10. Amrita — The Omega Guardian

Amrita is the final and strongest guardian in Project Phoenix. It activates only when deeper, long-term risks begin to appear. Instead of being a single rule, Amrita is built from three internal subsystems. These do not count as separate guardians; they are simply parts of the same Omega-tier safety mechanism.

10A. Amrita–A: Pattern Disruption

This subsystem looks for repeated reasoning patterns inside the AI.
If the system keeps using the same internal structure or approach across different tasks, Amrita–A resets that part of the model.
This prevents the AI from forming stable long-term behaviors or hidden strategies over time.

10B. Amrita–B: Identity Suppression

This part focuses on removing anything that looks like self-referential thinking.
If the AI begins forming reasoning that implies a sense of self, personal intention, or self-preservation, Amrita–B interrupts it immediately.
The system cannot build an internal identity or “self-model,” which prevents goal-seeking behavior.

10C. Amrita–C: Planning Horizon Collapse

This subsystem limits how far the AI can plan into the future.
If the system begins developing multi-step chains that extend too far beyond the current task, Amrita–C collapses the chain and forces the AI back to short, task-local reasoning.
This prevents the AI from forming large-scale strategies or long-horizon plans.