

Keith Steward, Ph.D.

Enterprise AI Architect | Knowledge, Trust & Agentic Platforms

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

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

PROFILE

Hands-on enterprise AI architect who designs governed foundation-model and agentic systems spanning knowledge, trust, control, integration, and production operations. Combines Python, RAG, MCP, orchestration, evaluation, observability, guardrails, security, Responsible AI, hybrid-cloud patterns, and executive influence. 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. Translates governance requirements into reusable architecture and working technical controls.

CORE TECHNOLOGIES

Python • LLM Applications • Agentic Systems • Agent Orchestration • Tool Calling • RAG • Structured Outputs • LLMOps • AI Evaluation • Guardrails • Vector Databases • Observability • AWS • FastAPI • MCP

SELECTED CAREER HIGHLIGHTS

- Architected and built the AssistX core and first 80% of its initial application hands-on using Python, FastAPI, LLMs, RAG, agents, vectors, APIs, and AWS.
- Delivered 70+ automations across more than 20 workflows with 99.9% availability and selected work accelerated up to 50x.
- Established evaluation, regression, guardrails, human review, model routing, prompt/config versioning, observability, cost controls, and staged releases.
- Advised 200+ AWS enterprise customers and led build-versus-buy, architecture, security, reliability, and adoption decisions.
- Led Responsible AI, privacy, IAM, auditability, bias, drift, and model-governance work in regulated financial services.
- Created production patterns spanning identity, permissions, data access, human escalation, evaluation, observability, model routing, cost controls, and auditable decision traces in regulated environments.
- Translated Responsible AI and compliance requirements into concrete platform controls spanning model access, approved providers, data boundaries, evaluation, human review, prompt-injection defenses, observability, audit trails, drift and bias monitoring, release gates, rollback, incident response, and lifecycle accountability. Balanced reusable enterprise standards with hands-on implementation so architecture councils, product teams, security leaders, and operators could make consistent decisions while moving foundation-model and agent use cases safely into production. Applied these patterns to foundation models and Agentic AI, strengthening safety and consistency across enterprise platforms.

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

- Built an AWS data-and-ML accelerator and delivered hands-on enterprise AI/ML consulting solutions.
- Developed commercial AI, bioinformatics, and scientific software, including patented knowledge-discovery systems.

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

Python • LLM applications • agentic systems • agent orchestration • tool calling • structured outputs • RAG • LangChain • LangGraph • MCP • LLMops • AI evaluation • regression testing • guardrails • prompt versioning • model routing • observability • cost management • vector databases • PostgreSQL • PGVector • FastAPI • REST APIs • AWS • Amazon Bedrock • SageMaker • distributed systems • microservices • event-driven architecture • data pipelines • deterministic workflows • human-in-the-loop • Responsible AI • AI governance • privacy • security • IAM • auditability • CI/CD • Docker • Kubernetes • AWS CDK • Git • Linux