

# AI & ML at QCon San Francisco 2025

## Sessions

### Designing AI Platforms for Reliability: Tools for Certainty, Agents for Discovery

**Monday Nov 17**

Modern AI platforms don't have to choose between deterministic precision and probabilistic exploration—they need both.

**Aaron Erickson**

Senior Manager and Founder of the DGX Cloud Applied AI Lab @NVIDIA, Previously Engineer @ThoughtWorks, VP of Engineering @New Relic, CEO and Co-Founder @Orgspace

---

### The Future of Engineering: Mindsets That Matter When Code Isn't Enough

**Monday Nov 17**

Since the first compiler, software has been a stack of human-friendly abstractions translated into machine instructions. The engineers who understood at least some of what was going on under the hood were essential—indispensable even. But now?

**Ben Greene**

4X Founding CTO, Co-Founder and CTO @Tessi, Co-Creator of FreeFormula.Exchange, CTO-in-Residence at Techstars Boston

---

### Directing a Swarm of Agents for Fun and Profit

**Monday Nov 17**

Coding agents are a new tool, which many of us are trying to figure out how to use effectively.

**Adrian Cockcroft**

Technology Advisor and Consultant @OrionX.net, Previously VP Open Source and Sustainability @Amazon, Cloud Architect @Netflix, Distinguished Engineer @eBay

## Engineering at AI Speed: Lessons from the First Agentially Accelerated Software Project

**Tuesday Nov 18**

Claude Code is the first developer tool built specifically to maximize AI development velocity.

**Adam Wolff**

Engineer and Individual Contributor to Claude Code @Anthropic, Previously @Robinhood, @Facebook

---

## Engineering AI for Creativity and Curiosity on Mobile

**Tuesday Nov 18**

This talk shares practical lessons from building production-grade AI for creativity and curiosity on mobile devices.

**Bhavuk Jain**

Tech Lead @Google

---

## Modernizing Relevance at Scale: LinkedIn's Migration Journey to Serve Billions of Users

**Tuesday Nov 18**

How do you deliver relevant and personalized recommendations to nearly a billion professionals—instantly, reliably, and at scale? At LinkedIn, the answer has been a multi-year journey of architectural reinvention.

**Nishant Lakshmikanth**

Engineering Manager @LinkedIn, Leading Infrastructure for "People You May Know" and "People Follows", Previously @AWS and @Cisco

---

## Improving Meta Generative Ad Text using Reinforcement Learning

**Tuesday Nov 18**

Reinforcement Learning with Performance Feedback (RLPF) unlocks a new way of turning generic GenAI models into customized models fine-tuned for specific tasks. This approach is especially powerful when combined with in-house data and performance metrics.

**Alex Nikulkov**

Research Scientist (RL lead for Monetization GenAI) @Meta

---

## **Designing Fast, Delightful UX With LLMs for Mobile Frontends**

**Tuesday Nov 18**

Delivering AI-powered features in mobile apps is not just about calling an LLM API. It is about crafting fast, reliable, and engaging user experiences.

**Balakrishnan (Bala) Ramdoss**

Senior Android Engineer @Amazon - Building Camera-Based AI Features, Specializes in Scalable Solutions for Complex Challenges

---

## **From Content to Agents: Scaling LLM Post-Training Through Real-World Applications and Simulation**

**Tuesday Nov 18**

This talk presents a comprehensive journey through modern AI post-training techniques, from Pinterest's production-scale content discovery systems to enterprise agent training through Veris AI's simulation.

**Faye Zhang**

Staff Software Engineer @Pinterest, Tech Lead on GenAI Search Traffic Projects, Speaker, Expert in AI/ML with a Strong Background in Large Distributed System

**Andi Partovi**

Co-Founder @Veris AI, Making AI Agents World-Ready

---

## **Dynamic Moments: Weaving LLMs into Deep Personalization at DoorDash**

**Tuesday Nov 18**

In this talk, we'll walk through how DoorDash is redefining personalization by tightly integrating cutting-edge large language models (LLMs) with deep learning architectures such as Two-Tower Embeddings (TTE) and Multi-Task Multi-Label (MTML) models.

