

Ken Stager

AI Systems Architect

Greenfield, MA 01301 | (267) 226-0736 | ken@kenstager.com | linkedin.com/in/ken-stager

Enterprise AI systems architect working at the platform layer. Sole AI Solutions Architect behind a production AI platform at Digital Realty, and concurrently contracted by Microsoft as an AI Architect to carry an internal agent from beta to full production.

An architect who ships: routing, retrieval, and evaluation built and measured as engineered components, not diagrammed and handed off. Architecture review boards, code review across the department, and governed AI infrastructure for the engineering teams that build on it.

EXPERIENCE

AI Solutions Architect

Digital Realty Trust, Inc. | June 2025 – Present

- Sole AI Solutions Architect for Praxo, an enterprise AI platform on Azure — in production, available to ~4,000 employees, with ~600 monthly active users in a recent month. Microsoft 365 integration for mail, calendar, and files runs in production on an on-behalf-of flow.
- Tasked by leadership to create the AI API Developer Portal, a governed gateway brokering the organization's AI models to internal engineering teams; lead architect, live in production. Approval-gated provisioning, one-time encrypted key delivery, on-demand rotation, and per-subscription token metering with cost showback, under least-privilege roles where provisioners cannot read keys.
- Architected the multi-agent retrieval layer now live in production — parallel domain agents across 5 domains, Reciprocal Rank Fusion (RRF) merge, and a prompt composer. Measured routing accuracy at 88% and retired the legacy fallback path from 32% to 0%, validated in shadow mode behind feature flags.
- Accepted a measured +7% end-to-end latency cost to cut response-generation time 18%, where users actually feel it; parallelizing router classification returned an observed 500–1,500 ms per query.
- Built the reliability scaffolding the platform ecosystem depends on — 3,000+ automated tests, shadow-mode comparison, and feature-flagged rollout — so behavior is validated before activation rather than after.
- Present architecture at review boards, review code across most of the department, and led a platform security audit and remediation program.

AI Architect (Contract)

Microsoft Corporation | April 2026 – July 2026

- Carried an internal AI agent from beta to full production — beta flag removed and streaming rolled out to production.
- Originated and built the agent's evaluation platform end to end: harnesses exercising the real published agent through the full production path, independent quality instruments whose deltas had to agree, a statistics layer, and a cost model.
- Authored the go/no-go readiness assessment behind the production rollout, using paired tests, false-discovery control, and explicit practical-significance thresholds.

AI Solutions Architect (Independent Consulting)

Self-employed | November 2023 – June 2025

- Delivered AI-driven scheduling automation for a healthcare organization under NDA, reducing operational backlog without adding staff while maintaining HIPAA-aligned data handling.
- Implemented OCR-based document processing with Amazon Textract, turning unstructured healthcare forms into structured data for downstream workflows.
- Built an end-to-end news analytics pipeline — discover, extract, cluster, deduplicate — for near-real-time triage.

Customer Success Manager

BlueConic, Inc. | February 2022 – November 2023

- Onboarded enterprise customers on a Customer Data Platform (CDP), translating complex data models into practical enablement.
- Advised global clients on first-party data strategy and privacy compliance, including GDPR and CCPA.

Owner and Business Manager

Penn360 | August 2016 – December 2021

- Founded and ran an architectural photography and marketing-media business serving architects, designers, and realtors; owned client contracts, operations, and finances, and led a team of three.

SKILLS

AI Systems — LLM systems engineering, multi-agent orchestration, agent routing, intent classification, guardrails, structured outputs, Azure OpenAI, Semantic Kernel

Retrieval — Retrieval-Augmented Generation (RAG), hybrid retrieval, vector search, embedding models (text-embedding-3-large), Reciprocal Rank Fusion (RRF), GraphRAG, context compression, entity resolution, pgvector

Reliability & Evaluation — evaluation harnesses, shadow-mode comparison, regression testing, pytest, feature flags, safe rollout, observability, structured logging, schema validation

Cloud & Platform — Azure Functions, Azure Container Apps, Azure Cognitive Search, Azure Cosmos DB, Azure Blob Storage, Azure AD, AWS (Amazon Textract), Infrastructure as Code (Bicep), PostgreSQL, Docker, Azure DevOps

Languages & Frameworks — Python, TypeScript, SQL, FastAPI, Flask, React, Node.js

EDUCATION

Columbia University, Fu Foundation School of Engineering and Applied Science | May 2024 Advanced AI and Machine Learning