

Gary Robinson III

AI & QA Engineering Leader · LLM Evaluation · Test Automation

Vienna, VA · grobinson3@gmail.com · (703) 261-5577 · linkedin.com/in/garyrobinsoniii · garyrobinsoniii.com

SUMMARY

QA engineering leader with 15+ years advancing quality, automation, and security across web, mobile, and enterprise platforms, now focused on AI-native quality: LLM evaluation, hallucination detection, prompt and RAG pipeline testing, and agentic test generation. GenAI certified, with a track record of building QA practices from scratch, leading teams on federal and commercial systems, and integrating AI into everyday test design, data analysis, and defect triage. Active security clearance.

EXPERIENCE

Founder, AI Consultant & Principal Quality Architect

Apr 2026 – Present

GR3 VerifAI

- Established an AI-focused QA consultancy delivering LLM evaluation, AI data quality assurance, and GenAI testing strategy to federal, healthcare, and commercial clients.
- Engineered an API test automation suite with Playwright and Gemini, reaching 100% coverage of 50+ endpoints in one week.
- Deployed three AI agents that autonomously plan, generate, and repair test cases, improving suite reliability and throughput.
- Deliver AI-native quality services including LLM output assessment, hallucination detection, prompt testing frameworks, and RAG pipeline validation.

QA Engineering Team Lead

Jul 2023 – Mar 2026

Steampunk · USDA / FEMA National Flood Insurance Program

- Led a team of four engineers and stood up the QA practice, processes, and tooling that launched a flood insurance quote application in seven months.
- Built and maintained 200+ data-driven API tests in Postman covering every application endpoint.
- Adopted Katalon and Axe DevTools, then trained the team on functional, regression, automation, and Section 508 accessibility testing.
- Owned performance testing: defined coverage, automated nightly and weekly runs, and delivered actionable reports each release.
- Wrote a Jupyter notebook in Google Colab to process thousands of address records for data-driven tests, lifting coverage 30%.
- Audited QA practices across 12 teams, identifying gaps and improvements needed for a Release-on-Demand cadence.

QA Lead

Apr 2022 – Mar 2023

Sentral · Sentral.com and Reservation Lookup

- Designed and automated test plans for the website and backend with Playwright and Postman.
- Implemented continuous integration and production monitoring with ChecklyHQ for real-time alerting on test results and stability.

QA Lead

Aug 2021 – Apr 2022

Clearcover · Car Insurance Platform

- Wrote automated, data-driven acceptance tests with Playwright, Jest, Postman, and CodeceptJS.
- Drove root-cause analysis and process improvements; led migration to a new test management tool with requirements traceability and automation integration.

QA Director / Scrum Master

Jan 2021 – Aug 2021

Flyreel AI

- Implemented test management, automation frameworks, and a GitHub Actions CI/CD pipeline with Jira and Slack integrations.
- Built Android mobile automation from scratch with TestProject, automating 90% of previously manual tests and cutting regression time by two days.
- Configured ChecklyHQ API performance tests to surface regressions immediately.

QA Lead / Software Engineer / Security Advocate

Feb 2015 – Jan 2021

Rally Health (now UnitedHealthcare / Optum Digital) · UHC Find & Price Care

- Developed automated UI and API tests in Robot Framework (Python/Selenium) and WebdriverIO (JavaScript); built and maintained Jenkins CI.
- Maintained the application threat model and ran OWASP ZAP scans ahead of weekly releases.
- Paired with developers on full-stack verification, validating front- and back-end changes before merge and occasionally fixing bugs.

QA Engineer II / QA Lead

Mar 2014 – Feb 2015

Amazon · Amazon Appstore

- Scoped features, planned tests, and drove automation for multiple large-scale Appstore and Amazon Mobile Android releases.
- Delegated within the QA team and reported results to leadership as an advocate for quality at every phase.

Earlier Experience

Senior QA Engineer / QA Lead · Taxi Magic (now Curb)

2013 – 2014

Principal SW Quality Engineer · RSA, The Security Division of EMC

2013

SQA Engineer / QA Lead · Reality Mobile

2011 – 2013

SQA Engineer / QA Lead · Symantec · Managed Security Services

2010 – 2011

Software Engineer — Test Analyst / PL/SQL Developer · Accenture

2007 – 2010

SKILLS

AI & LLM Quality	LLM evaluation, hallucination detection, prompt engineering & testing, RAG validation, AI agents, generative AI
Test Automation	Playwright, Katalon, Robot Framework, Selenium, Appium, CodeceptJS, WebdriverIO, Postman, JMeter
Quality Practice	UI & API automation, performance testing, Section 508 accessibility, penetration testing, DevSecOps, Agile
Platforms & Tools	GitHub & GitHub Actions, Jenkins, Docker, AWS, Google Cloud, Jira, TestRail, HP Quality Center, Splunk, Wireshark
Languages	JavaScript, Python, SQL / PL-SQL, Shell, HTML, JSON

EDUCATION & CERTIFICATIONS

B.S. Information Technology, Information Security · George Mason University

2012

Quality and Safety for LLM Applications · DeepLearning.AI

2026

GenAI Essentials Certified · ExamPro

2025

The Data Scientist's Toolbox · Johns Hopkins University

2016

Certified Application Tester · MIT / Accenture Solutions Delivery Academy

2009