, 2023. [PDF]
- Jipeng Z, **Xiaoqin F**. *Prosody Prediction with Discriminative Representation Method.* PRML, 2022. [PDF]
- **Xiaoqin F**, Xie R, et al. *Population Statistics Algorithm Based on MobileNet.* Journal of Physics: Conference Series (JPCS), 2019. [PDF]
- Zhiwei W, **Xiaoqin F**, et al. *Deep Knowledge Tracing with Side Information.* AIED, 2019. [PDF]
- Xie Rong, **Xiaoqin F**. *A Method of Quick Edge Detection Based on Zynq.* CCIOT, 2018. [PDF]
- Sheng J, **Xiaoqin F**, et al. *Research on the Internet of Things Platform for Smart and Environmental Protection.* CCIS, 2018. [PDF]

Patents

- all relevant to speech/NLP/LLM areas.

  main 6 First inventor: CN111078898A CN110970013A CN111145724A CN115470351A CN115470350A CN115547289A
- **other 4 Co-inventor**: CN111079428A CN111178042A CN115578998A CN116013251A

Personal Website

- **Website:** xqfeng-josie.github.io
- Contains selected projects (agents, speech/LLMs), publications, demos, and updates.