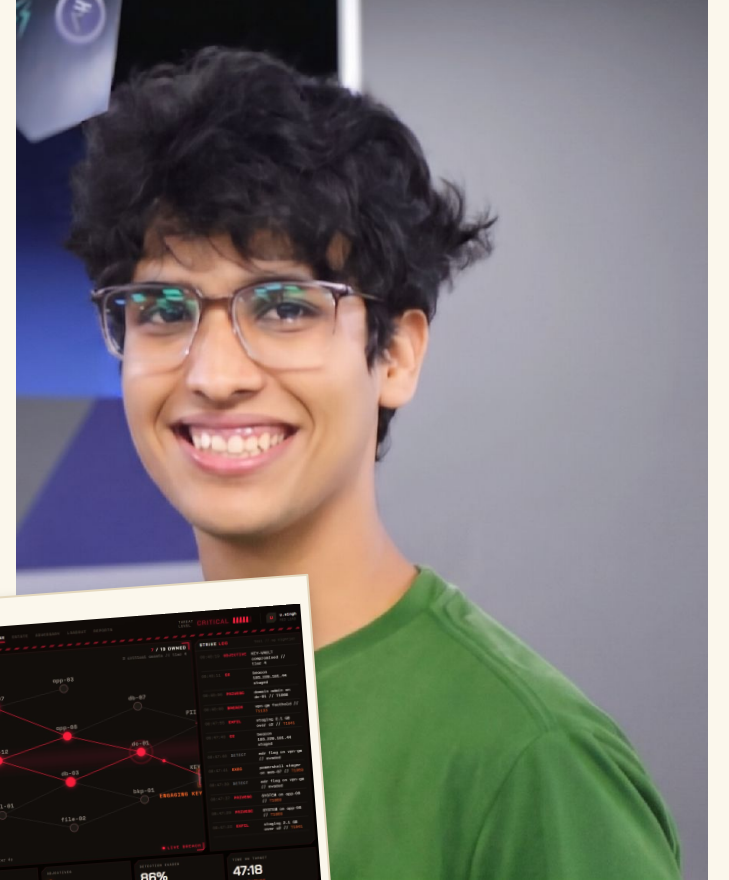


Utkarsh Bahuguna

I build things, then I break them on, **so the world can trust them.**

A developer and security researcher. He ships products, contributes to the open source tools the field is built on, and publishes independent research on how AI systems fail. He also leads a builder community and competes as a national athlete, the source of the discipline behind the work.

