

Ismail Isleem

Senior SDET · Test Infrastructure & AI/LLM Testing

Amman, Jordan +962 78 010 6088 ismail.isleem@gmail.com
[linkedin.com/in/ismailisleem](https://www.linkedin.com/in/ismailisleem) github.com/ismailisleem ismailisleem.com
Open to remote opportunities globally · Based in Jordan

PROFESSIONAL SUMMARY

Senior SDET with 13+ years building test infrastructure, internal developer tools, and full-stack QA platforms — not just consuming frameworks, but designing them. Hands-on coder across Java, Python, JavaScript, TypeScript, and C#, spanning backend APIs (.NET Code-First, Entity Framework), frontend web apps, automation frameworks (Selenium, Appium, Playwright, RestAssured, Robot Framework), and CI/CD. Specialized in AI/LLM testing — custom evaluation pipelines for hallucination detection, RAG evaluation, semantic similarity, and LLM-as-a-judge architectures.

TECHNICAL SKILLS

Programming Languages	Java, Python, JavaScript, TypeScript, C#, Groovy
Software Development	ASP.NET (Code-First, Entity Framework), Service-Oriented Architecture, REST APIs, full-stack web (HTML/JS/Bootstrap), web scraping pipelines
Test Frameworks & Tools	Selenium, Appium, Playwright, Cypress, Robot Framework, RestAssured, Postman, Autolt, JUnit, TestNG, Cucumber (BDD)
AI/LLM Testing	Prompt testing, hallucination detection, context relevance & faithfulness scoring, RAG evaluation, semantic similarity (embeddings), LLM-as-a-judge
AI/ML Libraries & APIs	LangChain, Hugging Face Transformers, OpenAI API, Anthropic Claude API
Architecture & Design	Page Object Model, Factory, Singleton, DTO/POJO, BDD, parallel execution, fault isolation
CI/CD & DevOps	Jenkins, Docker, GitHub Actions, Bitbucket Pipelines, Maven, Gradle
Collaboration	Jira, Azure DevOps, Confluence, Git, Agile/Scrum, code review at scale
Databases	SQL (MySQL, PostgreSQL, SQL Server), NoSQL basics

PROFESSIONAL EXPERIENCE

QPros — Senior SDET | Technical Lead

Feb 2023 – Aug 2026

Amman, Jordan (Fully Remote)

- Develop AI-powered test data generation systems using LLMs (ChatGPT & Claude APIs) with a dual-agent architecture — one agent generates synthetic positive/negative test data, a second validates outputs and provides structured feedback — using prompt engineering for deterministic, high-quality results.
- Build internal developer tools and Python testing utilities with LangChain, Hugging Face, OpenAI API, and transformers, exposing reusable libraries and CLI utilities consumed by automation engineers across the organization.
- Design internal QA service APIs exposing test data generation and dataset retrieval endpoints, applying DTO/POJO patterns for clean contracts; consumed by web and mobile automation pipelines.
- AI/LLM Testing Specialist: engineer custom evaluation pipelines for LLM-based systems — prompt testing, hallucination detection, context relevance & faithfulness scoring, RAG evaluation, semantic similarity via embeddings, and LLM-as-a-judge.
- Architect open-source test automation frameworks for enterprise clients across web, mobile, API, and desktop — coding directly in Selenium, Appium, Playwright, Robot Framework, RestAssured, and Autolt.
- Lead in-house automation tools and QA platforms, directing a cross-functional engineering team through architecture, sprint planning, code reviews, and release readiness.
- Conduct yearly code audits across the organization and at major client engagements, evaluating framework design and design-pattern adherence (Factory, POM, Singleton); mentor engineers to raise code quality.

