Nancy Singh

Testing tools & AI reliability | Delhi, India | Open to remote and relocation
singhnancy2004@gmail.com | +91 98211 86608 | nancy6547.github.io
github.com/nancy6547 | linkedin.com/in/nancy-singh-84367722a | medium.com/@singhnancy2004

Education

Sant Longowal Institute of Engineering and Technology (SLIET) — Punjab, India
B.E. Computer Science and Engineering (lateral entry) – CGPA 7.28/10; SGPA 5.70 → 8.60 (sem. III–VII) — 2023–2026

Delhi Skill and Entrepreneurship University (DSEU) — New Delhi, India
Diploma in Digital Electronics

Projects

Zeitaku | *TypeScript, TypeScript Compiler API, Playwright, Chrome DevTools Protocol, Vitest* — GitHub

- Built a codemod on the **TypeScript Compiler API** that finds hard-coded `waitForTimeout` sleeps in Playwright tests and splices in network-aware waits while preserving the file's formatting.
- Replaced fixed sleeps with **Chrome DevTools Protocol** request tracking: on throttled Slow 3G, fixed 2 s and 3 s sleeps failed **100%** of runs while Zeitaku failed **0%**; **45% faster** than a safe fixed wait on a normal network.
- Scanned **7 open-source codebases (76k files)** and found **240** hard-coded sleeps; AST classification leaves alone the **44** that guard "should not happen" assertions.

Junban | *TypeScript, Chrome DevTools Protocol, Playwright, delta debugging* — GitHub

- Built a tool that finds the **order of API responses** that breaks a flaky Playwright test: it holds each response in the browser (CDP `Fetch` domain) and releases them in chosen orders.
- On races that appear only a few times in a thousand runs, it caught **6/6** planted bugs by the **5th run** each, while plain reruns caught **2/6 in 3,000 runs**.
- Shrinks each failure to one rule with **delta debugging** and writes a Playwright repro test that failed **20/20** times for every bug.

DocMind | *Python, FAISS, sentence-transformers, Tesseract OCR, Llama 3.3 (Groq), Streamlit* — Live Demo | GitHub

- Built and deployed a **RAG** app that answers questions about any PDF, running **OCR only on pages without a text layer** so digital PDFs process in seconds.
- Added hybrid **BM25 + embedding** search: the right chunk lands in the top 3 for **96%** of eval questions vs **88%** with embeddings alone (scanned pages: 50% → 83%), plus a relevance guard that skips the LLM when nothing matches.

HalluGuard + Kensa | *Python, PyTorch, Hugging Face, NLI* — Live Demo | GitHub | Kensa

- Built a per-sentence hallucination detector for RAG answers; moved from cosine similarity to an **NLI cross-encoder (DeBERTa-v3)**, benchmarked on **RAGTruth** with thresholds tuned on the train split only.
- Built **Kensa**, a metamorphic test runner for RAG apps (reword, add distractor, delete evidence, edit the fact) that delta-debugs each failure to the minimal documents causing it.

Open Source (10 PRs merged)

Google Puppeteer | *TypeScript, Node.js, CDP, WebDriver BiDi* **(4 merged)** — #15587 | #15590 | #15595 | #15697

- Added `wsOptions.handshakeTimeout` so `connect()` fails fast on a silent host; `page.setGeolocation(null)` removes the override over CDP and BiDi (requested by the maintainer); `request.respond()` skips undefined headers on CDP; BiDi `respond()`/`abort()` no longer reject for requests that are already gone. Each with tests.

Google Chrome DevTools MCP | *TypeScript, Puppeteer, CDP* **(4 merged)** — #3148 | #3144 | #3198 | #3199

- Promise rejections show their real reason; Lighthouse explains unsupported URLs; redirect chains no longer list later hops; console args (Maps, Sets, DOM nodes, cycles) no longer print {}. #3199 was selected for the **XERJ Corpus Hub**, a library AI coding agents search before writing code.

Hugging Face sentence-transformers | *Python, PyTorch, learning-to-rank losses* **(1 merged)** — #4204

- Fixed `PListMLELambdaWeight` passing the padding mask instead of rank positions to a custom discount fn, which removed the position discount and zeroed the loss of shorter lists; shipped with a failing test.

Confident AI DeepEval | *Python, LLM evaluation* **(1 merged)** — report #3469 | fix #3470

- JSON datasets saved by DeepEval itself crashed on reload (null `tools_called`); reported and fixed it with a round-trip test.

Bugs found, fixed upstream | *Microsoft Playwright, Twenty (YC), Documenso* — #43459 | #28342 | #3445

- Playwright codegen double-escaped `${`, crashing generated tests; Twenty's Links field dropped slashes in query/hash; Documenso's E2E suite had redundant sleeps (found by Zeitaku). Reported with root causes; fixed by maintainers.

Technical Skills

Testing & Browser: Playwright, Chrome DevTools Protocol, Vitest, GitHub Actions, TypeScript Compiler API (AST)
ML & RAG: PyTorch, Hugging Face, sentence-transformers, FAISS, LangChain, NLI cross-encoders, TensorFlow Lite
Languages & Tools: TypeScript, Python, Kotlin, JavaScript, Git, Node.js, Express, Streamlit, Tesseract OCR, Linux

Achievements

RedBus Data Decode Hackathon: Top 500 of 8,000+ teams | **NCC "C" Certificate** (National Cadet Corps)