Alex Kashi

Staff Software Engineer and **Machine Learning Researcher** with **10+ years** shipping scalable ML systems and mobile apps. Founded Spur a social app with **350k+ downloads**. Led teams across engineering, product, marketing and design.

alexkashi.com | akashi@berkeley.edu | linkedin.com/in/alexkashi | github.com/alexkashi

Education

Harvard University — **GPA: 3.9**
M.S. Computational Science and Engineering — *Cambridge, MA*

University of California Berkeley — **High Honors (Top 10%) - GPA: 3.9**
B.S. Electrical Engineering and Computer Sciences — *Berkeley, CA*

Experience

Lucid Motors — May 2022 – Present
Staff Engineer — iOS & ML Perception | Swift, SwiftUI, Python, PyTorch, C++ — *Newark, CA*

- Led rewrite from Flutter to **SwiftUI**, integrating native features like **3D Models (Scene Kit)**, Live Activities, Siri, widgets, and Apple Watch support to enable **vehicle access, monitoring, and remote control** via mobile.
- **Reported directly** to the **Senior VP of Digital** to align development with strategic initiatives.
- Developed a **SwiftUI component toolkit** and frameworks for **testing, localization, persistence, and networking**, while authoring an **MVVM architecture usage guide** to enhance productivity and streamline onboarding.
- Led weekly **architecture and code reviews**, mentoring iOS engineers and guiding Android porting to **Kotlin**.
- Designed a **Bluetooth passive entry system**, improving reliability to **99%**, surpassing the 95% industry standard.
- Built an **H.264 WebRTC client** to stream live video from the car's cameras to a phone for security monitoring.
- Developed a parking slot detection system by fusing **four surround-view cameras** with a **multi-task transformer**.
- Built and integrated **polygon-based object detection** and **semantic segmentation** heads onto the **transformer**.
- Conducted **ablation studies** to optimize architecture, boosting **F1** from **0.89 to 0.96** and speed by **41% (16 FPS)**.
- Curated real and synthetic datasets with **Sim2Real** techniques to improve generalization across conditions.
- Exported and deployed model to production using **ONNX and TensorRT** on **Nvidia Jetson Orin**.

Harvard Medical School — May 2023 – Aug 2023
Machine Learning Research Scientist | Python, PyTorch — *Cambridge, MA*

- Adapted **AlphaFold 2** a protein prediction model for RNA structure prediction, boosting F1-score from **0.47 to 0.52**.
- Developed **synthetic data training algorithms**, significantly **reducing** the **generalization gap** between RNA types.

Spur — Jan 2019 – Dec 2022
Founder CEO/CTO | Swift, Kotlin, GraphQL, Javascript, Lambda, DynamoDB, Elasticsearch — *Sunnyvale, CA*

- Founded and developed a **native social media app for iOS and Android** that garnered over **350k downloads**, blending Instagram and Tinder features with real-time chat, customizable profiles, matches, likes, and IG stories.
- Admitted into the **Harvard Innovation Labs Venture Program** and selected for the semester-long **Harvard Business School accelerator** course, Field X.
- Crafted and executed a comprehensive **go-to-market strategy, pitch deck, and business plan**, which were presented to **angel investors and venture capitalists** to secure **pre-seed funding**.
- Engineered a **microservice-based backend** utilizing **AWS** technologies including **DynamoDB, Elasticsearch, S3, Lambda, EC2, Kinesis, and Rekognition**.
- Directed a multidisciplinary **team of over five professionals**, in **marketing, competitive analysis, and design.**

Stanford — Sep 2017 – Sep 2021
Research Associate | Python, TensorFlow, C — *Stanford, CA*

- Developed a **ML-based diagnostic model** for Myalgic Encephalomyelitis (ME) using impedance-over-time data from a novel biosensor; results **published in *PNAS***.
- Authored a **statistical theory** on the origin of ME based on Hardy-Weinberg Equilibrium, published in ***Diagnostics***.
- Led a team on **cell semantic segmentation, cutting computation time by 94%** and **boosting accuracy by 4%**.
- **Mentored** computer vision and machine learning **interns**, supported by the **Stanford Bio-X program**.

