

ANTON GULIN

AI QA Architect | Senior SDET | AI Test Engineer | Lead QA Automation Engineer

Playwright + TypeScript + Python + AI Agents | LLM Evaluation | Former Apple SDET
10+ Years Architecting Enterprise Test Systems | 4 Years Building AI-Assisted QA

ALSO KNOWN AS: Software Development Engineer in Test | Test Automation Architect | Quality Engineer | Senior QA Engineer | Automation Lead | Quality Engineering Lead

San Diego Metropolitan Area (Remote)

916-800-4085 | i@anton.qa

linkedin.com/in/antongulin | anton.qa | github.com/antongulin

SUMMARY

AI QA Architect and Senior SDET with 10+ years engineering enterprise-scale test systems at Apple, healthcare payers, and Fortune 500 companies — and 4 years (since 2022) integrating AI into production test automation. One of the first QA engineers to combine AI agents, MCP (Model Context Protocol) servers, and LLM-powered workflows with Playwright, TypeScript, and Python, building end-to-end (E2E) test systems where AI explores and deterministic checks decide. Hands-on LLM evaluation practice: golden datasets, evaluation harnesses, hallucination detection, and hard asserts that AI-generated code cannot negotiate with. Deep healthcare-domain QA experience (Medicare Advantage, HIPAA-regulated systems). MBA in Economics and Project Management aligns test architecture to business velocity, team ROI, and delivery outcomes. Track record: 300% test growth, 50% faster deployments, 35% lower maintenance, 60% broader regression coverage.

EXPERIENCE

LEAD QA / SDET

Stenograph | Jun 2025 - Present | Remote

- Built the Playwright/TypeScript test automation framework from scratch for a web-based application used by court reporting and stenography professionals, establishing the foundation for scalable UI automation
- Integrated AI-assisted test development into the team workflow (Claude Code, GitHub Copilot, Playwright MCP), accelerating test creation and maintenance while keeping deterministic, human-owned assertions as the release gate
- Designed agentic test flows where an AI agent explores critical user workflows and fixed checks against seeded data (golden datasets) validate outcomes, catching AI false positives before CI
- Implemented and maintained CI/CD test execution workflows, improving regression coverage and enabling faster feedback during development and release cycles
- Partnered with developers and product stakeholders on test strategy and feature validation; drove QA best practices across the team, including framework structure, test reliability improvements, and defect reporting processes

SOFTWARE ENGINEER IN TEST

HFD | Contract | Jan 2025 - May 2025 | United States (Remote)

- Built and maintained automated test coverage using C# and Playwright for a medical web-based application, improving regression coverage across core workflows in a HIPAA-regulated environment
- Used AI pair-programming (GitHub Copilot, Claude) to accelerate C#/Playwright test development, with review discipline on all AI-generated test code
- Supported functional, integration, and regression testing efforts, improving release stability and defect detection; partnered with developers to investigate failures and validate fixes

QA TESTER, FINANCE DELIVERY TEAM

Blue Cross Blue Shield of Nebraska | Mar 2024 - Dec 2024 | Remote

- Built and maintained Playwright-based automated test scripts across web applications, APIs, and Power Apps within a .NET (C#) environment, strengthening regression coverage and release confidence
- Tested Azure Data Factory, Informatica Intelligent Cloud Services (IICS), and Azure Data Lake data flows, using SQL for data validation and transformation checks across environments
- Supported healthcare initiatives including Medicare Advantage, validating enterprise integrations across C# .NET Core, REST web services, and Azure services to meet regulatory requirements
- Applied AI-assisted test case generation and defect triage to healthcare data workflows, pairing LLM-generated scenarios with SQL-validated expected results

QA TESTER, DYNAMICS CRM TEAM

Blue Cross Blue Shield of Nebraska | Sep 2023 - Dec 2023 | Remote

- Designed and executed automated test cases for web applications, APIs, and Power Apps within a .NET (C#) environment, improving test repeatability and coverage

- Used Playwright for browser automation, writing and maintaining complex end-to-end test scripts for enterprise application workflows in a healthcare-regulated environment