**Sudeep Das**

Head of Machine Learning and Artificial Intelligence, New Business Verticals @DoorDash, Previously Machine Learning Lead @Netflix, 15+ Years in Machine Learning

**Pradeep Muthukrishnan**

Head of Growth for New Business Verticals @DoorDash, Previously Founder & CEO @TrustedFor, 15+ Years in Machine Learning

---

## **Automating the Web With MCP: Infra That Doesn't Break**

**Tuesday Nov 18**

AI agents are only as strong as the infrastructure beneath them. In this talk, we'll walk through the architecture behind Browserbase's model context protocol (MCP), built to support stateful browser automation at scale.

**Paul Klein**

Founder @Browserbase, previously Director of Self-Service & Engineering Manager @Mux, Co-Founder & CTO @Stream Club, Technical Lead @Twilio Inc.

---

## **Producing the World's Cheapest Tokens: A How-to Guide**

**Wednesday Nov 19**

AI inference is expensive, but it doesn't have to be. In this talk, we'll break down how to systematically drive down the cost per token across different types of AI workloads.

**Meryem Arik**

Co-Founder and CEO @Doubleword (Previously TitanML), Recognized as a Technology Leader in Forbes 30 Under 30, Recovering Physicist

---

## **Choosing Your AI Copilot: Maximizing Developer Productivity**

**Wednesday Nov 19**

The AI coding agent landscape evolves weekly. This talk compares today's frontrunners, shows where each shines, and shares prompts, policies, and "rules templates" that turn code suggestions into production-

quality output.

**Sepehr Khosravi**

Machine Learning Platform Engineer @Coinbase, Award Winning Instructor @UC Berkeley - Gen-AI Bootcamp, Founder @ AI Scouts

---

## Accelerating LLM-Driven Developer Productivity at Zoox

**Wednesday Nov 19**

Over the past year, Zoox has invested in integrating Large Language Models (LLMs) into the entire developer lifecycle through a companywide initiative called Zoox Intelligence (ZI).

**Amit Navindgi**

Staff Software Engineer, Developer Experience @ Zoox, Leading Applied AI Initiatives

---

## The Ironies of A<sup>2</sup> I<sup>2</sup>

**Wednesday Nov 19**

In this talk, we'll explore some of the "ironies" of automation—and now, artificial intelligence—in their interactions with software operators (i.e. you), especially during high consequence, high tempo situations (aka incidents).

**J. Paul Reed**

Staff Incident Operations Manager @Chime

---

## Powering the Future: Building Your GenAI Infrastructure Stack

**Wednesday Nov 19**

Behind every productivity leap is a rock-solid platform. Go under the hood with Intuit's GenOS team to see how vector stores, prompt management, RAG pipelines, and agent orchestration come together to serve ~100 million users.

**Merrin Kurian**

Distinguished Engineer @Intuit

---

## From ms to $\mu$ s: OSS Valkey Architecture Patterns for Modern AI

Wednesday Nov 19

As AI applications demand faster and more intelligent data access, traditional caching strategies are hitting performance and reliability limits.

**Dumanshu Goyal**

Uber Technical Lead @Airbnb Powering \$11B Transactions, Formerly @Google and @AWS

---

## AI-Driven Productivity: From Idea to Impact

Wednesday Nov 19

In this session you'll learn how product leaders turn GenAI enthusiasm into an enterprise-ready blueprint for real productivity gains.

**Jyothi Nookula**

Product Leader Director with 13+ Years Driving AI Product & Platform Innovation, Previously @Meta, @Amazon, and @Etsy

---

## Trustworthy Productivity: Securing AI-Accelerated Development

Wednesday Nov 19

Autonomous agents operate in a continuous loop: perceive context → reason → execute tools → observe. Each edge creates distinct attack surfaces. This talk maps vulnerabilities—memory poisoning in context ingestion, goal hijacking during reasoning, blind execution at the action stage.

**Sriram Madapusi Vasudevan**

Senior Software Engineer @AWS Agentic AI, Previously Core Team @AWS SAM, AWS Cloudwatch, Core Developer @Openstack

---