

Kaitlyn Chen

kaitlynchen@gmail.com • (832) 907-6777 • github.com/kcitlyn • linkedin.com/in/kaitlynchen

EDUCATION

University of Texas at Austin, Cockrell School of Engineering

December 2028

B.S. in Electrical and Computer Engineering, Engineering Honors (Computer Architecture and Embedded Systems)

GPA: 3.8

RELEVANT EXPERIENCE

Amazon (AWS): Software Engineer Intern, GenAI Developer Tools

Summer 2026

Tech: TypeScript, Java, MCP, AWS Lambda, CloudWatch

Seattle, WA

- Architected a **self-augmenting agent** runtime that removes the fixed-at-startup tool-set constraint bounding today's LLM agents, letting a coding agent install a capability and use it mid-session with no restart, via a proxy that spawns and lifecycle-manages child MCP servers as live subprocesses with namespaced routing, handshake timeouts, crash detection, and automatic cleanup.
- Independently designed, built, and validated an end-to-end **agentic-AI toolchain** that lets generative-AI coding agents (e.g. Claude Code) autonomously discover, install, and operate tools from a ~4,000-server enterprise AI registry, transforming a human-only CLI into a self-service capability layer for **LLM agents** (TypeScript, 3 services, ~200 automated tests).
- Engineered **context-budget-aware tool** retrieval over the ~4,000-server space so an agent selects relevant capabilities without flooding its context window, using a **BM25** and fuzzy-matching index with relevance ranking, **TTL caching**, and harness-tuned result caps that serve warm queries in under 10ms versus a ~2s cold build.
- Hardened the agent against prompt injection and tool poisoning through untrusted-output handling, input allowlisting, shell-free execution, a non-repudiation audit log, and least-privilege tool scoping that limits each bundled dependency to an allowlisted subset of tools, designed under a generative-AI threat model.
- Designed a **serverless telemetry pipeline** (AWS Lambda and CloudWatch metrics) to measure how AI agents adopt and use the tooling, engineering a lazy, coalescing in-process buffer with a fail-safe flush so observability can never add latency to or break a tool call.

Texas EcoCar — Connected and Automated Vehicles (CAV) Team

Spring 2026

Tech: C/C++, MATLAB/Simulink, Python, Gaussian Process Regression, Machine Learning, Controls

Austin, TX

- Engineered a data-driven EPS steering controller using **Gaussian Process regression** to replace a gain-scheduled PID, eliminating manual calibration while matching lane-tracking performance (0.23 m RMSE); trained plant model on 1,500 samples in 26 seconds with 0.0004 rad prediction accuracy
- Designed real-time fault detection system using **GP prediction variance** as an autonomous confidence metric; detected out-of-envelope operation with 11,600x variance increase, triggering automatic safety fallback to proportional control
- Built system identification pipeline generating 678,000 data points across 24 operating conditions (6 speeds × 4 maneuver profiles); validated GP uncertainty bounds with **100% coverage within 2σ**

Texas Guadalupe — Hyperloop Engineering

Spring 2026

Tech: C, STM32, ADC, DMA, UART, Git

Austin, TX

- Developed multi-channel **STM32** sensor acquisition system using ADC with **DMA circular buffering** for real-time Hall effect speed measurement; implemented voltage-to-Gauss calibration pipeline derived from datasheet characterization
- Documented pod-wide embedded systems architecture defining 9 distributed STM32 controller nodes, CAN bus message routing topology, and LoRa telemetry link to ground station; produced system hierarchy diagrams and subsystem interface specifications for team-wide reference

PROJECTS

TrainYourFace — Offline Facial Recognition Tool (Python, OpenCV, dlib)

January 2025

- Built offline facial recognition system in Python (OpenCV + dlib) computing 128-dimensional ResNet face embeddings with Euclidean distance matching for real-time multi-face identification; designed for deployment in offline environments without cloud dependencies
- Implemented persistent embedding store and configurable threshold sweep to evaluate match sensitivity; tested across distance thresholds to balance false acceptance vs. missed identification for per-environment deployment tuning

PolyScribe — Offline STT + Translation (Python, Vosk, Argos, pyttsx3)

February 2025

- Built fully offline speech-to-text and translation pipeline using Vosk STT and Argos NMT with model-agnostic initialization — users download any supported language model and size, and the pipeline handles loading, configuration, and runtime setup automatically; published as open-source tool with external feature requests
- *Implemented a real-time multithreaded producer/consumer pipeline (audio → Vosk STT → Argos translation) with TTS playback, structured exports, and fault recovery for device/model failures*

Astrarium (HackTX 2025) — LLM-Orchestrated Backend (FastAPI, PostgreSQL)

October 2025

- Built FastAPI + PostgreSQL backend serving LLM-generated practice questions with structured output validation; implemented timeout/fallback guards and structured logging to maintain stability under unreliable LLM response latency

AWARDS & TECHNICAL SKILLS

- **AI/Agentic:** MCP (Model Context Protocol) • LLM agents • prompt engineering • evals/benchmarking • structured-output validation • RAG
- **Embedded/Hardware:** STM32 • ARM Cortex-M0 • CAN bus • I2C/SPI/UART • ADC/DMA • real-time systems • FSM design • timers/interrupts • GPIO • PCB layout • datasheet bring-up • Raspberry Pi • MATLAB/Simulink • KiCad • LTSpice
- **ML/Perception:** OpenCV • Gaussian Process regression • real-time video inference • system identification • dlib • ResNet • Vosk (STT) • Argos (NMT)
- **Languages:** C/C++ • Python • MATLAB • JavaScript/TypeScript • Assembly (ARM Cortex-M0, LC-3) • LaTeX • Mandarin (conversational)
- **Systems/Software:** Linux • Git • GDB • CMake • Valgrind • concurrency/multithreading • FastAPI • PostgreSQL • REST APIs • [Next.js](#) • logic analyzer • oscilloscope
- **Awards:** NCWIT Aspirations in Computing, National Honorable Mention & Houston Affiliate Winner • Engineering Honors Scholarship • National First-Gen Recognition