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 — Stanford, CA
Ph.D. in Electrical Engineering · M.S. in Statistics — *2014 – 2020*

Olin College of Engineering — Needham, MA
B.S. in Electrical and Computer Engineering — *2010 – 2014*

Experience

Nexa AI — Cupertino, CA
Head of AI/ML — *Feb 2025 – Present*

- **Semantic Search Innovation**: Invented a local file-search semantic search engine with structured metadata support and @-search features, enhancing real-time query resolution and usability for customer-facing demos.
- **From 0 to 1**: Led a 4-member team to build and ship an on-device RAG system in under 3 months, powered by continuous regression testing from day one to support rapid iteration and measurable weekly quality gains.

Apple — Cupertino, CA
Machine Learning Engineer — *Jan 2020 – Feb 2025*

- **Infrastructure Modernization**: Revamped internal testing pipelines for query understanding and ranking, cutting release cycles from weekly to daily and increasing launch reliability for WWDC-highlighted features.
- **Scalable Evaluation**: Designed generative and retrieval-based evaluation frameworks for industrial-scale search systems; methodologies and insights were shared at top-tier software engineering conferences (*ICSE*, *FSE*).
- **Technical Leadership**: Defined team roadmaps and halved manual triage overhead year-over-year, enabling peers to pursue novel evaluation directions and earn recognition at internal AI/ML conferences.

Stanford University — Stanford, CA
Graudate Research Assistant — *Sep 2014 – Feb 2020*

- **Full-Stack Data Science**: Built an interactive visualization tool integrating backend graph algorithms for 30,000+ Gene Ontology terms— that provides association discovery with rigorous false discovery rate control.
- **Scalable Graph Learning**: Advanced graph-based unsupervised learning pipelines for million-size high-dimensional data, leading to high-impact publications across *Nature*, *Nature Methods*, *Cell*, and *NeurIPS*.

Olin College of Engineering — Needham, MA
Undergraduate Researcher — *Sep 2010 – May 2014*

- **Graph Theory**: Resolved distance-2 graph coloring challenges for specialized graph families (5 publications).
- **Information Theory**: Developed stochastic geometric models for wireless network interference (3 publications).

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