

Keith Steward, Ph.D.

AI Automation & Quality Director | Evaluation & Release Safety

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

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

PROFILE

Hands-on AI automation and quality leader who builds the evaluation, testing, observability, and release controls that make agentic systems trustworthy in production. Personally implemented LLM-as-judge evaluation, prompt regression, guardrails, prompt-injection screening, synthetic-data testing, monitoring, auditability, CI/CD, and incident-response practices for a live AI platform. Combines Python, RAG, agents, regulated-system experience, and life-sciences software depth with a record of establishing new technical functions and mentoring engineers. Builds practical quality systems that accelerate delivery while preserving trustworthy production behavior.

SELECTED CAREER HIGHLIGHTS

- Built evaluation infrastructure for a production AI assistant platform, including structural checks, LLM-as-judge, human review, automated regression tests, synthetic test data, quality thresholds, and repeatable validation workflows.
- Implemented prompt-injection screening, guardrails, observability, auditability, model routing, cost controls, security, CI/CD, release, rollback, and incident-response practices for 70+ automations across more than 20 workflows.
- Personally built the AssistX core and first 80% of its initial application with Python, FastAPI, agents, RAG, PostgreSQL/PGVector, APIs, Docker, AWS, and CDK; sustained 99.9% availability.
- Established and grew engineering capabilities from scratch, recruiting and mentoring 11 engineers at Intelligent DataWorks and previously building AI/ML, data, cloud, and software teams as large as 25.
- Delivered regulated privacy and governance systems at MassMutual, including model transparency, bias, drift, monitoring, auditability, data governance, and a privacy platform that automated approximately 90% of a manual workflow.
- Brings a Ph.D. in Molecular Biology and commercial life-sciences software experience spanning bioinformatics, pharmaceutical research, scientific workflows, and regulated healthcare technology.

PROFESSIONAL EXPERIENCE

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

Build and operate a production AI platform while personally owning architecture, implementation, engineering standards, evaluation, delivery, 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.
- Own a \$900,000 budget, staffing, product and technical roadmaps, architecture, releases, cloud operations, and executive decisions across the venture.
- Initiated strategic acquisition discussions regarding IDW's AI platform as part of a planned transition to a 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

- Original developer of Ingenuity Pathway Analysis at Ingenuity Systems, a commercial bioinformatics platform adopted by biotechnology, pharmaceutical, academic, and government research organizations.
- Developed bioinformatics analysis software at Wyeth Research and led development of CambridgeSoft's Biological Notebook for pharmaceutical and biotechnology customers.

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 • pytest • AI agents • agent testing • agent lifecycle management • RAG evaluation • prompt regression • LLM-as-judge • accuracy benchmarking • safety benchmarking • bias detection • golden-set testing • synthetic data • automated testing • quality thresholds • release gates • CI/CD • GitHub • observability • monitoring • auditability • guardrails • prompt-injection screening • LangChain • LangGraph • RAG • embeddings • vector databases • PostgreSQL • PGVector • FastAPI • REST APIs • AWS • SageMaker • Amazon Bedrock • Docker • AWS CDK • Responsible AI • AI governance • model explainability • drift monitoring • security • technical documentation • regulated software • bioinformatics