

Jinda Zhang

(236)-965-3670 | jindaz.personal@gmail.com | jindaz.github.io | [Github](#) | [Linkedin](#) | Burnaby, BC

OBJECTIVE

Experienced Software Engineer 3 years professional experience seeking Software and Machine Learning opportunities.

SKILLS

Python (Flask), Java (SpringBoot, JDBC, Redis), C++, R, Distributed Systems (Microservice, Kubernetes, Docker, REST API, AWS, Google Cloud Platform, Azure), Devops (Slurm, Bash, Linux), Web (JavaScript, TypeScript, NodeJS, Vue, React, ReactNative), Database (MongoDB, SQL), Git, AI (Speech, NLP, RAG, Agent, LLM)

EDUCATION

- **Northeastern University** Sept 24
Master in Computing Vancouver, Canada

EXPERIENCE

- **Software Developer Intern, Machine Learning** Sept 25 - Present
Kinaxis Remote, Canada
 - Developed and maintained **Distributed RAG system** in dev environments using **Docker, K8s, Helm, PostgreSQL**.
 - Explored **LLM** integration for **agent** personalization; implemented **prompt engineering, memory** injection.
 - Configured authentication and secrets management for backend services by integrating **Azure AD (OAuth2), Vault, and CosmosDB/Azure Blob Storage**, ensuring secure token exchange and API access.
- **R&D Engineer Intern** May 25 – Aug 25
Ansys, Synopsys Vancouver, Canada
 - Optimized **distributed HPC benchmark pipelines** using **C++, Python, Bash, and SLURM**
 - Identified and visualized 3 scalability bottlenecks in **parallel computing**, improving performance by **~200%**.
 - Built **cross-platform monitoring tools** with **FastAPI** and **Streamlit** to track **parallel computing performances**
- **Research Assistant (Machine Learning Engineer)** Oct 23 - Oct 25
SLangLab (Contract) Vancouver, Canada
 - Architected distributed **LLM**-based ASR pipeline with transfer learning optimization across **speech** domains
 - Developed a **multimodal** LLM system utilizing HPC clusters, achieving **4x** faster training and inference throughput
 - Proposed a novel **augmentative communication architecture** integrating **Flask, React.js, Docker, and S3**
- **Researcher** May 25 - Sept 25
PathOnAI (non-profit) Remote, Canada
 - Built unified **web agent** evaluation framework supporting 5+ models with standardized trajectory processing
 - Implemented hybrid evaluation system with text validation, screenshot analysis, and **graph**-based comparison
 - Developed production-ready testing pipeline with **Streamlit** dashboard, automated scripts
- **Software Engineer Intern** Dec 24 - Mar 25
One Community Global (non-profit) Remote, Canada
 - Built dynamic UI modules for a web dashboard using **React.js** and **Node.js RESTful APIs** for project tracking.
 - Implemented **50+ unit tests** with **Jest**, boosting code coverage to **92%**; ensured quality with **Agile Scrum** sprints.
- **Research Assistant (Software Engineer)** Jan 24 - Apr 24, Mar 25 - Sept 25
Human-AI Collaboration Lab (Contract) Vancouver, Canada
 - Designed **Agent** using **MongoDB**, supporting **10K+** contextual interactions for conversational AI agents in **VR**.
 - Integrated **3** major APIs into an LLM-powered **RAG** system, improving retrieval latency by **35%**.
 - Achieving human-agent communication through cloud-based LLM integration and real-time Speech-to-Speech API
- **Research Assistant (Software Engineer)** Nov 24 - Mar 25
MOD Squad Lab, Advisor: Prof. Yvonne Coady Victoria, Canada
 - Conducted **RAG evaluation** using **RAGAs** on **10** evaluation metrics, identifying key performance trends in LLM
 - Integrated **vLLM** for **scalable inference**, achieving **200%** speedup on long-context tasks
- **Freelance Software Engineer** Summer 22
UDEVU (Contract) Remote
 - Designed and developed engaging 3D web applications using **Three.JS** for client: [Josh Struve \(Udevu.dev\)](#).
 - Provided consultative support on **search engine optimization**, optimizing web presence for better visibility.

