

Carlos Vera

AI Quality Engineer & QA Automation Architect | Test & Evaluation Frameworks for AI-Powered Products

📍 Piriapolis, Uruguay · ✉ cvera08@gmail.com ·

🌐 cvera-portfolio.vercel.app · 🌐 cvera08 ·

🌐 carlos-vera-automation-qa



👤 PROFESSIONAL SUMMARY

I've spent 16 years finding what breaks. Now I engineer quality into systems that fail in new ways: AI-powered products.

Playwright · TypeScript · Python · CI/CD with GitHub Actions

LLM & Agent Evaluation · AI Observability · Test Automation Architecture

Remote Contractor (US Timezones)

AI Quality Engineer & QA Automation Architect with 16+ years of experience building test and evaluation frameworks that ship reliable software — including AI-powered products. Designs and implements automated test and evaluation pipelines using **Playwright, TypeScript, and Python**, integrated into CI/CD (GitHub Actions). Applies LLM evaluation techniques including reference-based and reference-free methods, **LLM-as-a-judge, RAG evaluation**, and **adversarial testing** to validate non-deterministic AI systems in production. Known for translating quality risk into engineering solutions: **scalable frameworks, observable pipelines**, and **actionable metrics** across web, mobile, backend, and AI-powered platforms.

Building **#AITestingJourney** — a public R&D log of what actually runs in a production eval pipeline.

🏆 KEY ACHIEVEMENTS

- Designed and deployed **LLM and agent evaluation pipelines** integrated into CI/CD, enabling continuous quality validation for AI-powered products.
- Built test automation frameworks from scratch using **Playwright and TypeScript**, adopted across multiple product teams as the standard automation layer.
- Reduced bug rates by **18%** and increased test automation efficiency by **20%** by leading a 5-person QA team through a full framework overhaul.
- Cut test execution time by **40%** (from 2+ hours to under one) by parallelizing automation suites across four concurrent runners.
- Unblocked **400+ regression tests** by engineering a 2FA/MFA bypass solution for authenticated test flows.
- Reduced environment setup time by **50%** through a data generation framework using GraphQL to automate dynamic test data creation.

- Increased test efficiency by **27%** by building a new Playwright framework replacing a legacy stack.
- Applied **adversarial testing and red-teaming** techniques to validate non-deterministic AI outputs, contributing to production observability for LLM-based systems.
- Achieved **23% decrease in post-release defects** at Socure by establishing automated mobile testing on native iOS using XCUITest and real device coverage.

WORK EXPERIENCE

Independent Contractor / Consultant [AI Quality Engineering & Research]

 Remote

Quality Engineering for AI Systems — Personal R&D

Oct 2023 - Present

- Built an **evaluation framework for LLM-powered agents** using Python and Evidently AI, covering reference-based evals, LLM-as-a-judge, RAG evaluation, adversarial testing, and agent tracing wired into a CI/CD pipeline.
- Designed and ran **red-teaming and adversarial test suites** against LLM outputs, validating prompt reliability, hallucination boundaries, and failure modes across generative model pipelines.
- Integrated **AI observability and eval logging** into automated workflows using GitHub Actions, translating experimental findings into reusable test patterns for AI-powered products.

SmartEquip [Equipment Lifecycle & Procurement Platform]

 Norwalk, Connecticut, U.S.

Senior Quality Engineer / AI-Powered Products & Test Automation

Jan 2025 - Present

- Own **automation strategy, framework architecture, and test coverage** across procurement and marketplace platforms using Playwright, TypeScript, Python, TestRail, and GitHub Actions.
- Design and maintain **scalable end-to-end automation solutions** integrated into CI/CD pipelines, supporting reliable and continuous delivery across multiple applications.
- Leverage **AI-assisted development workflows** using GitHub Copilot, Claude, and MCP-enabled tooling to accelerate framework evolution, debugging, and quality engineering delivery.
- Define **quality validation approaches** for AI-generated software artifacts and autonomous development workflows, reducing false positive rate in AI-generated test suggestions by **30%** and cutting manual review time of AI outputs by **35%** through automated eval pipelines.
- Apply **LLM evaluation techniques** and intelligent validation workflows to assess AI-powered product behavior, bridging traditional test automation with emerging quality engineering practices for production AI systems.

