



# Steven Leon

📍 Queens, NY 11385 📞 347-612-7281 ✉️ [Steven\\_leon@live.com](mailto:Steven_leon@live.com)

🔗 <https://github.com/stevenleon30>

🔗 <https://www.linkedin.com/in/leonsteven>

## TECH STACK

**LLM Platforms:** OpenAI, Anthropic Claude  
**Evaluation:** Promptfoo, YAML Configurations  
**Automation:** Playwright (UI & API tests), VS Code  
**Version Control:** GitHub, GitHub Copilot-assisted scripting  
**Languages:** JavaScript, HTML, CSS, JSON  
**Testing Tools:** BrowserStack, Jira  
**DevOps:** CI/CD workflows, release gating

## EDUCATION

**Bachelor of Arts** | Liberal Arts And General Studies

**Adelphi University, Garden City, NY**

Earned a Bachelor of Arts in Liberal Arts & Theater Studies at Adelphi University, focusing on performance, creative expression, and communication. Gained experience in stage production, improvisation, and collaboration while developing strong presentation and storytelling skills applicable to both artistic and professional settings.

**Certificate of Technical Studies** | Quality Assurance And Software Testing Principles

**Per Scholas, New York, NY**

Gained QA fundamentals in test planning, case design, and execution for manual testing. Hands-on with bug tracking in Jira/Trello, exploratory and functional testing, and cross-browser/platform validation. Experienced in Agile/Scrum, test documentation, and reporting, with strong collaboration and problem-solving skills to ensure quality software delivery.

AI Test Automation of Large Language Models

10/2025

**Engenious University, Online**

Through Engenious University, I've gained hands-on experience with Promptfoo for test automation, reporting, and red team evaluations. I've set up environments via CLI (Homebrew), created YAML and .env files in VS Code, and modified configurations to optimize LLMs (temperature, cost, latency). I've also tested models for factuality, hallucinations, and reasoning, while creating and refining system prompts for better alignment.

## CORE COMPETENCIES

**Conversational AI** Regression Frameworks  
**LLM Evaluation** & Hallucination Detection

**PROFESSIONAL  
EXPERIENCE**

**Prompt Engineering & System Prompt Optimization**  
**Playwright Automation Testing**  
**GitHub-Based CI Validation**  
**JSON Schema & Structured Output Validation**  
**API & Integration Testing**  
**Risk-Based Release Gating**

**AI Reliability & Conversational QA Engineer** | Engenious University  
AI Systems & LLM Projects | 2025–2026

- Designed and executed 150+ structured LLM regression tests across safety, ambiguity handling, hallucination detection, and structured output enforcement.
- Built Promptfoo evaluation pipelines with rubric-based scoring and release thresholds ( $\geq 90\%$  compliance pass rate).
- Conducted 50+ adversarial prompt simulations to detect injection and unsafe output vulnerabilities.
- Tuned model parameters to reduce output variance and improve determinism by  $\sim 25\%$ .
- Implemented structured JSON validation checks ( $\geq 99\%$  schema integrity) for downstream system reliability.
- Developed Playwright-based browser automation tests validating AI-generated UI workflows and integration stability.
- Integrated automated tests into GitHub workflows for pre-release validation and regression detection.
- Used GitHub Copilot within VS Code to accelerate automation scripting and improve test coverage efficiency.

**Quality Technician** | Boyce Technologies  
Oct 2025 - Present | New York, New York

- Inspect laser-cut, waterjet, and fabricated metal components for surface defects, dimensional accuracy, and overall workmanship.
- Review shop drawings, blueprints, and engineering specifications to verify design intent, material requirements, and critical dimensions.
- Use precision tools (calipers, micrometers, gauges) to confirm tolerances and alignment with technical drawings.
- Verify and document product dimensions, weights, and calculations to ensure consistency across production batches.
- Perform in-process quality checks to detect deviations early and prevent rework or nonconforming parts.
- Work with production, engineering, and fabrication teams to resolve discrepancies, apply corrective actions, and maintain quality standards.

- Conduct root-cause analysis on defects, identify trends, and recommend long-term process improvements.
- Maintain detailed QA reports, inspection logs, and nonconformance records for traceability and compliance.
- Ensure all manufactured and assembled components meet company, client, and regulatory safety requirements before release.
- Perform field quality inspections on MTA subway projects to confirm installed systems meet design, safety, and performance criteria.
- Support continuous improvement by providing feedback from field and manufacturing inspections to enhance design and production efficiency.

### **Quality Assurance Analyst | Creative Circle**

March 2025 - June 2025 | New York, New York

- Execute test cases to validate functionality, usability, and responsiveness across devices.
- Ensure smooth user experience and stability after updates on Vercel.
- Conduct cross-browser testing utilizing BrowserStack.
- Proficient in digital review tools such as Jira, GitHub, and Canva.
- Document and track bugs in GitHub with clear reproduction steps and severity levels.
- Strong proofreading and content validation skills, with over 10 years ensuring written and visual materials meet brand and style standards.
- Re-test resolved issues to prevent regressions and maintain core functionality.
- Identify and report performance issues, collaborating with developers for optimization.
- Experience producing, editing, and validating branded marketing materials, including email campaigns, landing pages, and social assets.
- Maintain and refine test cases, findings, and best practices for continuous improvement.

### **Senior Quality Assurance Analyst | Vox Media Inc**

Oct 2016 - Dec 2023 | New York, NY

- Reviewed pull requests, test cases, and updated feature logic within GitHub repositories. Collaborated with cross-functional teams to ensure that every code change adhered to testing protocols.
- Executed detailed testing of the Verge website hosted on the Vercel platform by assessing core functionality, page load times, image rendering, and overall user experience. Identified performance bottlenecks and collaborated with developers to troubleshoot and optimize the website for seamless, cross-device usability.

- Maintained issue tracking in Jira, ensuring clear documentation, attachments, and accurate assignment of tickets.
- Supported company growth by participating in interviews, hiring, and onboarding of new team members.
- Collaborated with project managers and tech leads to define project scope, prioritize issues, and align sprint goals.
- Identified and reported visual inconsistencies, user experience issues, and usability problems, enhancing the CI/CD pipeline for The Verge team.
- Demonstrated new product features and functionality to project stakeholders.
- Produced and maintained technical documentation for project reference.
- Proposed development and testing improvements to positively impact usability, function and performance.
- Supported the successful launch of numerous product releases by coordinating QA efforts, including test planning, execution, and reporting.
- Participated in available trainings to improve software knowledge.

**Quality Assurance Coordinator | Thrillist Media Group**

Mar 2015 - Feb 2016 | New York, NY

- Maintained and updated detailed test matrices in MS Excel, refining test cases and data to ensure accuracy and effectiveness.
- Collaborated with senior designers to review mockups, validate builds against designs, and confirm testing requirements.
- Participated in daily Scrum meetings to plan testing activities, coordinate deployments, and address high-priority updates.
- Performed exploratory and edge-case testing to identify potential application issues outside defined scopes.
- Utilized BrowserStack for cross-browser testing on macOS and Windows across Chrome, Firefox, Safari, and relevant browsers.