



Ezekiel Olson

AI Engineering Manager, Generative AI Evaluation

☎ (630) 555-3172 ✉ ezekiel.olson@example.com

🌐 linkedin.com/example/ezekielolson 📍 Naperville, IL 60540

STRENGTHS

- ✓ **Evaluation Strategy**
Built layered checks for uncertain model outputs. Product leaders gained clearer evidence before release decisions.
- 👥 **Team Development**
Set practical goals and coached engineers through difficult AI testing work. Peers became go to resources.
- 🔍 **Incident Learning**
Connected traces, logs, and deployment history during reviews. Repeated failures became useful automated scenarios.
- 🛠️ **Cross Team Trust**
Shared quality metrics with product and platform partners. Conversations moved from opinion toward evidence.
- ⚙️ **Delivery Discipline**
Used feature flags and progressive releases during rapid experimentation. Teams kept momentum without losing control.

SKILLS

AI Engineering Leadership

GenAI Evaluation Test Automation

Software Quality Assurance LLMs

SLMs Multimodal Models

Agentic Systems

Prompt Engineering RAG

Vector Databases Embeddings

LoRA Fine Tuning

Human Validation

AI Safety Evaluation

SUMMARY

AI engineering leader with 12 years of **software quality** and platform experience, including more than five years guiding teams across test automation, validation, observability, and **generative AI evaluation**. Leads enterprise software and GenAI delivery across hybrid cloud and edge connected environments through disciplined SDLC practices, CI/CD, progressive deployment, incident management, and measurable quality controls. Builds evaluation strategies for non deterministic systems using deterministic checks, probabilistic assessment, RAG quality measures, safety testing, human review, A/B testing, tracing, telemetry, and **root cause analysis**. Hands on with LLMs, SLMs, multimodal and **real time models**, agentic systems, prompt engineering, vector databases, embeddings, LoRA, MCP, A2A, and synthetic data. Creates clear priorities, develops engineers, and connects technical evidence with product risk, release confidence, and enterprise outcomes.

EXPERIENCE

Director, AI Quality Engineering

Praxis AI Platforms 📅 February 2023 - Present 📍 Chicago, IL

Lead an AI quality engineering group responsible for enterprise agents, digital assistants, RAG services, and multimodal capabilities across cloud and edge connected products. Set technical direction, define evaluation evidence, develop engineering capability, and align release decisions with product risk across Azure, AWS, and GCP.

- Led layered evaluation framework for agents and **RAG** services, improving release evidence for non deterministic outputs.
- Defined safety checks, retrieval measures, human review, and model grading for enterprise GenAI capabilities.
- Scaled **CI/CD** quality gates, **feature flags**, canary releases, and regression suites across multiple AI products.
- Connected Langfuse, Phoenix, tracing, and **telemetry** with prompts, retrieval context, tools, and downstream failures.
- Guided engineers through MCP, A2A, vector search, **embeddings**, chunking, LoRA, and real time model testing.
- Partnered with product leaders to set **quality metrics**, risk thresholds, and release evidence across cloud environments.

Senior Engineering Manager, Software Quality and AI Validation

Atlas Industrial Software 📅 May 2020 - January 2023 📍 Oak Brook, IL

Managed **software quality**, automation, and validation engineers supporting cloud services and embedded edge applications for connected industrial products. Built team capability, standardized SDLC and release practices, and connected validation evidence with production risk across several product lines.

- Managed quality engineers supporting cloud services and embedded edge applications for connected industrial products.
- Built API, event flow, and device interaction tests, integrating coverage into CI/CD pipelines and release controls.
- Introduced blue/green and canary **validation** playbooks with **feature flags** before broad production releases.
- Led incident root cause reviews using logs, distributed traces, test history, and deployment changes.
- Converted recurring failures into automated scenarios and expanded validation coverage for complex workflows.
- Developed engineers through mentoring, architecture reviews, **quality metrics**, and cross product release standards.