SR. SDET

Magnit Global | Jul 2022 - Aug 2023 | Remote

- Led a test automation team across US-based and offshore engineers, improving delivery coordination and quality standards
- Developed a TypeScript/Playwright test automation framework from scratch for functional, integration, and end-to-end (E2E) test coverage
- Analyzed CI/CD results in GitHub Actions and Azure DevOps, identifying risk, reporting defects, and communicating critical coverage gaps
- Early adopter of AI pair-programming for test code (GitHub Copilot, 2022); set team guidelines for reviewing and validating AI-generated tests
- Stabilized test suites across production and non-production environments; mentored junior engineers and drove process improvements that increased quality and delivery velocity

SDET (SOFTWARE ENGINEER IN TEST)

Magnit Global | Feb 2022 - Jul 2022 | Remote

- Led API testing for microservices using Java and RestAssured, expanding service-level coverage across core integrations
- Developed comprehensive API test cases and scenarios, analyzing responses to identify deviations, defects, and bottlenecks; automated repetitive API validation tasks
- Partnered with developers and product managers to strengthen test strategy and feedback loops; monitored API performance and reliability

LEAD SOFTWARE ENGINEER IN TEST

CooperVision | Feb 2021 - Feb 2022 | San Diego Metropolitan Area

- Promoted to QA Automation Lead/SDET after delivering 40% faster test cycles; led a cross-functional automation team (US + offshore) under Agile/Scrum, expanding regression coverage by 60% across the product suite
- Built and expanded the Playwright/TypeScript automation framework with Page Object Model architecture and standardized locator conventions, improving maintainability and team adoption
- Leveraged Docker and Azure DevOps to streamline CI/CD, improve reporting accuracy, and accelerate debugging and issue resolution

SOFTWARE ENGINEER IN TEST

CooperVision | Aug 2020 - Feb 2021 | San Diego Metropolitan Area

- Spearheaded the full migration from C#/Selenium to TypeScript/Playwright, improving test execution speed by 40% and maintainability across 5 projects
- Developed the Test Automation Framework from scratch using Playwright with Page Object Model architecture, reducing test development time by 30% and establishing scalable patterns the team still uses today
- Implemented Docker containers and Azure DevOps CI/CD pipelines, improving build reliability by 25% and reducing debug time by 20%; expanded automation coverage by 35%
- Developed REST API tests in C# that caught 22% more integration defects pre-release, closing validation gaps between UI and backend layers

SOFTWARE ENGINEER IN TEST

OnCore Consulting | Nov 2019 - Aug 2020 | Greater Sacramento

- Architected CI/CD-integrated test automation across 4 cross-functional teams, reducing release cycle time by 30% and establishing the QA engineering standard for the organization
- Led QA test stack strategy as primary engineer for Java-based automation tools, improving test pass rate by 22%; requirements analysis decreased escaped defects by 40%
- Mentored junior automation engineers in Agile testing practices, increasing team productivity by 35% and earning promotion within 8 months

QA AUTOMATION ENGINEER

OnCore Consulting | Apr 2019 - Nov 2019 | Greater Sacramento

- Delivered automated testing solutions for the California State Department of Conservation using Java, Selenium, Maven, TestNG, and Azure DevOps CI/CD, improving release frequency by 25%
- Developed REST API test suites with Java and RestAssured, reducing API regressions by 40%; established Postman/Swagger documentation standards
- Translated complex government business requirements into functional tests spanning UI, database, and API layers, reducing post-release defects by 30%

SDET / QA AUTOMATION ENGINEER

Apple | Jun 2018 - Apr 2019 | San Francisco Bay Area

- Spearheaded cross-platform testing for pre-release Apple.com web pages, including the Apple Card project serving millions of visitors during product launches, improving launch readiness by 30%
- Enhanced the Python/Selenium/PyTest test automation framework and expanded the automated test suite by 40%, ensuring zero-defect tolerance for launch-day pages
- Executed comprehensive cross-browser compatibility testing across all major devices and browsers, reducing cross-browser defects by 25%
- Coordinated with Apple's global localization team to verify translations and regional content across dozens of markets, cutting localization defects by 20%
- Mentored junior QA engineers and enforced code quality through rigorous code reviews, accelerating team onboarding by 35%

QA AUTOMATION ENGINEER

Williams-Sonoma | Jul 2016 - Jun 2018 | San Francisco Bay Area

- Developed automated test suites using Java, Selenium, TestNG, Maven, and Jenkins for a multi-brand e-commerce platform, increasing regression coverage by 40%
- Implemented BDD test suites with Cucumber and Gherkin, turning business requirements into executable specifications and reducing defect leakage to production by 25%
- Identified 120+ critical defects pre-release in payment and checkout flows, preventing revenue loss across millions in daily transactions

