

Professional Summary

Senior AI/ML leader with 14+ years of experience building and deploying production AI systems across computer vision, NLP, and multimodal AI. Expert in search systems, generative AI, and content-based retrieval at scale, with proven track record in model development, dataset curation, and team leadership. Deep expertise in machine learning infrastructure, distributed systems, and full-stack development spanning mobile, web, and AI-powered products.

Professional Experience

Aviary 2024 – Present

Head of AI / ML

- Leading research and development of next generation video search and understanding technologies

Snap Inc. 2016 – 2024

Staff Machine Learning Engineer & Tech Lead Manager

- Launched Snapchat's first generative AI-powered visual enhancement system and led development of in-house video captioner
- Developed billion+ scale vector similarity search engine powering visual search and ML pipelines
- Led development of emoji-aware OCR service improving spam detection by 168% and reducing moderation costs by \$130k/yr
- Owned ML infrastructure track, defined hiring criteria, grew team from 0 to 10 engineers; reduced costs by \$2M+ annually

Amazon 2011 – 2016

Software Development Engineer II

- Developed product de-duplication system and led launch of 30+ new product categories

Education

Carnegie Mellon University 2009 – 2011
M.S. Computer Science - Language Technologies Pittsburgh, PA

University of Texas at Arlington 2004 – 2009
B.S. Software Engineering & Physics Arlington, TX

Technologies & Skills

- **AI/ML:** Multimodal AI, Computer Vision, NLP, LLMs, Agents, Information Retrieval
- **Infrastructure:** Python, Go, AWS, GCP, Vector DBs, Kubernetes, Ray, LangChain, LlamaIndex
- **Leadership:** Team Management, Technical Strategy, System Design, Technical Hiring & Recruitment

Select Publications & Talks (18 Papers, 11 Patents, h-index: 11, Citations: 900+)

- Multimodal Structured Extraction for Self-Querying Music Video Retrieval (Demo @ ISMIR 2024)
- Video Enriched Retrieval Augmented Generation using Aligned Video Captions (MRR @ SIGIR 2024)
- Multimodality-guided Image Style Transfer using Cross-modal GAN Inversion (WACV 2024)
- FORB: A Flat Object Retrieval Benchmark for Universal Image Embedding (NeurIPS 2023)
- Video matching with a messaging application (US Patent 11,823,456, 2023)
- Open-Vocabulary Object Detection Using Captions (CVPR 2021)
- MLOps at Snapchat: Continuous Machine Learning with Kubeflow & Spinnaker (Kubecon 2020)
- Build accurate training datasets with Amazon SageMaker Ground Truth (AWS re:Invent 2019)