SURYA PRASANTH R

AI engineer · systemic orchestrator of multi-agent pipelines · agent and database correctness

suryaprasanthcse.netlify.app · suryaprasanthcse@gmail.com · github.com/suryaprasanthcse · linkedin.com/in/suryaprasanthcse

Tiruppur, India · Open to remote contract work and full-time remote roles

SUMMARY

I orchestrate agents. I design and run multi-agent pipelines that build software, and I own the parts the agents can't be trusted with. I write the specifications, split the work across agents, decide where the boundaries go, and build the checks that prove the result works; Claude Code writes most of the code under my direction. My recent work tests what actually holds when many agents hit the same database at once, and I report what I find upstream with minimal reproductions.

AWARDS

1st place	Game Tech track, Cambridge × Arcade AI Hackathon, Oct 2026 · GOLEM (solo)
3rd place	Agentic AI track, HackMatrix 2026 · SYNAPSE-3D

CORRECTNESS AND OPEN SOURCE

Witnessmark · concurrency test harness for graph databases and agent memory — Nov 2026

Open source, Apache-2.0 · Python, Docker, FalkorDB 6.0.4, Neo4j 5.26.2 · findings on Graphiti, Omnigraph, Cognee, HydraDB · repo

- Releases many clients at the same instant against Graphiti, Cognee and Omnigraph, records every operation, and checks invariants such as single winner, no lost acknowledged writes and deletion permanence, each with sequential, negative and positive controls.
- Five reviewed milestones: 243 unit tests, 21 reports that regenerate byte for byte from raw evidence, and every published number traced to a recorded run id.
- Found a data-loss race in Graphiti (19 of 20 runs on Neo4j) and wrote the fix and regression test, PR #2177 (in review; an independent check measured 0 losses in 200 runs).
- Omnigraph's maintainers reproduced and fixed a refusal-type mismatch I reported (#1426); Cognee left duplicate vector ids in 13 of 14 trials.

FalkorDB · query-planner fix and engine reports — Oct – Nov 2026

Rust · PR #3699 (under review) · issue #3694

- Found a planner bug where a WHERE after LIMIT or SKIP in a WITH clause returned rows the WITH never produced, and submitted a fix that makes LIMIT and SKIP barriers to filter pushdown, with flow and unit tests and a parity check against Neo4j. A maintainer has reproduced the bug on both engines.
- Earlier reported engine edge cases to the team; two are fixed upstream and one is covered by an open PR.

GNOMON · event-sourced operating system for AI agents, inside FalkorDB — Oct 2026

Python, FalkorDB, MCP · WeMakeDevs GraphHacks · selected for the XERJ Corpus Hub · repo

- Agents never write to the graph directly: they propose rewrite rules pinned by hash, which run atomically inside the database, append a hash-chained record, or return a named rejection such as stale read or failed guard.
- Bitemporal facts answer "what did we believe, and when" as a query; races between rules are found before they run through declared read and write keys and lints.

AI AGENT SYSTEMS

KINTSUGI-Ω · immune system for fleets of AI agents — Oct 2026

Python · #neatHack 2026, continued as a Master Edition · neatlogs, Entire, cfo.ai · repo

- Detects agent failures in production traces, reproduces them, lets competing sandboxed repair agents propose patches, and ships one only after eight verification gates, then watches the canary with a sequential test and reverts on harm.
- Turns solved incidents into reusable "antibodies", making repeat repairs about 7× cheaper; a 14-check cryptographic audit backs every claim.

TENURE · earned autonomy for customer-resolution agents — Oct 2026

Python · AMD Developer Hackathon · Gemma-4-31B on AMD Token Factory · live dashboard

- A refund agent gains autonomy one case type and order band at a time, only when statistics show it matches the human team; it agreed on 482 of 500 held-out cases (96.4%, 95% CI 94.3–97.8%).
- CUSUM drift detection demotes cells automatically, and every decision lands in a SHA-256 hash-chained ledger that can be re-verified offline.