Intel **May 2017 – Sep 2017**

Software Engineer - R & D | Java, Android, C, C++ *Santa Clara, CA*

- Developed an Android NDK **MapReduce** framework using pthreads and C++ templates.
- Applied it by implementing **k-means clustering** on location data for network data reduction.

Early Engineering Experience **2014 – 2017**

Software Engineer Intern - Intel, Zspace, Mux Wiring | Java, Android, C/C++ *Santa Clara, CA*

- Enhanced **I2C and UART drivers**, extending support to **touch screens and cameras** (Intel).
- Built **gaze-tracking** system using **deep learning and CNNs** achieving **93%** accuracy (Zspace)
- Built an **interpreter** translating human-readable code to machine instructions (Mux Wiring).

Projects

Dubbles.ai | Python, PyTorch, Swift, JavaScript, Next.js **Apr 2024**

- Co-founded and shipped **AI-powered text-to-video app** with precise **lip and voice synthesis** of celebrities.
- Built a **serverless inference pipeline** with content moderation on replicate for fast, cost-efficient AI video generation.

Microsoft Capstone: Socioeconomic Future of Kenya | Python, PyTorch **Dec 2022**

- Analyzed Kenya's socioeconomic indicators with **Bayesian models**, including fertility, education, and GDP.
- Presented an interactive **dashboard to Microsoft** on population and industrialization trends, aiding strategic decisions.

Compressing and Accelerating Stable Diffusion | Python, PyTorch **Dec 2022**

- Accelerated **Stable Diffusion**, achieving an **80% memory reduction** and a **4x speed increase** through **mixed-precision quantization** and **Flash Attention** maintained image quality.

Visual-Inertial Odometry on the MIT Racecars | C++, ROS **Dec 2021**

- Implemented **Visual-Inertial Odometry (VIO)** on the MIT Racecar using an IMU, depth camera, and LiDAR.
- Enabled **real-time SLAM, loop closure**, and mesh creation via **GPU acceleration** on embedded Nvidia TX2.
- Validated VIO system against OptiTrack, **achieving precise path estimation and waypoint following**.

Publications

Diverse Database and Machine Learning Model for RNA Structure Prediction — Harvard - 2024

A nanoelectronics-blood-based diagnostic biomarker for ME/CFS — Stanford - 2019

The IDO Metabolic Trap Hypothesis for the Etiology of ME/CFS — Stanford - 2019

Technical Skills

Programming Languages: Python, Swift, Java, Kotlin, C++, C, JavaScript

Databases: PostgreSQL, DynamoDB, Elasticsearch, SQLite, CoreData

Networking/Communication Protocols: REST, GraphQL, gRPC

Developer Tools: Git, Docker, JIRA, Postman, Figma, Weights & Biases, Amplitude, Hugging Face

Technologies/Frameworks: PyTorch, TensorFlow, SwiftUI, UIKit, AWS, React, Next.js, Vercel, Supabase, Linux, ROS

Spoken Languages: Native English Speaker, Conversational in Spanish

Relevant Coursework

- Computer Vision
- Reinforcement Learning
- Visual Navigation (SLAM)
- Performance Computing
- Machine Learning
- Operating Systems
- Artificial Intelligence
- Internet Architecture
- Computer Security
- Efficient Algorithms
- Feedback Control
- Robotic Manipulation

Affiliations / Awards

Nexus Hackathon 3rd Place — Lucid Motors

Eta Kappa Nu IEEE Honor Society Member — UC Berkeley

Escape from Alcatraz Triathlon Finisher — San Francisco

Half-marathon Finisher x2 — Boston

Valedictorian — Fremont High School in Sunnyvale

Academic All American — USA Water Polo