

# 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

- **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

*Machine Learning Engineer*

Cupertino, CA

Jan 2020 – Feb 2025

- **Modernized internal evaluation pipelines** for query understanding and ranking services to accelerate software deployment velocity (weekly to daily), enhancing the quality and stability of new features showcased at WWDC.
- **Developed both generative and retrieval-based methods** to evaluate ML systems at scale, applied them to industrial-scale search engines and shared the early results at top software engineering conferences (*ICSE*, *FSE*).
- **Built project roadmaps and reduced manual triaging time** by nearly 50% year over year, building capacity for colleagues to innovate on new evaluation initiatives and achieve recognition through internal AI/ML conferences.

### Stanford University

*Graduate Research Assistant*

Stanford, CA

Sep 2014 – Feb 2020

- **Full-Stack:** developed an interactive tool to visualize 30,000+ Gene Ontology terms and 70,000+ genes, enabling power analyses and simulations across high-throughput genomic data to control the false discovery rate.
- **Applied Research:** Built upon graph-based unsupervised learning methods to develop a scalable pipelines with million-scale sample-size data sets, resulting in publications in top scientific journals (*Nature*, *Nature Methods*, *Cell*, *NeurIPS*).

### Olin College of Engineering

*Undergraduate Researcher*

Needham, MA

Sep 2010 – May 2014

- **Graph Theory:** Discovered the optimal distance-2 coloring assignments (which applies to the radio frequency assignment problem) for various graph families with combinatorial algorithms and co-authored five publications in discrete math journals.
- **Information Theory:** Derived mathematical solutions for stochastic geometric models to analyze wireless network interference with multi-antenna systems, resulting in three first-author publications 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>