Xiaoqin Feng

xqfeng-josie.github.io | xiaoqinf@usc.edu | +1(213)636-5366 Los Angeles, CA

Curriculum Vitae

Introduction

Research Interests: Generative AI, LLM-powered agents, human-AI collaboration, and multimodal modeling (speech/text). I am a Master's student at the University of Southern California (USC) focused on building reliable, deployable agentic systems, with experience spanning speech and multimodal modeling, data pipelines, evaluation, and production ML systems.

Education

M.S. in Artificial Intelligence **Sep. 2025 – Present**
University of Southern California (USC) *Los Angeles, CA, USA*

M.S. in Software Engineering **Sep. 2016 – May. 2019**
Beijing University of Technology BJUT; GPA: 3.4 *Beijing, China*

B.E. in Computer Science **Sep. 2012 – Jul. 2016**
Southwest Minzu University SMU; GPA: 3.66, "Top 5" of 154 students *Chengdu, China*

Professional Experience

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

- **Consultant**: Optimized production speech models and LLM-based multi-agent workflows; improved offline evaluation, data pipelines, and deployment reliability.

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

- **Agents**: Built and deployed a production-grade LLM-powered agent framework for speech-centric applications (tool calling, multi-agent coordination, safety guardrails, runtime monitoring) [www.moyin.com].
- **LLMs & Audio**: Integrated LLMs with speech/audio pipelines (ASR, TTS, prompt generation) to ship AI-driven product features in production [www.dupdub.com].
- **Data Engineering**: Designed and maintained scalable multimodal data pipelines (text, audio, image) for model training and evaluation, including data collection, versioning, and automated quality checks at large scale.
- **Evaluation**: Developed an end-to-end evaluation framework for multimodal generation, combining offline benchmarks and regression testing to ensure model quality and production stability.
- **Collaboration**: Partnered with product managers, researchers, and engineers to turn AI prototypes into production systems; mentored junior engineers and interns.

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

- **NLP/Speech Synthesus:** Owned and optimized core speech algorithms (multilingual TTS, text normalization, G2P, prosody prediction, emotion analysis, rhythm prediction) for both foundational NLP components and product-facing speech features.
- **Backend:** Designed and built production speech backend services, exposing scalable TTS/NLP APIs that powered a consumer AI product with tens of millions of users.
- **Services:** Improved reliability of large-scale speech services to 99% availability through deployment hardening, monitoring, performance optimization, and production incident response.
- **Collaboration:** Mentored 3 interns on Speech/NLP projects (annual program).

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

- **Graph-based Learning**: Developed a Deep Knowledge Tracing (DKT) pipeline using graph-based embeddings and distance metrics to model and evaluate student knowledge progression.
- **Research Outcome**: Co-authored a paper accepted at *AIED 2019*; received the Outstanding Intern award for research contributions.

Projects

AI-Powered Audio Visualization Platform [QinZhi Creative] **Nov. 2025 – Present**

- Developed and launched a web-based audio processing platform that utilizes AI algorithms to generate high-fidelity, dynamic visualizations from speech and subtitles.
- Architected a scalable cloud infrastructure on Google Cloud Platform (GCP); implemented serverless backend services to handle audio rendering and media processing pipelines with high availability.
- Engineered an end-to-end dubbing workflow, integrating speech API with an interactive frontend to provide professional-grade audio editing and visualization tools for creators.

Skills

- **Languages**: Python, C++, Bash, SQL
- **Backend / Systems**: backend services, APIs, distributed systems, AI model serving, monitoring, CI/CD
- **Data Engineering**: large-scale datasets, dataset construction, data quality and versioning.
- **AI / ML**: LLM integration, agent systems, speech & NLP, multimodal learning, benchmark evaluation.
- **Dev/ML Tools**: Git, Docker, Linux, Kubernetes, cloud infrastructure; PyTorch, TensorFlow, vLLM

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.

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

Personal Website

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