

Keith Steward, Ph.D.

AI Tooling Director | Agent Evaluation & Reliable Platforms

Boston, MA | 508-318-8035 | keith.steward@gmail.com

[LinkedIn](#) | [IntelligentDataWorks.ai](#) | [IDW LinkedIn](#)

PROFILE

Player-coach AI leader who turns customer and product signals into reliable agent architectures, evaluation systems, and measurable platform improvements. Combines hands-on RAG, agent orchestration, prompt and planner optimization, LLM-as-judge evaluation, observability, guardrails, and production operations with team leadership and executive communication. Personally built the core and first 80% of a production AI platform, then led 11 engineers supporting 70+ automations across more than 20 workflows with 99.9% availability. Builds practical feedback loops that improve quality, latency, reliability, safety, and adoption.

SELECTED CAREER HIGHLIGHTS

- Built and developed engineering organizations as large as 25 people, including managers and team leads, with responsibility for hiring, onboarding, mentoring, performance reviews, staffing, delivery standards, career development, and accountable execution.
- Architected and operated production AI, data, and cloud platforms spanning agentic AI, RAG, MLOps/LLMOps, evaluation, observability, auditability, security, CI/CD, automated promotion, and resilient operations.
- Advised 200+ enterprise customers and C-suite stakeholders on AI/ML, data, platform architecture, modernization, governance, adoption, and measurable business outcomes across concurrent client engagements in complex, regulated enterprise environments.
- Created global AWS AI/ML and Big Data communities totaling 180+ specialists, produced 170+ enablement assets, and established technical standards and reusable practices for customer-facing engineers.
- Built the AssistX core and first 80% of its initial application hands-on, then led 11 engineers around 70+ automations across more than 20 workflows with 99.9% availability.
- Built closed-loop quality practices that connect production telemetry, customer feedback, failure analysis, offline and online evaluation, prompt and planner changes, retrieval quality, model routing, release gates, and incident learning. Used measurable reliability, latency, cost, safety, adoption, and task-completion signals to prioritize platform improvements while coordinating engineering, product, governance, and customer-facing teams around explicit tradeoffs and accountable delivery. This included model evaluation, retrieval architecture, prompt optimization, production reliability, and metrics used to prioritize improvements.

PROFESSIONAL EXPERIENCE

Technical Co-Founder | Intelligent DataWorks | Boston, MA | Aug 2023 to Oct 2026

Built and operated a production AI platform while personally owning architecture, implementation, engineering standards, evaluation, delivery, governance, security, and cloud operations.

- Recruited and led 11 engineers selected from more than 400 applicants; established onboarding, architecture training, Agile sprints, mentoring, code review, delivery priorities, and production standards.
- Built the AssistX core and first 80% of its initial application hands-on using Python, FastAPI, LLMs, agents, RAG, PostgreSQL/PGVector, APIs, Docker, AWS, and CDK.
- Delivered 70+ automations across more than 20 workflows, accelerating selected work by up to 50x while sustaining 99.9% availability.
- Established AI evaluation, LLM-as-judge, guardrails, observability, auditability, model routing, cost controls, security, CI/CD, release, and incident-response practices.
- Owned a \$900,000 budget, staffing, product and technical roadmaps, architecture, releases, cloud operations, and executive decisions across the venture.
- Completed planned transition from operating CEO responsibilities to pursue a long-term senior AI/engineering leadership role.

Sr. Director of Engineering | SuccessKPI | Remote | Jun 2022 to Aug 2023

Built and led engineering for a new AI-powered workforce-management platform within a cloud-native enterprise SaaS company.

- Recruited and mentored eight engineers and one manager, with responsibility for performance reviews, compensation decisions, delivery standards, staffing, architecture, and roadmap execution.
- Directed a distributed Agile organization through production launch for four large customers, including the U.S. Departments of State and Education.
- Designed forecasting, machine-learning, optimization, and scheduling capabilities integrated with Amazon Connect, Genesys, and Talkdesk.

Sr. Director, Data Analytics & Operations | TriMark USA | Mansfield, MA | Aug 2021 to May 2022

Built the cloud data organization and operating model supporting enterprise analytics across a \$1B+ company and 11 divisions.

- Upskilled an initial four-person analytics team and expanded it into a seven-member AWS engineering organization with additional contractors.
- Aligned executives around more than 50 use cases and delivered the first enterprise data lake within ten months, reducing reporting from roughly two months to one day.
- Owned staffing, performance reviews, architecture, delivery planning, risk, change management, executive reporting, and cross-functional priorities.

Engineering Lead, Cloud Data Privacy & AI Governance | MassMutual | Boston, MA | Nov 2020 to Aug 2021

Led architecture and engineering across privacy, governance, Responsible AI, data, and cloud initiatives in a regulated enterprise.

- Guided five initiatives spanning policy automation, model transparency, bias, drift, monitoring, human oversight, auditability, data governance, and governance APIs.
- Partnered with Enterprise Architecture and SRE on AWS, security, IAM, CI/CD, event-driven, container, and infrastructure-as-code standards.
- Delivered a cloud-native privacy platform within eight months that automated approximately 90% of a formerly manual workflow.

VP of AI/ML | NorthBay Solutions | Boston, MA | Jan 2019 to Nov 2020 | (Role ended in company-wide COVID layoff).

Recruited from AWS by the CEO to build an enterprise AI/ML consulting practice and reusable cloud delivery capabilities.

- Led an 18-person distributed engineering team delivering the Data Lake Accelerator while remaining hands-on in initial product architecture and implementation.
- Advised enterprise clients on AI, ML, data, platform strategy, architecture, and production delivery across concurrent consulting engagements.
- Combined executive advisory, staffing, reusable standards, cloud architectures, and production systems for enterprise customers.

Sr. Specialist Solutions Architect, AI/ML & Big Data | Amazon Web Services | Boston, MA | May 2015 to Dec 2018

Helped establish and scale AWS AI/ML and data specialist capability while advising enterprise customers on production adoption.

- Advised more than 200 enterprises across North America on AI/ML, data platforms, architecture, governance, cloud adoption, and business outcomes.
- Founded and led technical communities totaling 180+ specialists, created 170+ enablement assets, and established standards and reusable client practices.
- Delivered 36 invited presentations, trained 180+ architects, and earned AWS ML Elite Gold and Platinum recognition.

EARLIER RELEVANT EXPERIENCE

- Led a 25-person team, including managers, through development and commercial launch of CambridgeSoft's Biological Notebook for pharmaceutical and biotechnology customers.
- Managed cloud and device-software engineering for an FDA-cleared connected dental product, including a sub-team manager and an AWS/Cassandra production platform.

EDUCATION & RECOGNITION

Ph.D., Molecular Biology / Microbiology & Immunology, University of Western Ontario. AWS ML Elite Gold and Platinum; AWS re:Invent Rockstar; named inventor on multiple patents.

TECHNOLOGIES

engineering leadership • manager development • hiring • onboarding • performance management • career development • staffing strategy • engineering standards • client delivery • executive advisory • AI SDLC • AI governance • Responsible AI • AI evaluation • LLM-as-judge • observability • auditability • production safety • agentic AI • AI agents • RAG • MLOps • LLMOps • platform architecture • pipeline architecture • cloud infrastructure • operational resilience • Python • PostgreSQL • PGVector • AWS • Google Vertex AI • SageMaker • Amazon Bedrock • Docker • Kubernetes • AWS CDK • CI/CD • Git • Linux • LangChain • LangGraph • MCP • data pipelines • executive communication