PROJECTS

• Digital Business Management Platform

Tools: *Spring Boot, Spring MVC, MyBatis, Redis, JWT, Vue.js, Nginx, MySQL, WebSocket, Maven, Swagger, Apache POI, AWS* [G]

- Developed full-stack enterprise platform using **SpringBoot**, **MyBatis**, **Redis** with **JWT** authentication, **RESTful** APIs, and real-time analytics serving 10+ business modules
- Implemented microservices architecture with **Nginx** load balancing, payment integration, automated Excel reporting, and cloud storage using **AWS**

• Voice Collector

Tools: *Flask, React.js, Docker, AWS S3* [G]

- Developed robust **Speech Recognition** and Voice Conversion tool, utilizing **Flask** and **React.js**, to capture valuable audio data from users with unique speech patterns, promoting accessibility and inclusivity in speech technology.
- **Containerized** backend functionalities using **Docker**, orchestrated with **Docker Compose** for easy deployment and management, incorporated **Minio** for secure storage of audio recordings with **AWS S3** compatibility.

• Athena AI Research Assistant

Tools: *Typescript, Javascript, Node, Express, React* [G]

- Built cross-platform AI-powered desktop application using **Electron**, **React**, and **TypeScript** with **Gemini API**
- Architected event-driven system with IPC communication, **real-time processing**, and **asynchronous** AI processing utilizing **React Query state management**, singleton design patterns, queue-based solution for multi-modal AI

Side Projects: [Financial Analyzer](#) (NLP, OCR), [Firelook](#) (Android), [Around](#) (GoLang, ElasticSearch), [Ticket](#) (Java, EC2, MySQL), [Titan](#) (MapReduce, MongoDB), [Shiny-Instagram](#) (Java), [Sparkling-Memos](#) (Java), [HairSalon](#) (JS)

AWARDS

University of Victoria Graduate Entrance Award, UVic Research Fellowship, Khoury SEED Grant (FA24), Research Assistant (SU24), Teaching Assistant (SU24), West Coast Research Grant (SP24), Khoury Research Apprenticeship (SP24, Top 1%), Research Assistant (23), Teaching Assistant (23), Teaching Assistant (19)

OUTREACH

[Inside the Spring 2024 Khoury Research Apprenticeship Showcase](#)

June 24

[Northeastern University Researchers Present Research Findings, Receive Honors at ACM CHI](#)

May 24

PUBLICATIONS

Multimodal LLMs for Speech Processing

- J Zhang, A Mohan. [Towards atypical speech transcription using LLM-based ASR](#) (Interspeech 25 Oral)
- M Hui, J Zhang, A Mohan (2025). [Enhancing AAC Software for Dysarthric Speakers in e-Health Settings: An Evaluation Using TORGO](#) (IEEE ICC'25, *Responsible for experiments, analysis*)

Agentic AI systems and applications

- Yaoyao Qian, Yuanli Wang, Jinda Zhang, Yun Zong, Meixu Chen, Yifan Zeng, Jindan Huang, Danqing Zhang (2025). [WebGraphEval: Multi-Turn Trajectory Evaluation for Web Agents using Graph Representation](#). *MTI-LLM @ NeurIPS 2025*.
- H Wan, J Zhang, AA Suria, B Yao, D Wang, Y Coady, M Prpa (2024). [Building LLM-based AI Agents in Social Virtual Reality](#). In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI LBW 24)*
- Yaoyao Qian, Yuanli Wang, Jinda Zhang, Yun Zong, Meixu Chen, Yifan Zeng, Jindan Huang, Danqing Zhang (2025). [WebGraphEval: Multi-Turn Trajectory Evaluation for Web Agents using Graph Representation](#). *Submitted to AAAI 2026 Demonstration*.
- M Pan, P Liu, J Zhang, B Yao, D Wang, M Prpa. [Reducing Foreign Language Speaking Anxiety with an LLM-based Embodied Conversational Agent in Social VR: A Longitudinal Study](#). *Submitted to CHI 26*.
- D Jacoby, K Smith, N Inbar, C Farr, J Wong, S Khutorna, T Fang, Z Li, J Zhang, and Y Coady. [Tech Tales: Digital Storytelling with the Royal BC Museum](#). *Appy Hour, SIGGRAPH 2025*.
- (Early Works) Jinda Zhang. [Enhancing Tree Type Detection in Forest Fire Risk Assessment: Multi-Stage Approach and Color Encoding with Forest Fire Risk Evaluation Framework for UAV Imagery](#). *arXiv 2407.19184*.