

Justin Thomas

AI-DRIVEN ARCHITECTURE & ENGINEERING LEADERSHIP

Ex-AWS · Stanford AI/NLP certified · Fortune-100 engineering leader. Two decades building and leading backend platforms, SRE, and edge — now shipping generative AI, resilient APIs, and agents.

📍 United States justin.thomas1@gmail.com [linkedin.com/in/justin-thomas](https://www.linkedin.com/in/justin-thomas) stackoverflow.com/users/491770
westcot.io

LEADERSHIP

Engineering leadership · org scaling (2 → 22) · hiring & mentoring · operational excellence · SRE

TECHNICAL

AI agents OpenCode

Skill creation RAG / LLMs

Kubernetes / GKE GCP

AWS CI/CD · GitLab

nginx gateways Pub/Sub

BigQuery Postgres

Node.js TypeScript

Java Go

CERTIFICATIONS

- › Artificial Intelligence Professional Program
- › XCS224N — NLP with Deep Learning (Stanford)
- › XCS221 — AI: Principles & Techniques (Stanford)
- › Mathematics for Machine Learning
- › Commercial Pilot License

EXPERIENCE

●

Ford Credit
Principal Architect
May 2024 — Present · United States

- Lead architecture for Ford Credit's emerging AI space — driving **90% adoption of AI/agentic coding** across engineering.
- Shipped a reusable **RAG call-center platform in 2 months**, integrating knowledge bases to streamline support workflows.
- Lead AI transformation org-wide — routinely presenting to **30–70 engineers** in the cross-functional AI Tech Guild and Architecture Forum.
- Shipped pilot **AI agent flows** integrating Jira and other sources to automate updates and ticket compliance.
- Helped **hire 4 AI engineers** and set software-excellence standards: CI/CD, linting, and structured AI/ML hiring panels.

●

Amazon Web Services (AWS)
Software Engineering Manager
Oct 2022 — Apr 2024

- Led 5 engineers delivering third-party integrations for **CodeCatalyst** & CodeSuite.
- Drove region expansion across **5 global regions**, aligning 30 services with 16 teams for on-time re:Invent launches.
- Built a Kotlin ticket-ingestion system; ran security reviews; mentored and promoted 2 engineers.

●

Verizon Media (→ Yahoo / Edgio)
Director of Engineering
Oct 2018 — Sep 2022 · Los Angeles, CA

- Built and scaled SRE, platform & EdgeCompute orgs from **2 to 22 engineers**, growing internal leaders.
- Migrated **500 engineers** to Kubernetes + GitLab CI/CD — Time-to-Ship from **3 months to 1 day**.
- Launched **"Functions at Edge,"** a globally available serverless edge product.
- Implemented SLOs that cut time-to-detection from 15 min to **2 min**.

EDUCATION

Georgia Institute of Technology

M.S. Computer Science · expected 2028

North Carolina State University

B.S. Philosophy & Biochemistry · minor Computer Science



Red Bull Media House
Director of Engineering → Lead Software Engineer

Jan 2013 – Oct 2018 · 5 yrs 10 mos · Santa Monica, CA

- Designed an **nginx API gateway** for Red Bull TV's microservices — API latency **700ms → 100ms**.
- Migrated the platform from Java to **Golang: 8 → 1,000 RPS** (125× throughput), operational cost –90%.
- Moved infra from Heroku/AWS to **GKE**; adopted Pub/Sub, CloudSQL & BigQuery over Kafka, cutting DevOps cost 80%.
- Stood up Grafana & Kibana monitoring (–50% detection time); data-mined artist social signals to sign 10 emerging artists.

EARLIER

Chromatik — Lead Software Engineer 2011–2013

Front- & back-end for a sheet-music reader + recording/sharing platform (Backbone/canvas, Node.js, RabbitMQ, MongoDB).

Rally Software — Sr. Software Engineer, AgileZen 2010–2011

Shipped Rally Scrum-tool features (Java, Oracle, Spring) and rewrote the AgileZen backend + its Node.js/RabbitMQ notification service.

Verizon Business — Sr. Software Engineer 2006–2010

Redesigned network-management architecture; Lucene search engine **20s → 10ms**; data load 7 min → 5 s; automated ticketing.

bioMérieux — Software Engineer 2006

Ruby on Rails & Spring web apps generating pathogen growth curves.

Honeywell Aerospace — Software Intern 2000–2001

Built "six-pack" test-bench visualization software for CFM turbine engine controls on the fuel-control test line.

SELECTED WORK — WESTCOT.IO

Hermes Observability — open-source AI-agent ops stack github

Prometheus/Grafana cost, latency & error monitoring for LLM-agent sessions — per-session cost metrics, alerting & a deadman heartbeat. Extracted into a reusable repo.

Agent Deck Bake-off — multi-LLM agent pipelines LLMs

Agents that generate complete slide decks; benchmarked four frontier models on one brief for cost and quality.

Quadcopter Drop Bay — servo-actuated mechatronics 3D · servo

Designed, 3D-printed & programmed a servo-released payload bay for an autonomous drone — CAD-modeled, with an interactive three.js kinematics viewer.