Jiuzhou Fu

📱 +1 (206) 123-4567
✉ jiuzhou@uw.edu
🌐 alpha3-3.github.io
⭕ Alpha3-3

Research Interests

Applied Mathematics, Vector Databases, Information Retrieval, Large Language Models (LLMs), Natural Language Processing (NLP), Machine Learning Systems, Artificial Intelligence (AI).

Education

2022–Present   **B.S. in Mathematics**, *University of Washington*, Seattle, WA
Coursework includes: Two-year Honors Math Sequence (Calculus, Differential Equations, Real & Complex Analysis), Abstract Algebra, Topology & Geometry, Probability, DBMS, AI, and Machine Learning.

Publications

ICDM 2025   **Jiuzhou Fu**, Luanzheng Guo, Nathan Tallent, Dongfang Zhao. *ProHD: Projection-Based Hausdorff Distance Approximation*. Oral Presentation at the **IEEE International Conference on Data Mining (ICDM) 2025** (top-tier data mining conference, 13% acceptance rate for full research paper). Link to be added

Under review   **Jiuzhou Fu**, Dongfang Zhao. *QPAD: Quantile-Preserving Approximate Dimension Reduction for Nearest Neighbors Preservation in High-Dimensional Vector Search*. Under review at **IEEE International Conference on Data Engineering (ICDE) 2026**. https://arxiv.org/abs/2504.16335

Under review   Yuhan Huang, Xiaoxi Hu, **Jiuzhou Fu**, Fei Chen, Jingming Cao, Naibang Wang. *SSIDN: A Small-Sample Incremental Diagnosis Network for Railway Turnout Actuators*. Under review at **IEEE Instrumentation and Measurement Technology Conference (i2MTC) 2026** (The flagship conference of the IEEE Instrumentation and Measurement Society). link to be added

Award

ICDM 2025 Student Travel Award (Funded by NSF)

Research

Nov 2024 – Present   **Undergraduate Researcher**
High-Performance Data-Intelligence Computing (HPDIC) Lab, *University of Washington*, Seattle, Principal: Prof. Dongfang Zhao

- **ProHD: Projection-Based Hausdorff Distance Approximation (ICDM 2025)** — proposed a projection-guided subset-selection algorithm combining centroid intersections and principal components to approximate the Hausdorff distance up to 10–$100\times$ faster with bounded error; work supported by the **U.S. Department of Energy's Pacific Northwest National Laboratory (PNNL)**, Texas A&M ACES cluster (NSF ACCESS), and Chameleon testbed; presented in Washington, D.C.

- **QPAD: Quantile-Preserving Approximate Dimension Reduction for Nearest Neighbors Preservation in High-Dimensional Vector Search (under ICDE review)** — developed an unsupervised dimensionality-reduction method maximizing neighborhood margin to preserve approximate nearest-neighbor (ANN) structure, with formal proofs of boundedness, compactness, and indexability in ANN systems.

- **Directed Hypergraph Compression and Reordering (under preparation)** — extended a state-of-the-art undirected hypergraph reordering and compression algorithm to directed hypergraphs, enabling dynamic maintenance, reduced L1 cache misses, task speedups, and a smaller storage footprint for large-scale graph systems.

- **Techniques and Stack: Linux/VM** environment; **C/C++** (*STL, OpenMP, Makefile*); **Python** (*multiprocessing, psutil, pandas, NumPy, SciPy*); *FAISS* (ANN search); *matplotlib* (visualization); *mpi4py* and *PyOpenCL* (parallel computation); *Conda* (environment management).

Mar 2025 – **Paid Research Intern**
Present **School of Vehicle and Mobility**, *Tsinghua University*, Beijing

- ○ **SSIDN: A Small-Sample Incremental Diagnosis Network for Railway Turnout Actuators** - Conducted research on **sensor error injection and fault diagnosis** for time-series data from railway switch machines (RSM) and industrial robots, constructing datasets simulating sensor abnormalities (open circuit, periodic interference, impulse noise, etc.) to evaluate model robustness. The dataset supports training of the proposed network for end-to-end classification under realistic sensor failures.
- ○ Implemented data preprocessing and **error-injection pipelines** in **Python** (*NumPy*, *pandas*, *PyTorch*) for vibration and control-signal time series; performed ablation studies and feature visualizations. One more co-authored papers on fault diagnosis with sensor reliability modeling is in preparation.

Dec 2023 – **Undergraduate Researcher**
Dec 2024 **Washington Experimental Mathematics Lab (WXML)**, *University of Washington*, Seattle

- ○ Conducted theoretical and computational research on the **Concentration of Measure** phenomenon via the Curie–Weiss model and high-dimensional spheres, deriving analytic asymptotics and probabilistic limits, and validating them through **Monte Carlo simulations** and symbolic computation in *Wolfram Mathematica*.
- ○ Presented the project poster at the **UW Undergraduate Research Symposium** as part of the **Undergraduate Research Program**, highlighting rigorous mathematical reasoning combined with computer-aided experimentation. Project link

Internship

Jun – Aug **Prompt Engineer Intern, MaaS Department**, *Z.AI*, Beijing, China
2025
- ○ Worked on the **REDNOTE** (Xiaohongshu overseas version) AI Translation System Team, focusing on multilingual prompt design, retrieval-augmented generation (**RAG**) pipelines, and translation strategy development across cross-lingual domains.
- ○ Designed and deployed **Dify-based prompt-engineering workflows**, supporting **instruction fine-tuning** of multimodal LLMs (GLM-4v/4.5v) for contract understanding from PDF/image inputs and structured JSON output generation.
- ○ Designed the **PRD** and **MVP** for an LLM prompt-sharing website; implemented workflows for prompt-to-HTML rendering, multi-speaker podcast generation, and TTS stress testing for production readiness.

May – Aug **Data Science Intern**, *Real AI*, Beijing, China
2024
- ○ Processed large-scale unstructured web data using **Python** (*pandas*, *NumPy*) and *SQL*, including automated data-cleaning pipelines for schema standardization, null-value imputation, and cross-table validation.
- ○ Helped design and deploy a multi-turn prompt-engineering and retrieval pipeline for an enterprise policy-advising LLM, integrating natural-language understanding, semantic retrieval, and consistency checks to enhance factual reliability and reduce hallucination.

Aug – Oct **Cybersecurity Intern, Security BU**, *QAX (Qi An Xin Group)*, Beijing, China
2024
- ○ Participated as a **Blue-Team Cybersecurity Junior Specialist** in red-blue adversarial exercises, performing real-time threat detection, vulnerability assessment, and network forensics using QAX's defense platform.
- ○ Utilized *Burp Suite*, *sqlmap*, *AWVS*, and *Kali Linux* to identify injection, authentication, and configuration vulnerabilities; wrote technical reports and mitigation documentation for incident response.