

# Jason (Junjie) Zhu, Ph.D.

jasonjunjiezhu.com

Email: junjie.zhu.jason@gmail.com

Mobile: 650-285-7123

## SUMMARY

---

I am a curiosity-driven, scientifically trained builder with 10+ years of experience in AI/ML, statistics, and graph algorithms. I have had the fortune to collaborate with world-class researchers and top-tier product teams to drive meaningful, collective impact—reflected in 10,000+ citations to my publications. Passionate about complex challenges and high-agency environments, I architect and implement scalable solutions across emerging domains, from multi-modal RAGs and intelligent search to biomedical discovery.

## EDUCATION

---

### Stanford University

*Ph.D. in Electrical Engineering · M.S. in Statistics*

Stanford, CA

2014 – 2020

### Olin College of Engineering

*B.S. in Electrical and Computer Engineering*

Needham, MA

2010 – 2014

## EXPERIENCE

---

### Nexa AI

*Head of AI/ML*

Cupertino, CA

Feb 2025 – Present

- Leadership: Spearheading a small team to optimize Gen-AI edge inference across CPU/NPU/GPU platforms.
- Local RAGs: Pioneering privacy-first RAG systems with lightweight AI models and edge vision processing.
- Agentic Systems: Architecting next-generation AI applications using emerging protocols (e.g., MCP, A2A).

### Apple

*Machine Learning Engineer*

Cupertino, CA

Jan 2020 – Feb 2025

- Mentorship: Empowered engineers to achieve recognition through internal and external conference publications.
- Synthetic Data Generation: Pioneered robust frameworks using advanced dimensional perturbation techniques.
- Preference Learning: Designed offline A/B testing to capture preference dynamics and distribution shifts.
- Query Intent Testing: Built high-performance Java pipelines for Maps Search query understanding evaluation.
- Ranking Triage: Proposed and implemented linear-time algorithms to debug and improve multi-ranker systems.

### Stanford University

*Research Assistant*

Stanford, CA

Sep 2014 – Feb 2020

- Graph Visualization: Engineered interactive visualizations for Gene Ontology analysis.
- Statistical Inference: Enhanced multiple-hypothesis testing to eliminate data snooping bias.
- Unsupervised Learning: Developed novel dimension-reduction techniques for cellular modeling.
- Selective Inference: Implemented advanced methods for tissue-specific eQTL analysis.
- Sequence Alignment: Enhanced DNA sequence alignment performance through C/C++ optimization.

### Olin College of Engineering

*Research Assistant*

Needham, MA

Sep 2010 – May 2014

- Graph Theory: Resolved distance-2 graph coloring challenges for specialized graph families.
- Information Theory: Developed stochastic geometric models for wireless network interference.

## SELECTED PUBLICATIONS

---

1. Automatically Authoring Regression Tests for Machine-Learning-Based Systems. *ICSE*, 2021
2. Progenitor identification and SARS-CoV-2 infection in human distal lung organoids. *Nature*, 2020
3. Exploratory gene ontology analysis with interactive visualization. *Scientific Reports*, 2019
4. Visualization and analysis of sc-RNA-seq data by kernel-based similarity learning. *Nature Methods*, 2017

Full list shown on Google Scholar: <https://scholar.google.com/citations?user=2EasRdEAAAAJ&hl>