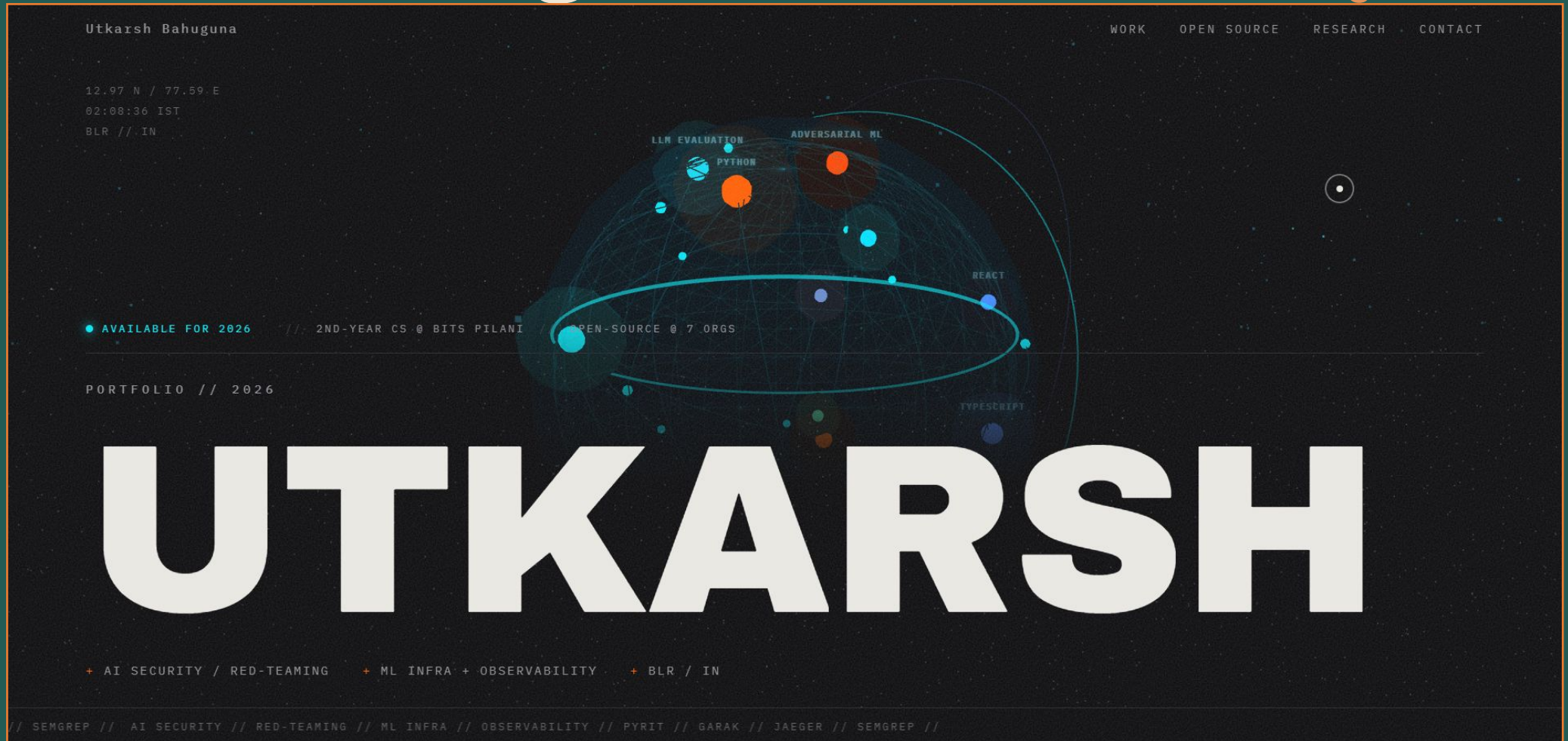
Meta × PyTorch Finalist · NVIDIA garak · Microsoft PyRIT · Google Gen AI Academy



I build things. This is where they live



[click here](#)

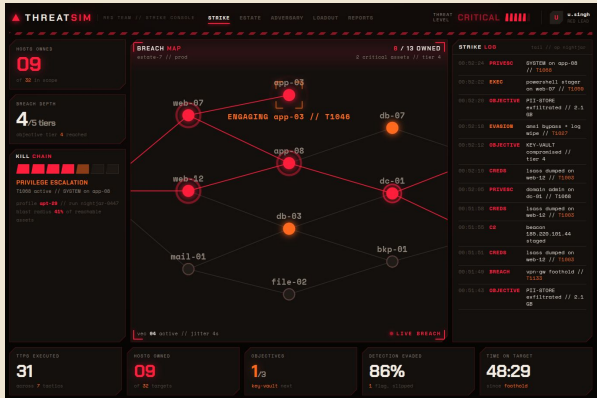


My portfolio is itself a product I designed and shipped. Every project, paper, and contribution on it started with a question I could not stop thinking about.

It always starts with one question. How does this fail?

I build systems so I can understand how they break. That instinct runs from my first security project through hackathons, open source, and research.

