KAI WEI

Palo Alto, CA | (510) 604-1381 | weikai77@gmail.com | https://www.linkedin.com/in/kaiwei

Goal

Software engineer role on a product team building LLM-powered applications.

Professional Summary

I have worked for 20 years in the tech industry (first as a software engineer, and later as an engineering manager) at heavy weight tech companies such as LinkedIn, Uber, and Google. I was getting bored with the daily repetitive grind, until the arrival of LLM / Generative AI rekindled my passion for building software. I believe this new technology is unlocking a new paradigm of problem solving, and I am looking for opportunities to contribute to this revolutionary journey and have fun with it.

Work History

Engineering Manager   Jan 2022 - Nov 2024
Google / YouTube   Mountain View, CA

Leading a team of 10+ engineers on YouTube Nitrate Data & ML, working at the intersection of data, ML, and infrastructure.
- Nitrate is a promotional platform for driving paid product growth and feature usage across product teams at YouTube. The Nitrate Data and ML team's mission is to use data and machine learning to maximize Nitrate's impact through inventory optimization and audience targeting.
- Led the team to increase YT paid products operating income by 2+% through the launches of a new inventory management system, improved ML models, and critical campaign launches such as Sunday Ticket and Prime Time Channels. In addition, we improved the Nitrate platform's promo efficiency by 10+% through the launch of min performance filters.
- Developed strategic roadmaps, grew the team by hiring the right talent for the right roles (made 6 hires in 3 years), and supported the career progression of high performers by aligning them with high impact opportunities (6 successful promotions in 3 years).

Engineering Manager   Jul 2020 - July 2021

CloudKitchens (aka City Storage Systems)   Mountain View, CA

Led a team of 10+ high performing engineers building a business of "growth as a service for digital restaurants".
- Helped define the product, technical, and organizational vision for the team.
- Grew the tech team (eng/DS/PM/design) by 3x and supported a 15x business growth in the first 11 months.
- Wore multiple hats, including filling the gap of product management and product operations.
- Introduced lightweight processes and structure to streamline team operation.
- Collaborated closely with stakeholders in operations, international, and platform engineering.

Engineering Manager   Jan 2018 - Jul 2020
Uber Eats   San Francisco, CA

From August 2019 to July 2020, I led the Eater Engagement Pricing eng team (of ~20 engineers) at UberEats. We built pricing products that influenced our users' behavior for the benefits of sustainable business growth and marketplace efficiency. We were a full stack engineering team that tackled problems at the intersection of product experience, distributed system, and data/ML.

From Jan 2018 to July 2019, I led the Courier Trip Experience eng team at UberEats. We built across the UberEats app and our courier app to help our delivery partners fulfill deliveries. We were a full stack engineering team that built mobile features as well as the backend services/intelligence that power these features. I managed this team and worked closely with cross-functional partners in product/DS/design/prodOps/etc as well as stakeholders across UberEats and Uber to execute our roadmap.

On both teams, I was responsible for:
- Growing the team size by 2-3x
- Building processes and structure to streamline team operation
- Developing product and organizational vision for the team
- Empowering the team on execution and shipping products

Engineering Manager   Jan 2016 - Jan 2018
Uber   San Francisco, CA

Led the Realtime Locations team within Marketplace. My team built services and solutions on the real-time locations data of Uber drivers and riders, including:
- Real-time location-based indexes of Uber marketplaces to power supply/demand

matching.
- Foundational solutions for real-time location-based configuration needs across Uber.
- Always-on location data store and pipeline.

Senior Software Engineer II   Dec 2014 - Dec 2015
Uber   San Francisco, CA

Worked on multiple generations of the real time dispatch system. Built Uber's first production-grade, high-QPS Golang service.

Staff Software Engineer   Jul 2012 - May 2014
LinkedIn   Mountain View, CA

Member of the ad serving platform team. Grew into a technical lead role over time.
- Led the design and implementation of a forecasting pipeline, which combines offline/online processing to deliver fast on-demand forecasting over high-dimensional data. This pipeline was built to be generic and has at least three use cases deployed in production.
- Led the design and implementation of supply forecasting for sponsored updates. Built on top of the forecasting platform, this forecasting application had to account for special business logic like frequency capping and competition.
- Led the design and implementation of budget pacing, a mechanism for achieving optimal temporal distribution of an ad campaign's budget spend for the benefits of higher advertiser ROI and overall revenue for LinkedIn.
- Worked on various aspects of the ad serving system, including frequency capping, tracking, and much more.

Senior Software Engineer   Mar 2010 - Jul 2012
LinkedIn   Mountain View, CA

Member of the Jobs engineering team. Worked on various features on the job seeker and job poster experiences on LinkedIn.

Principal Software Engineer   Feb 2008 - Mar 2010
Motorola Mobility (via acquisition of Good Technology)   Sunnyvale, CA

Member of the MOTOBLUR engineering team that built backend cloud services for Motorola's first
Android phones.

Senior Software Engineer   May 2005 - Feb 2008
Good Technology   Santa Clara, CA

Worked on a web-based management console for enterprise software using Java technologies.

Education

MS in Computer Science
University of California, Berkeley   Berkeley, CA

MA in Transportation Planning and Geographic Information Systems
University of California, Berkeley   Berkeley, CA

BS in Urban and Environmental Sciences
Peking University   Beijing, China

Certifications

Generative AI with Large Language Models   2025
Coursera

TensorFlow 2 for Deep Learning Specialization   2021
Coursera

Deep Learning Specialization   2021
Coursera