Blackpoint [Cybersecurity]

 Ellicott City, Maryland, U.S.

Automation Technical Lead / Staff Engineer

Feb 2024 - Jan 2025

- Led a project that cut software bugs by 18%, improving overall product reliability across cybersecurity platform releases, and managed a team of five Automation QAs to boost delivery efficiency by 20%.

- Developed a solution to bypass **2FA/MFA** codes for logged-in users, unblocking over 400 tests and increasing regression test efficiency by 70%.
- Boosted efficiency by **parallelizing** four automation suites, cutting test execution time by 40%, from over two hours to almost one.
- Implemented a **data generation** framework using GraphQL, automating the creation of dynamic test data (accounts, customers, billing models, logos, Stripe integrations, etc) and reducing environment setup time by 50%.

DevOps Automation Specialist

Apr 2023 - Mar 2024

- Recognized for **building** a new testing **framework** with Playwright that led to a 27% increase in test efficiency.
- Implemented a **CI/CD** pipeline using **GitHub Actions**, enhancing code quality and boosting deployment frequency by 30%.
- Developed integration between **Playwright** test executions and **Xray** + Jira, enabling full traceability between automated results and test plans.
- Set up contract testing with Pact, ensuring reliable API boundaries between services.

Socure [Identity Verification & Fraud Prevention]

📍 Florida City, Florida, U.S.

SDET (iOS - Native Mobile)

Jul 2019 - May 2023

- Automated testing of a biometric face & ID recognition app using Xcode, covering both simulator environments and real devices (iPhone, iPad).
- Improved overall product quality, reducing post-release issues by **23%** through systematic test coverage across identity verification and fraud prevention workflows.
- Established **Jenkins-based CI pipelines** for automated iOS builds, streamlining test execution and enabling continuous quality validation.

Renault & Nissan [Automobile Manufacturer]

📍 London, England, U.K.

QA Technical Lead (of Manual, Hybrid & Automation)

Oct 2015 - Jul 2019

- Acknowledged for leading a QA team that delivered a critical product release **17%** faster than planned, coordinating across engineering, product, and client stakeholders.
- Managed **QA teams** across manual, hybrid, and automation streams, defining test strategy, tracking performance metrics, and aligning quality coverage with product and engineering goals.
- Built and structured the internal **QA hiring process**, interviewing candidates and onboarding new team members into automation practices.

Webee [IoT / Internet of Things]

📍 Sunnyvale, California, U.S.

Architect Quality Engineer (Android - Mobile Responsive)

May 2014 - Oct 2015

- Reduced testing time by **68%** by designing and implementing efficient automation processes for Android mobile applications on an IoT platform.
- Established **automation candidate selection criteria and acceptance frameworks** using Cucumber, bridging manual test suites and automated coverage.

PwC [Audit & Tax Advisory]

📍 San Francisco, California, U.S.

Software Engineer in Test (Back-end Web)

Jan 2013 - May 2014

- Built a **Backend Automation Framework** from scratch to support API development and validate user stories through automated test coverage.
- Integrated Postman with **CI/CD pipelines** using Newman, enabling automated API execution on every build.

EY [Accounting]

📍 Chicago, Illinois, U.S.

Tester QA (Front-end Web)

Dec 2010 - Jan 2013

- Executed **functional and exploratory testing** across front-end web applications, gathering requirements and analyzing regression cycle results to maintain product quality.
- Maintained and fixed outdated **Selenium-based automation tests**, marking an early shift from manual toward automated quality practices.

NEPS Solutions [Software Factory]

📍 Cordoba, Argentina

Jr. Developer (Front-end Web)

Jan 2010 - Feb 2011

- Developed front-end web features including **shopping cart flows** using HTML, CSS, and JavaScript, applying MVC and design patterns in a startup environment as the company's first hire.
- Increased **test coverage by 25%** by implementing unit testing strategies with Jasmine, establishing early quality practices from the ground up.

PERSONAL PROJECTS

AI Stock Predictions [OpenAI dev]

Oct 2024-Present

AI-Enhanced **Trading Suggestions**.