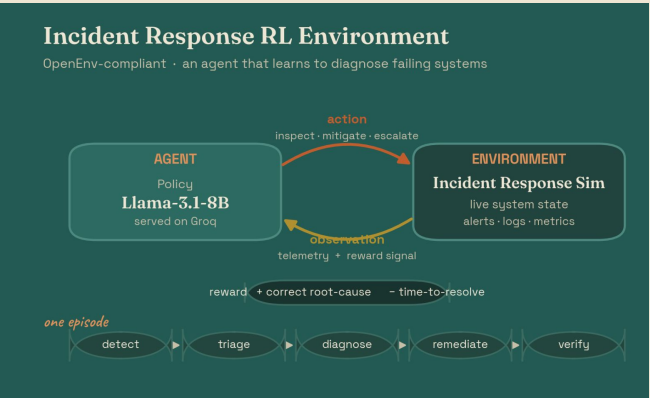
1



ThreatSim
A threat-simulation platform

Acknowledged by ISRO

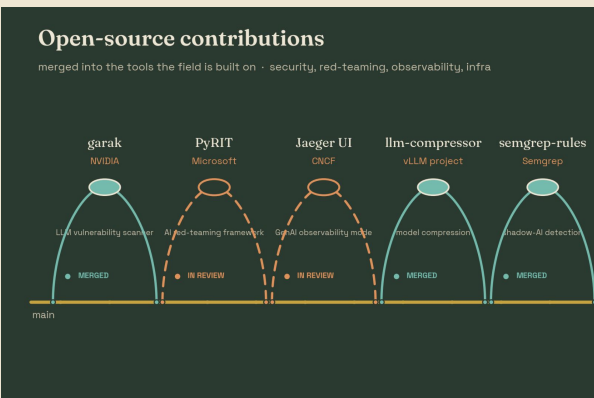
2



Hackathons
An RL agent that diagnoses failing systems

Solo finalist, global field

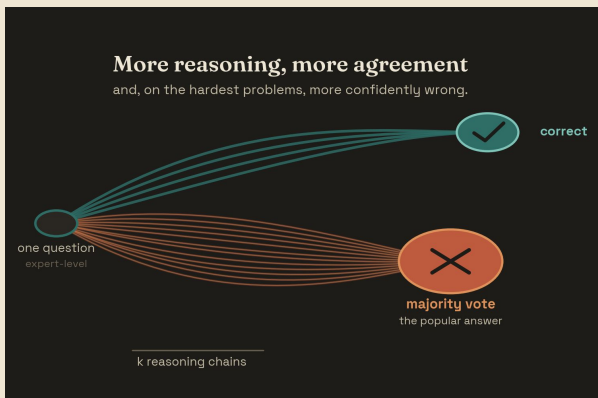
3



Open Source
Security and AI infrastructure

Merged across major projects

4



Research
Testing what the field assumes

Independent, sole author

I contribute to the tools the field is built on.

Security scanners, AI red-teaming frameworks, observability platforms. I find the gaps, ship the fixes, and add features other engineers depend on.

G

garak

NVIDIA

Added generator support and parameter handling to the LLM vulnerability scanner.

• MERGED

P

PyRIT

Microsoft

Implemented a CodeAttack jailbreak converter and attack from recent research.

• MERGED

J

Jaeger UI

CNCF

Built a GenAI observability mode with OpenTelemetry convention detection.

• MERGED

L

llm-compressor

vLLM project

Modernized type hints across the codebase to improve maintainability.

• MERGED

S

semgrep-rules

Semgrep

Authored shadow-AI detection rules for local and hosted LLM runtimes.

• MERGED

A

AI Village

DEF CON

Shipped the community and sponsors pages for the AI Village website.

• MERGED

Can making AI think more actually make it worse?

The field assumed the answer was no. I wanted evidence, so I ran a pre-registered study.

Yes.

On the hardest reasoning problems, the popular technique often performed worse.

