

# ZILING ZHANG

Gen AI Architect / Data Scientist / Robotacist

@ zhang.ziling@gmail.com

720-939-9292

8792 Flattop St, Arvada, CO 80007

zilingz

US Citizen

## EXPERIENCE

### Advisory Data Scientist

Automation Technical Implementation Group | Ricoh Graphic Communications

Aug 2022 - present

Boulder, CO

- Lead AI Coding Assist pilot with Claude Code for 280+ builders. Architect AWS Bedrock cloud infrastructure with SSO auth, per-user cost tracking, OpenTelemetry observability.
- Architect enterprise Gen AI transformation. Built Grace AI platform serving 280+ builders with Claude on AWS Bedrock, Docker microservices (Open WebUI, LiteLLM, vLLM), MCP tool integrations (Jira, Confluence, GitLab).
- Manage NVIDIA GH200x4 GPU cluster serving local LLMs, embedding, reranking, and image generation models, powering enterprise AI agents across multiple departments.
- Manage NVIDIA RTX Pro 6000 Blackwell cluster for simulation-based robotics and computer vision software CI pipelines.
- Prototype Universal Robot (UR20) book palletizer. ROS2, computer vision, Cognex barcode scanner, FastAPI backend, React/Three.js digital twin (Cobot UI).
- Hands-on expertise across NVIDIA's three-computer paradigm: GH200 (datacenter AI), RTX Pro 6000 Blackwell (Omniverse simulation), Jetson (edge robotics).
- Colorado School of Mines field session: Nvidia Cosmos robotic simulation data generation, Omniverse ray traced printshop digital twin.
- Design Agentic Coding training programs. Hackathon judge, CU Boulder.
- Develop serverless ETL pipeline aggregating printer/sensor timeseries data. AWS Athena, Lambda. Supporting predictive maintenance inference with SageMaker serverless.

### Staff/Advisory Imaging Science Engineer

Boulder Development Section | Ricoh Graphic Communications

Aug 2019 - Aug 2022

Boulder, CO

Computer Vision + Automation R&D for Continuous Feed Industrial Printer

- Detect nozzle defect by training a semantic segmentation Deeplab v3+ neural network, accelerate with Nvidia TensorRT int8 quantization, achieve 37x speedup.
- Prototype nozzle uniformity algorithms, accelerate with SIMD and threading, achieve 40x speedup. Intel MKL/TBB, OpenCV, Visual Profiler, Concurrency Visualizer.
- Architect image processing pipeline that can withstand 7GBytes/s scanner data, perform realtime defect detection and compensation to achieve superior print quality.
- Standardize the team's coding practice, enforce modern C++ RAII style, C++ Core Guidelines, Google python style / pylint, modularize code into testable liblets.
- Setup CMake, google test, GitLab CI, docker runners for high velocity iteration.
- Filed 28 patents, granted 10, achieved Master Inventor title, fastest in Ricoh Boulder.

### Robotics Research Assistant

Human Centered Robotics Lab | Colorado School of Mines

Aug 2018 - Dec 2019

Golden, CO

- DARPA subteranean challenge Mines team, SLAM LiDAR 3D mapping ROS node.
- Toyota InfoTechnology Center, V2X IoT multi-vehicle collaborative perception.
- Metcalfe Archaeology, long term place recognition AR iOS app (Swift, OpenCV C++).
- Kinect + FLIR lipton sensor fusion, forehead temperature measurement TurtleBot.

## SKILLS

### Gen AI Stack

Claude Code Claude Agent SDK  
Bedrock vLLM LiteLLM

### Accelerated Compute

CUDA TensorRT Warp  
CuTile MKL TBB SIMD

### ROS Stack

Omniverse OpenCV Isaac Sim  
Newton Jetson Python  
C++20

### Cloud Stack

AWS Serverless SageMaker

### Frontend Stack

React TypeScript Three.js

## EDUCATION

### M.S. Computer Science

Colorado School of Mines

Aug 2018 - Dec 2019

- Robotics, HPC ★ GPA 3.96
- Courses: Computer Vision, Parallel Computing, Robotics, Computer Architecture
- Paper: Collaborative Localization for Occluded Objects in Connected Vehicular Platform  
4th author, 2019 IEEE 90th Vehicular Technology Conference

### B.S. Applied Physics

University of Science and Technology of China

Sep 2007 - June 2011

- Condensed Matter Physics
- Paper: Speed of Last Vehicle Feedback Strategy in Intelligent Transportation Systems  
3rd author, International Journal of Modern Physics C, 2011

## Scientific Translator (English -> Chinese)

### Freelance Contracts

📅 Jan 2012 – Aug 2018

📍 Remote

💻 (NetEase Inc.) Translate OpenCourseWare, Coursera lectures subtitles, 1.87 million words.

📖 (Yeeyan.org) Book translation, Net Smart by Howard Rheingold, MIT press.

---

## Software Developer

### JindianTuteng Talent information solution Inc.

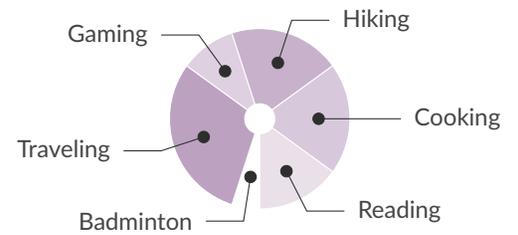
📅 Oct 2011 – June 2012

📍 Guangzhou, China

💻 Rule based transaction fraud detection for Foreign Exchange Management.  
Java, J2EE, JSP, IBM WebSphere, IBM DB2.

## INTERESTS

---



## CONFERENCES

---

- 📅 SIGGRAPH 2024
- 📅 NVIDIA GTC 2025
- 📅 NVIDIA GTC 2026
- 📅 SIGGRAPH 2026