Jason (Junjie) Zhu, Ph.D.

https://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 collaborated with world-class researchers and top-tier product teams to drive meaningful, collective impact—reflected in 10,000+ citations to my co-authored publications. Passionate about complex challenges and high-agency environments, I architect and implement scalable and reliable solutions with measurable impact across emerging domains, including multi-modal RAGs, intelligent search, and 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 semantic file-search engine with structured metadata and @-search support, enhancing real-time query resolution and usability in customer-facing demos.
- **From 0 to 1**: Led a 4-member team to develop and ship an on-device RAG system in under 3 months—powered by continuous regression testing and weekly iteration to accelerate quality improvements.

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

- **Infrastructure Modernization**: Revamped internal testing pipelines for query understanding and ranking, reducing release cycles from weekly to daily and improving launch stability for WWDC-featured products.
- **Scalable Evaluation**: Designed generative and retrieval-based frameworks to evaluate ML systems at scale; shared methodologies at top-tier software engineering conferences (*ICSE*, *FSE*).
- **Technical Leadership**: Defined roadmaps and reduced manual triage time by 50% year-over-year, enabling teammates to explore new evaluation strategies and gain recognition at internal AI/ML conferences.

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

- **Full-Stack Data Science**: Developed an interactive tool to visualize and perform power analysis on 30,000+ Gene Ontology terms—enabling large-scale association discovery with controlled false discovery rate.
- **Scalable Graph Learning**: Developed graph-based unsupervised learning pipelines for million-scale, high-dimensional datasets—resulting in publications in *Nature*, *Nature Methods*, *Cell*, and *NeurIPS*.

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

- **Theoretical Research**: Developed combinatorial algorithms for graph coloring and stochastic geometric models for wireless interference, leading to 8 publications—5 in discrete math journals and 3 in flagship IEEE conferences.

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