Independent Research · Sole Author · Pre-Registered Study

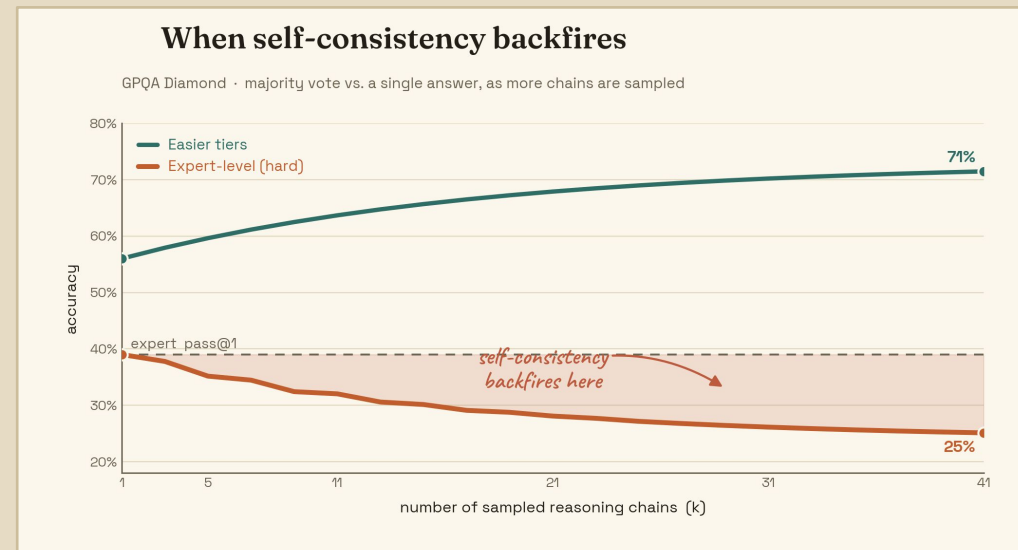


Fig. 01 · GPQA Diamond, accuracy under self-consistency

PAPER · PREPRINT

When Self-Consistency Backfires

Independent · Sole Author · Pre-Registered



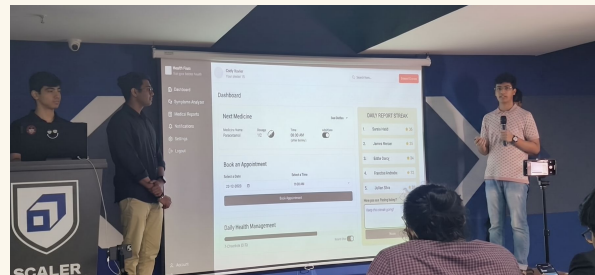
Fig. 01 · VYOM



featured



podcast



hackathon

06 LEADERSHIP

“I didn't want a title. I wanted a room where people could build.”

VYOM is my attempt to remove the luck from finding the right room.

12

MEMBERS

9

EVENTS

200+

PARTICIPANTS

I run a community for builders, organize events, and bring people together around shipping real things. The best opportunities in my life came from being in the right room. I wanted to build that room for others.

Where the habits get practiced.

I compete as a national-level athlete and play four instruments. Sport taught me how to lose and improve anyway. Music taught me patience. It is the same deliberate practice I bring to the work, just a different arena.

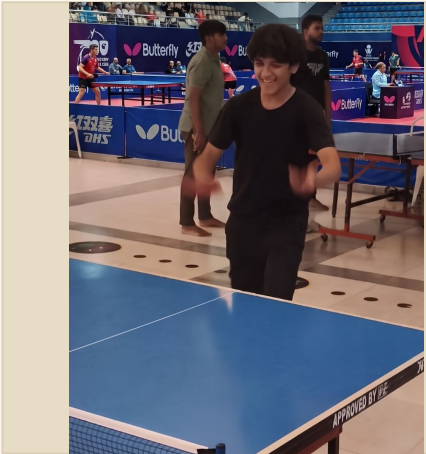


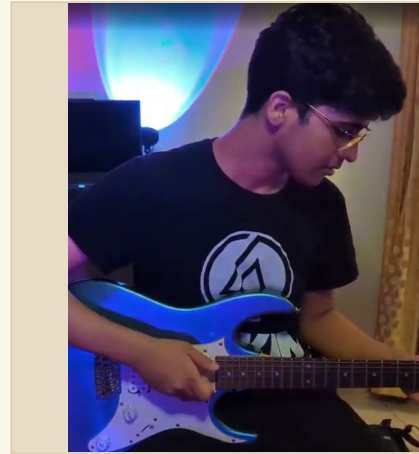
Table Tennis

national & state



Tennis

competitive



Guitar

daily practice



Studio

kalimba, keys, ukulele

In ten years, AI will run things that **matter.**

Infrastructure. Security. Decisions about real people.

If people remember one thing about me, I hope it's that I built things, then broke them on purpose, so the world could trust them.

From my first security prototype to the systems people depend on.

Still building. Still learning. Still asking questions.

· [Email](#) · [Linkedin](#) · [Github](#) · [Portfolio](#)

Utkarsh Bahuguna