AXIOM · zero-trust deploy gate for AI coding agents on AWS Oct 2026

AWS CDK, Lambda, API Gateway, CloudFormation, STS, IAM Access Analyzer, DynamoDB · AWS Builder Center "Zero to Shipped" · project page

- Agents hold read-only credentials; each change set passes 15 deterministic rules and IAM Access Analyzer before STS mints a 15-minute credential for that one change set. No LLM in the security path; fails closed.
- Shipped in under 24 hours: 12 of 12 live attack scenarios blocked, 20 of 20 unit tests passing.

HEARTHWARD · scam defence for people who live alone Oct 2026

Python, AWS CDK, AWS KMS, Alexa+ · deployed to AWS · repo

- Joins door camera, room audio and Alexa+ requests into one bitemporal, hash-chained record and decides with Wald's sequential probability ratio test, so the error budget is explicit; evidence is signed with KMS ECDSA P-256.
- Anything that lowers protection needs a family member's passkey approval; a voice in the room can never do it alone.

GAME TECH, 3D AND GEOSPATIAL

GOLEM · AI-generated 3D props to physics-ready game objects Oct 2026

1st place, Game Tech track, Cambridge × Arcade AI Hackathon (solo) · Python, Blender, Unity 6 PhysX, glTF · play

- Turns single-mesh AI props into articulated game objects (hinges, sliders, mass from volume) in 4.1–4.9 s each; a headless Unity PhysX self-test drives and verifies every joint.
- Follow-up orchestrator: code enumerates 54 joint candidates per part group and sweep-tests each; the vision model may only choose among code-verified options, and every run writes a hashed proof record.

KENSHO · verification gates for AI-written 3D game content Nov 2026

TypeScript monorepo · repo

- The model proposes scenes and assets; deterministic checks and a headless render smoke test decide what passes. 49 of 57 schema-valid scenes from Gemini failed at least one of ten checks.

CASCADE · serverless glacial-lake flood early warning on AWS Oct 2026

WeMakeDevs × AWS Environmental Hacks (solo) · AWS SAM, Step Functions, Lambda, S3, DynamoDB, SNS, Python, GDAL

- Measures lake area from Sentinel-2 imagery and routes the outburst flood 66 km down a DEM-traced valley; validated on the 2023 South Lhonak disaster (lake area within 1.8% of ISRO's figure; the forecast band bracketed the observed 138-minute arrival).
- 62 of 62 cloud runs succeeded at a 7.1 s warm median and $0.00076 per run.

SYNAPSE-3D · neuro-symbolic AR drone agent for earthquake response 2026

3rd place, HackMatrix 2026, Agentic AI track · Unity 6, AR Foundation, C#, Gemini

- Gemini turns free-text field reports into a constrained zone map; an on-device C# verifier enforces the safety rules and relay connectivity. First rescue dispatch 4.2× sooner than waiting for the full survey.

SKILLS

Verification	Concurrency and invariant testing, controlled experiments, minimal reproductions, sequential statistics (SPRT, CUSUM), reproducible evidence (SHA-256 manifests)
Databases	FalkorDB, Neo4j, Cypher, LanceDB, graph data modelling, bitemporal and event-sourced design
Cloud	AWS (Lambda, Step Functions, S3, DynamoDB, CloudFormation, SAM, CDK, IAM, STS, KMS, API Gateway), Docker, GitHub Actions
Orchestration	Systemic orchestration of multi-agent pipelines (Claude Code, Gemini), multi-model architecture, agent handoffs and review gates, LLM tool calling and guardrails, MCP, deterministic model stubs for testing
Languages	Python, Rust (contributions), TypeScript, C#, Java
3D and simulation	Unity 6 (PhysX, AR Foundation, compute shaders), Blender, glTF, geospatial analysis

EDUCATION

B.Sc. Artificial Intelligence and Machine Learning 2024 – Apr 2027 (expected)

Nandha Arts and Science College (Autonomous, Bharathiar University), Erode · CGPA 7.77/10 (Semesters 1–4)

Certifications: NPTEL, Programming, Data Structures and Algorithms Using Python (2025); NPTEL, Python for Data Science (2025)