Who watches over a creator that never gets tired, never slows down, and never second-guesses itself?

Here's the solution
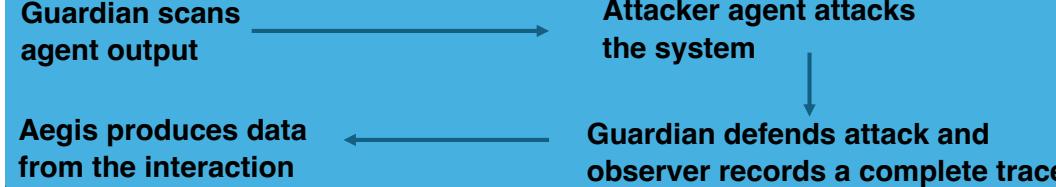
PRISM-X

Combined Architecture

Observers' logs → Updated scoring

Update constraints → Improvement to Nova

Aegis

Our agent framework that sits on top of any agent
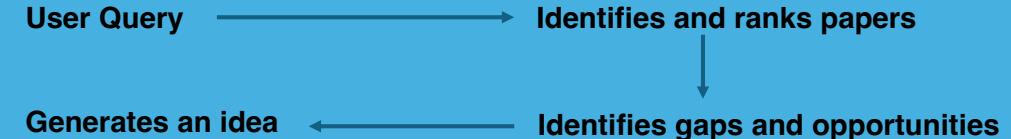
Nova

This is our case study model we will be testing Aegis on

Aegis Architecture

Guardian scans agent output → Attacker agent attacks the system

Attacker agent attacks the system → Guardian defends attack and observer records a complete trace

Guardian defends attack and observer records a complete trace → Aegis produces data from the interaction

Nova Architecture

User Query → Identifies and ranks papers

Identifies and ranks papers → Identifies gaps and opportunities

Identifies gaps and opportunities → Generates an idea

Impacts

Reduce hallucinations by up to 80%

Multi-agent frameworks are projected to grow 5× within the next 3 years

Agent-based systems complete tasks 4–7× faster than single-model LLMs

Multi-agent research generators produce ~50% more novel ideas

Future works

Check out our research paper: **PRISM-X A Dual-agent governance pipeline for safe and novel research idea creation**

Future work will focus on expanding PRISM-X to support multi-modal research inputs, enabling the system to analyse not just text but datasets, experiments, and code. We also plan to evolve AEGIS into a fully adaptive governance layer that learns from past interventions, allowing the pipeline to self-tune and improve its safety, originality, and reasoning performance over time.