Devarshi Ojha

Lead SDET | AI · Web · API · NFR

career.devarshi@gmail.com | Gurugram, India | linkedin.com/in/devarshi-ojha | github.com/code-vault

SUMMARY

Lead SDET with 9+ years of experience owning quality across **AI/ML, web, mobile and API systems**. Designs test strategies and automation frameworks in **Playwright (TypeScript)**, **Selenium**, **REST Assured and BDD**, integrated into CI/CD quality gates. **Owned and scaled AI evaluation frameworks** for production AI document-processing pipelines using golden datasets, per-class precision/recall/F1 and field-level validation. Experienced in performance and load testing, AWS and Azure cloud environments, and Agile (Scrum/SAFe) delivery. Set QA standards through a Quality CoE adopted by multiple business verticals.

EXPERIENCE

KlearNow.AI, *Lead Quality Assurance Engineer (AI Ops)* 06/2026 – 09/2026 | Gurugram

- Led a QA team of ~6 owning test strategy, test plans, risk-based testing and release sign-off for production GenAI document classification and extraction.
- Owned and strengthened the team's existing AI evaluation framework: golden dataset of ~2,000 real and 400–500 synthetic documents (8 classes, ~100 fields), with per-class precision/recall/F1, field-level exact-match and JSON schema checks as CI quality gates, catching region-specific misclassification before release.
- Cut evaluation run time from 5–6 hours to ~3 hours and made results easier to interpret and scale; piloted DeepEval (pytest-style LLM test cases in CI), RAGAS (retrieval and context-quality metrics) and Promptfoo (prompt and model comparison across document types), and integrated Arize for production observability (run results, quality-trend visualisation).
- Introduced spec-driven agentic test generation with Playwright Test Agents (TypeScript) and an Atlassian MCP server, creating ~500 test cases from Jira stories and Confluence specs and halving manual test design per sprint.
- Built REST Assured + TestNG API and integration tests for microservices and Kafka workflows, plus a parallel Appium framework for mobile.

Career Break, *Family Care, Upskilling & Freelance* 02/2025 – 06/2026

- Took a planned break to manage a family medical emergency (02/2025 – 01/2026), then upskilled in GenAI evaluation, Playwright agentic QA and AI-assisted engineering (02/2026 – 06/2026).
- Delivered 4 short freelance engagements for Southeast Asian clients via freelance platforms, covering Playwright test automation and API testing.

StatusNeo Consulting Pvt Ltd, *Senior Consultant* 07/2022 – 01/2025 | Gurugram

- Defined QA standards, best practices and reusable automation frameworks (Maven/npm) in the Quality CoE, adopted by 16 business verticals. Personally onboarded 7 verticals and trained 20+ QA engineers. The framework was later ported into StatusNeo's TestCraft product.
- Avis Budget Group mainframe-to-AWS modernization: built the test strategy for moving 6,000+ legacy test cases to ~1,200 layered tests across platform, experience, UI, database and GraphQL.
- Architected the UI framework in Selenium, then led the migration to Playwright (TypeScript), cutting full regression from 8 hours to 3 and flakiness by ~60%, with Docker-based CI infrastructure.
- Owned API and integration testing of core microservices: rate and currency APIs, Kafka reservation flows, PostgreSQL, and validation of bulk-replicated legacy DB2 data.
- Ran performance and load testing with JMeter and k6, and monitored with Grafana and Dynatrace across AWS environments.

Publicis Sapient, *Senior QE Associate* 06/2019 – 06/2022 | Noida

- Sole QA for a headless AEM CMS. Authored the test framework from scratch, covering author workflows, publisher APIs, dispatcher, CDN caching, and publisher and dispatcher performance.
- Contributed to mobile automation (Appium) across Android and iOS devices on BrowserStack and real devices.

I3infosoft Pvt Ltd, *QA Consultant* 04/2016 – 06/2019 | Noida

- Automated functional and regression suites with Selenium (Java) + TestNG for e-commerce web apps.

SKILLS

Automation: Playwright (TypeScript) · Selenium (Java) · REST Assured · Appium · Cucumber/BDD · TestNG · Pytest · GraphQL · Kafka

API & Integration: Postman/Newman · Swagger/OpenAPI · WireMock

AI/ML & GenAI Testing: Golden datasets · Precision/Recall/F1 · Confusion matrices · Field-level validation · JSON schema validation · LLM-as-judge · Prompt regression · Model comparison · CI evaluation gates · DeepEval · RAGAS · Promptfoo · Arize

Agentic & Spec-Driven QA: Playwright Test Agents · MCP (Playwright, Atlassian) · Claude Code · Cursor

NFR: JMeter · k6 · OWASP ZAP (basic) · Applitools · Lighthouse · Datadog · Grafana · Dynatrace · CloudWatch

CI/CD & Cloud: GitHub Actions · Jenkins · Azure DevOps · Docker · Kubernetes · AWS · Azure

Devices & Mobile: BrowserStack · LambdaTest · Sauce Labs · Appium Inspector · Android Studio/Xcode · Charles Proxy

Data & Reporting: PostgreSQL · pgAdmin · MongoDB · Kafka UI tools · Allure · Extent Reports

Test Management & Agile: Jira (Xray/Zephyr) · Azure Test Plans · Confluence · Scrum · SAFe

QA Practice: Test strategy and planning · Risk-based testing · Defect leakage control · Release sign-off · QA standards and CoE

Languages: TypeScript/JavaScript · Java · Python · SQL

EDUCATION

B.Tech, Computer Science, *AKTU* 2016