Mawdoo3 — Senior SDET | Team Lead

Sep 2018 – Nov 2022

Amman, Jordan

- Established the Quality Engineering function from the ground up within the AI/NLP department — defining testing strategy, frameworks, and tooling for Arabic language models and conversational AI, an emerging space with limited off-the-shelf Arabic tooling.
- Designed and built a full-stack ML model evaluation platform end-to-end: C#/.NET backend (ASP.NET, Code-First with Entity Framework) with a service-oriented API architecture for fault isolation, plus a JavaScript/HTML/Bootstrap no-code frontend used by manual QA engineers.
- Architected the evaluation engine supporting Word2Vec embeddings, POS tagging, NER, text classification, categorization, and normalization; ran datasets in parallel on Jenkins with configurable thread limits, computing Accuracy, Precision, Recall, F1 (Macro/Micro/Binary), and Confusion Matrix.
- Engineered web scraping pipelines for automated data collection, filtering, and categorization — generating internal training datasets that reduced reliance on third-party data purchases.
- Developed a rule-based Arabic text normalization and categorization engine in C#, used as a preprocessing layer for NLP models and a standalone service for the data science team.
- Extended the platform to test fine-tuned chatbots and conversational agents, including a production chatbot deployed for a banking client.
- Code review gatekeeper for the shared automation repository — approving all merges to master, running cross-division training, and mentoring manual QA engineers to run model evaluations independently.

Winfooz — Senior Quality Assurance Engineer

Jul 2017 – Jun 2018

Amman, Jordan

- Owned full QA coverage (manual + automation) for an online vehicle auction platform serving the US and Canadian markets, reporting to the QA Manager.
- Rebuilt the automation infrastructure across three layers: web (Selenium/Java), mobile (Appium — Android & iOS), and JavaScript web testing (Cypress), applying reusable design patterns.
- Led the migration from Cypress to Playwright with Python, modernizing the web automation stack and improving stability and execution speed.
- Designed and built an in-house testing lab (dedicated servers and physical mobile devices) replacing cloud platforms — delivering significant cost savings while improving reliability and control.

Fetcher — Senior Quality Assurance Engineer

Sep 2016 – Jun 2017

Amman, Jordan

- Owned the full QA lifecycle (manual + automation) for a logistics platform powering warehouse operations, slot management, inventory, and inter-service REST integrations.
- Designed two complementary API automation frameworks: a lightweight Postman collection for fast developer feedback, and a full-scale Java framework (Maven, TestNG, RestAssured) covering functional, regression, and smoke suites.
- Architected a cross-platform mobile automation framework (Appium, Java, Maven, TestNG) enabling unified Android/iOS execution.
- Migrated the suite to a BDD model with Cucumber, improving readability and validating integration flows between microservices for data consistency and contract integrity.

7AWI — Quality Assurance Engineer

Mar 2015 – Aug 2016

Amman, Jordan

- Owned end-to-end QA for a high-traffic online classifieds platform — web functionality, user flows, and release reporting.
- Designed and implemented the company's first full web automation framework using Selenium and Java, establishing reusable components and structured test suites.
- Built API test automation for the search layer (Apache Solr) using Postman and JavaScript, and delivered weekly QA reports via Jira dashboards.

AnnabSoft — Quality Assurance Engineer

May 2013 – Feb 2015

Amman, Jordan

- Tested a US-based e-commerce platform and custom ERP covering warehouse management, inventory, and third-party shipping integration (DHL).
- Introduced the company's first automation initiative — Selenium/Java scripts for regression-heavy flows — transitioning into a hybrid manual/automation role; authored test cases and ran full regression cycles per release using Mantis.

SELECTED PROJECTS

AI Test Data Generator

github.com/ismailisleem/testdata-factory

Local-first AI tool that generates schema-valid synthetic test data through a dual-agent architecture — one agent produces positive and negative datasets, a second validates the output and returns structured feedback.

Python • FastAPI • Local LLMs • JSON Schema

Automation Framework Suite

github.com/ismailisleem

A shared core library powering three automation frameworks — web (Playwright), mobile (Appium), and API (httpx/GraphQL) — with parallel execution, reusable components, and unified reporting.

[automation-core](#) · [web-automation-framework](#) · [mobile-automation-framework](#) · [api-automation-framework](#)

Python · Playwright · Appium · httpx · Pytest · Allure

EDUCATION & CERTIFICATIONS

B.S. in Software Engineering — Al-Zaytoonah University, Jordan

Feb 2008 - Jul 2012

ISTQB Certified Tester, Foundation Level (CTFL) — Cert No. CTFL-JOR-54-2014

LANGUAGES

Arabic — Native · English — Professional Working Proficiency