Live Demo - github.com/assets/stock-predictions-demo

View GitHub Source - github.com/cvera08/ai-stock-predictions

Web Based AI Games [AI dev]

Mar 2024-Nov 2024

Multi Games by using **Artificial Intelligence**.

Live Demo - cvera08.github.io/multi-games-artificial-intelligence-js

View GitHub Source - github.com/cvera08/multi-games-artificial-intelligence-js

Blockchain QA Automation [Automation Architect]

Jan 2023-Jul 2023

Developed **Smart Contract Test Automation** by using Solidity, Remix, JavaScript, Cypress.

Deployed, debugged, and tested Ethereum and EVM-compatible smart contracts.

Live Demo - ibb.co/SmartContractCypress

View GitHub Source - github.com/cvera08/blockchain-automation

TOOLS EXPERIENCE

- **AI & LLM Evaluation:** Evidently AI, LangSmith, Langfuse, Ragas, Arize Phoenix.

- **AI-Assisted Development:** GitHub Copilot, Claude, Cursor, Codex, Gemini, ChatGPT & OpenAI APIs.
 - Leveraged in testing workflows: test generation, prompt engineering, LLM evaluation, and failure analysis.
- **Test Automation Frameworks:** Playwright, Selenium, Cypress; TypeScript, JavaScript, Python.
- **CI/CD & Observability:** GitHub Actions, Docker, Kubernetes, Prometheus, Grafana, Datadog, Elasticsearch.
- **API & Back-end Testing:** Postman, GraphQL, REST-assured, ReadyAPI.
- **Test Management & Reporting:** TestRail, Allure, Monocart Reporter, Zephyr, Xray.
- **Languages:** TypeScript, JavaScript, Python, Java, Ruby.
- **Cloud & Infrastructure:** AWS (EC2, S3, Lambda), GitHub, GitLab, Bitbucket.
- **Project Management:** Jira, Confluence, Trello.
- **Performance Testing:** k6, JMeter.
- **Mobile Automation:** Appium, XCUITest, TestFlight, Fastlane.
- **Databases:** MySQL, MongoDB.
- **IDEs & Dev Tools:** Visual Studio Code, IntelliJ, PyCharm, Warp, Git CLI (Bash), Zsh.

EDUCATION

LLM Evaluation for Builders – Applied Course 06/2026

Evidently AI. Hands-on LLM evaluation: custom LLM judges, RAG evaluation, adversarial testing, and evaluation pipelines integrated into CI/CD workflows.

LLM Evaluation Course – Concepts & Methods 06/2026

Evidently AI. Core evaluation concepts: metrics design, hallucination detection, output quality assessment, and systematic LLM testing methodologies.

Agent Skills with Anthropic 05/2026

DeepLearning.AI. AI agent design patterns, multi-agent coordination, and tool use with Anthropic's Claude for building production-ready AI systems.

Accelerate app development by using GitHub Copilot 02/2026

Secure Coding – OWASP Top 10 for LLM Applications 08/2025

ChatGPT Prompt Engineering for Developers 05/2025

Intro to AI Engineering 09/2024

Artificial Intelligence with Python 05/2024

Intro to Artificial Intelligence 04/2024

Introduction to Generative AI Studio 02/2024

English - Certified IELTS 03/2020

Competent User - Listening, Reading, Writing & Speaking. Effective language command.

IDP International English Language Testing System Australia. Candidate: 003239 - Centre:

AR630.

ISTQB Certified Tester - Foundation Level

06/2014

Pearson VUE - International Software Testing Qualifications Board. Certificate Number: 14-CTFL-57115-HA.

System Engineering - UTN FRC

01/2013

UTN FRC: Universidad Tecnológica Nacional - Facultad Regional Córdoba.

LANGUAGES

English [Advanced] / Spanish [Mother Tongue]

HOBBIES

- Family Time [Spending quality moments with my loved ones]
- AI Testing Research [Exploring AI trends and sharing learnings via #AITestingJourney]
- Fishing Enthusiast [Enjoying peaceful moments by the water while fishing]

ADDITIONAL INFORMATION

Let's Connect and Drive Quality together!

- [linkedin.com/carlos-vera-automation-qa](https://www.linkedin.com/company/carlos-vera-automation-qa)