LANGUAGES



MY CAREER



- Director, AI Quality Engineering at Praxis AI Platforms (3.6 Years)
- Senior Engineering Manager, Software Quality and AI Validation at Atlas Industrial Software (2.7 Years)
- Principal Software Development Engineer in Test at Northline Cloud Systems (2.8 Years)
- Software Quality Engineer at Lakefront Systems Engineering (2.8 Years)

Principal Software Development Engineer in Test

Northline Cloud Systems 📅 July 2017 - April 2020 📍 Chicago, IL

- Designed test platforms for enterprise cloud applications spanning APIs, data pipelines, user workflows, and distributed integrations. Mentored test engineers, improved pre release evidence, and supported early machine learning evaluation through repeatable criteria and human review.
- Designed test platforms for cloud APIs, data pipelines, user workflows, and distributed enterprise integrations.
 - Integrated functional, reliability, and performance checks into build and deployment pipelines before release decisions.
 - Developed reusable test services and telemetry utilities that supported fault isolation across distributed systems.
 - Partnered with software engineers on testable interfaces, **observability** hooks, and deployment safeguards during development.
 - Evaluated NLP and machine learning services with curated datasets, acceptance criteria, and human output review.
 - Mentored test engineers on automation architecture, **root cause analysis**, software design, and risk based **validation**.

Software Quality Engineer

Lakefront Systems Engineering 📅 September 2014 - June 2017 📍 Evanston, IL

- Built automated quality practices for **enterprise software** delivered through structured SDLC cycles and frequent releases. Supported integration testing, release validation, defect triage, test data, and operational readiness across distributed application environments.
- Built automated integration and regression tests for enterprise products delivered through structured **SDLC** cycles.
 - Created service interface and data workflow harnesses with deterministic assertions and diagnostic logging.
 - Supported release **validation** and defect triage by linking behavior with code changes and test evidence.
 - Expanded continuous integration coverage for critical workflows used by development teams.
 - Maintained reliable test data and execution environments across evolving distributed application platforms.
 - Documented reproducible quality procedures for product requirements, testability, and operational readiness reviews.

PROJECTS

Enterprise GenAI Evaluation Framework 📅 2026

Built a layered evaluation approach for agents, RAG services, and multimodal systems using deterministic checks, probabilistic grading, safety review, retrieval measures, telemetry, and human validation.

Hybrid Cloud Release Quality Program 📅 2023

Established shared validation practices for cloud and edge connected products, combining CI/CD gates, feature flags, progressive deployment, incident learning, and production quality evidence.

LEADERSHIP & AWARDS

- AI Quality Leadership Recognition, Praxis AI Platforms, 2025
- Engineering Excellence Award, Atlas Industrial Software, 2022
- Technical Mentorship Recognition, Northline Cloud Systems, 2019

EDUCATION

Master of Science in Computer Science

Illinois Institute of Technology 🎓 GPA: 4.0 📅 2014 📍 Chicago, IL

Coursework: Artificial Intelligence, Software Engineering, Distributed Systems, Machine Learning

Bachelor of Science in Software Engineering

Iowa State University 🎓 GPA: 4.0 📅 2012 📍 Ames, IA

Coursework: *Software Testing, Database Systems, Algorithms, Systems Design*

CERTIFICATIONS

- AWS Certified Machine Learning Engineer - Associate 📅 2026
- Microsoft Certified: Azure AI Engineer Associate 📅 2025

TECHNICAL SKILLS

- **Cloud Platforms:** Azure, AWS, GCP
- **AI Cloud Services:** Azure AI Foundry, SageMaker, Bedrock
- **Generative AI Models:** LLMs, SLMs, multimodal models
- **AI Architectures:** RAG, vector databases, embeddings
- **Agent Platforms:** LangChain, LangGraph, MCP
- **Evaluation Platforms:** Langfuse, LangSmith, Humanloop
- **Evaluation Frameworks:** Ragas, DeepEval, Phoenix
- **Deployment Practices:** CI/CD, blue/green deployment, canary deployment
- **Quality Engineering:** test automation, software QA, validation
- **Observability Systems:** tracing, telemetry, incident management
- **Model Adaptation:** prompt engineering, LoRA, chunking
- **Testing Methods:** A/B testing, synthetic data, human review

PROFESSIONAL AFFILIATIONS

- Member, Association for Computing Machinery, AI and software engineering community
- Member, IEEE Computer Society, software quality and systems engineering community

ADDITIONAL INFORMATION

Work Status : Authorized to work in United States. No sponsorship required.

REFERENCES

AVAILABLE ON REQUEST