AI & OPEN-SOURCE PROJECTS

TOKENDYN0.COM — LLM BENCHMARK HARNESS | 2025 - Present

- Built and operate a public LLM evaluation harness benchmarking model throughput and latency across providers — the evaluation-harness discipline (repeatable runs, golden baselines, regression tracking) applied to AI models themselves

MULTI-AGENT LLM AUTOMATION SYSTEM | 2025 - Present

- Designed and operate a production multi-agent LLM system: n8n orchestration, model routing across multiple LLM providers, prompt guards, deterministic publish gates, audit logging, and human-in-the-loop escalation — autonomous by default, safe by design

OPENCODE-SKILL-CREATOR — AI AGENT SKILL EVALUATION TOOL | 2026 - Present

- Built and published an Apache-2.0 TypeScript tool that creates, tests, and optimizes OpenCode skills using evaluation sets, train/test splits, trigger-accuracy benchmarks, variance analysis, and human review; earned 136 GitHub stars
- Open source: <https://github.com/antongulin/opencode-skill-creator>

ROBINREVIEW.DEV — OPEN-SOURCE AI CODE REVIEW | 2026 - Present

- Built Robin, an MIT-licensed GitHub Action that reviews every pull request with AI, posts inline comments, supports multiple LLM providers, and runs with bring-your-own-key or free-model options
- Live site: <https://www.robinreview.dev/> | Source: <https://github.com/antongulin/robin>

ANTON.QA — AI QA ARCHITECT | 2026 - Present

- Publish weekly on AI testing architecture: LLM evaluation patterns, agentic test design, model-as-dependency practices (pinning, fallbacks, re-verification); first person to claim the "AI QA Architect" title on LinkedIn

SKILLS

AI & LLM Engineering: AI Test Automation, AI Agents / Agentic Testing, AI Agent Orchestration, LLM Evaluation (AI Evals), Evaluation Harness Design, Golden Datasets, LLM-as-Judge Reliability, RAG Evaluation, AI Red-Teaming, AI Guardrails, Hallucination Detection, Prompt Engineering, Generative AI (GenAI) QA, Claude Code, GitHub Copilot, MCP (Model Context Protocol) Server Integration, AI-Augmented Test Generation, Self-Healing Test Design, LLM-Powered QA Workflows

Languages & Frameworks: Playwright, TypeScript, Python, SQL, JavaScript, Java, C#, Selenium, PyTest, TestNG, RestAssured, Cucumber, Gherkin, NUnit, SpecFlow

Testing & Quality Architecture: End-to-End (E2E) Testing, API Testing, Test Automation Framework Design, Test Strategy, Page Object Model, BDD (Behavior-Driven Development), Regression Testing, Functional Testing, Cross-Browser Testing, Localization Testing, Performance Testing, Test Data Management, Test Observability & Reporting, Test Planning, Requirements Analysis, Manual/Exploratory Testing

DevOps & Infrastructure: CI/CD Pipeline Design, GitHub Actions, Azure DevOps, Jenkins, Docker, Kubernetes, Git, Azure Cloud, Linux, Jira, Postman, Swagger, Charles Proxy

Leadership & Strategy: Team Leadership (US + Offshore), Agile/Scrum, Mentoring, Stakeholder Management, QA Strategy, Project Management (MBA), Process Improvement, Cross-Functional Collaboration, Quality ROI Analysis

Domains: Healthcare (Medicare Advantage, HIPAA), E-commerce & Payments, Government, SaaS

EDUCATION

MASTER OF BUSINESS ADMINISTRATION, ECONOMICS AND PROJECT MANAGEMENT

Kazakh University of Economics, Finance and International Trade | 2010 - 2012

BACHELOR OF BUSINESS ADMINISTRATION AND MANAGEMENT

Karaganda State Industrial University | 2003 - 2008

CERTIFICATIONS

PCEP - Certified Entry-Level Python Programmer (PCEP-30-01)

Python Institute | December 2019

https://www.credly.com/badges/0a222958-24cc-4e11-b5e9-349383cf168e/linked_in_profile

VOLUNTEERING

Regular Blood Donor - American Red Cross